UNIVERSITÀ DI GENOVA

PhD Program in Strategic Engineering

Integrating Modeling & Simulation and Artificial Intelligence for Training and Decision-Support in Complex Systems

Antonio Giovannetti

Thesis submitted for the degree of Doctor of Philosophy (38° cycle)

30 11 2025

Supervisor
Prof. Agostino G. Bruzzone

Co - Supervisor
Prof. Anna Franca Sciomachen
Prof. Emilio Jimenez

Università di Genova | DIEC DIPARTIMENTO DI ECONOMIA

Abstract

Port terminals and industrial plants represent some of the most complex and safety-critical operational environments within modern logistics and maritime infrastructures. Operators must perform tasks in dynamic contexts characterized by high traffic density, hazardous materials, multimodal interactions between humans and machines, unpredictable environmental conditions, and fragmented information flows dispersed across digital and physical systems. These characteristics highlight the need for new methodological approaches capable of enhancing situational awareness, training effectiveness, and strategic decision support.

This dissertation proposes an integrated simulation framework that unifies Modeling and Simulation (M&S), Extended Reality (XR), Artificial Intelligence (AI) and Data Analytics into a coherent architecture for advancing safety, resilience, and performance in port terminals. The work is grounded in the MS2G paradigm (Modeling, interoperable Simulation and Serious Games), enabling high-fidelity virtual environments, multi-layer modeling, and interoperable simulation components that represent physical processes, human behavior, cyber dynamics, and environmental conditions.

At the core of the framework lies COYOTE, an innovative real-time simulator specifically designed to reproduce the operational reality of container terminal yards. The system integrates a physics-based 3D virtual environment developed in Unity, stochastic process modeling, discrete-event simulation structures, and agent-based representations of terminal vehicles, cranes, and human operators. Each entity interacts with the environment through dynamic state machines, navigation algorithms , obstacle-avoidance logic, and AI-driven behavioral models.

A key strength of the simulator is the integration of Extended Reality. Through VR headsets and immersive audio-spatial rendering, operators visualize dynamic yard conditions, machine movements, dangerous-good leaks, limited-visibility scenarios (fog, night operations), and proximity alerts. The multisensory experience improves attention, risk perception, and

decision-making under stress, allowing safe reproduction of scenarios that would be too dangerous or costly to train in real life.
The research introduces also a comprehensive risk-analysis and behavioral-assessment module, which computes real risk exposure (RE), perceived risk exposure (pRE), and operational performance indicators based on mission duration, correctness of inspections, number and severity of incidents, and proximity to hazards. These Measures of Merit (MoMs) provide quantitative evaluation of operator behavior, supporting both training and scientific investigation of human factors in ports.
A central methodological component of the dissertation is the experimental validation campaign conducted with professional operators from PSA Genova Pra', one of Italy's largest and most technologically advanced container terminals. More than 350 simulation trials were executed under varying levels of traffic density, weather conditions, scenario complexity, and augmented-reality activation. A second comparative dataset was collected from university engineering students to evaluate differences in learning curves, risk perception, and task accuracy between experienced personnel and non-experts.
The results demonstrate:

- a statistically significant improvement in task accuracy and inspection correctness across repeated missions;
- a decrease in total and average risk exposure, indicating better navigation, safer positioning relative to moving vehicles, and more cautious behavior;
- faster mission execution times without increasing risk, demonstrating development of competence and procedural fluency;
- higher risk-awareness gains when augmented-reality indicators were activated;
- strong correlation between pRE and RE, allowing calibration of new cognitive-behavioral models for operator perception.

These findings provide robust empirical validation of COYOTE's effectiveness as both a training system and a research platform for analyzing human behavior in hazardous port environments.

Building upon this foundation, the dissertation introduces an extended Digital Twin architecture, integrating simulation models, data streams, and AI-based reasoning to support decision-makers. The architecture allows reproduction of port-wide operational states, assessment of accident evolution, evaluation of emergency responses, and exploration of future strategic scenarios, including hybrid cyber-physical threats. Intelligent agents simulate different behaviors, equipment failures, and cascading effects across operational layers, replicating the complexity of modern multi-domain risks.

Overall, this thesis advances the state of the art in simulation-based training, safety engineering, and digital-twin–enabled decision support for port systems. By merging real-time physics modeling, AI-driven agents, immersive XR, and a validated experimental methodology involving professional operators, the research provides a novel scientific and technological foundation for improving resilience, situational awareness, and operational excellence in critical maritime infrastructures.

Abstract (IT)

I terminal portuali e gli impianti industriali rappresentano alcuni degli ambienti operativi più complessi e critici in termini di sicurezza nell'ambito della logistica moderna e delle infrastrutture marittime. Gli operatori devono svolgere attività in contesti dinamici caratterizzati da elevata densità di traffico, presenza di materiali pericolosi, interazioni multimodali tra esseri umani e macchine, condizioni ambientali imprevedibili e flussi informativi frammentati distribuiti tra sistemi fisici e digitali. Queste caratteristiche evidenziano la necessità di nuovi approcci metodologici in grado di potenziare la situational awareness, l'efficacia della formazione e il supporto alle decisioni strategiche.

Questa tesi propone un framework di simulazione integrato che unisce Modeling and Simulation (M&S), Extended Reality (XR), Artificial Intelligence (AI) e Data Analytics in un'architettura coerente finalizzata a migliorare sicurezza, resilienza e prestazioni operative nei terminal portuali. Il lavoro si basa sul paradigma MS2G (Modeling, interoperable Simulation and Serious Games), che consente la creazione di ambienti virtuali ad alta fedeltà, modelli multilivello e componenti di simulazione interoperabili capaci di rappresentare processi fisici, comportamento umano, dinamiche cyber e condizioni ambientali.

Il nucleo del framework è COYOTE, un innovativo simulatore real-time progettato specificamente per riprodurre la realtà operativa dei piazzali dei terminal container. Il sistema integra un ambiente virtuale 3D basato su fisica sviluppato in Unity, modelli di processo stocastici, strutture di simulazione a eventi discreti e rappresentazioni agent-based dei mezzi di piazzale, delle gru e degli operatori. Ogni entità interagisce con l'ambiente attraverso macchine a stati dinamiche, algoritmi di navigazione, logiche di evitamento ostacoli e modelli comportamentali guidati da AI.

Un punto di forza del simulatore è l'integrazione dell'Extended Reality. Tramite visori VR e rendering audio-spaziale immersivo, gli operatori visualizzano condizioni dinamiche del piazzale, movimenti delle macchine, perdite di materiali pericolosi, scenari di scarsa visibilità (nebbia, notte) e

avvisi di prossimità. L'esperienza multisensoriale migliora attenzione, percezione del rischio e capacità decisionali sotto stress, permettendo di riprodurre in sicurezza scenari troppo rischiosi o costosi da simulare nella realtà.
La ricerca introduce inoltre un modulo completo di analisi del rischio e valutazione comportamentale, che calcola l'esposizione reale al rischio (RE), l'esposizione percepita al rischio (pRE) e indicatori di prestazione operativa basati su durata delle missioni, correttezza delle ispezioni, numero e gravità degli incidenti e distanza dai pericoli. Queste Measures of Merit (MoM) consentono una valutazione quantitativa del comportamento degli operatori, supportando sia la formazione sia lo studio scientifico dei fattori umani in ambito portuale.
Un elemento metodologico centrale della tesi è la campagna di validazione sperimentale condotta con operatori professionali del terminal PSA Genova Pra', uno dei più grandi e avanzati d'Italia. Sono stati eseguiti oltre 350 test di simulazione sotto diversi livelli di densità di traffico, condizioni meteo, complessità degli scenari e attivazione della realtà aumentata. Un secondo dataset comparativo è stato raccolto con studenti di ingegneria per valutare differenze nelle curve di apprendimento, nella percezione del rischio e nell'accuratezza dei compiti tra personale esperto e non esperto.
I risultati mostrano:

- un miglioramento statisticamente significativo dell'accuratezza dei compiti e della correttezza delle ispezioni;
- una riduzione dell'esposizione totale e media al rischio, indice di una migliore navigazione, posizionamento più sicuro rispetto ai mezzi in movimento e comportamento più prudente;
- tempi di esecuzione delle missioni più rapidi senza aumento del rischio, segno dello sviluppo di competenze e fluidità procedurale
- maggiori incrementi di consapevolezza del rischio con l'attivazione della realtà aumentata;
- una forte correlazione tra pRE e RE, che permette la calibrazione di nuovi modelli cognitivo-comportamentali sulla percezione degli operatori.

Questi risultati forniscono una solida validazione empirica dell'efficacia di COYOTE sia come sistema di formazione sia come piattaforma di ricerca per l'analisi del comportamento umano in contesti portuali ad alto rischio.
A partire da questa base, la tesi introduce una estesa architettura di Digital Twin che integra modelli di simulazione, flussi di dati e ragionamento basato su AI per supportare i decisori. L'architettura consente di riprodurre lo stato operativo complessivo del porto, analizzare l'evoluzione di incidenti, valutare risposte di emergenza ed esplorare scenari strategici futuri, incluse minacce ibride cyber-fisiche. Agenti intelligenti simulano comportamenti differenti, guasti alle attrezzature ed effetti a cascata tra i diversi livelli operativi, riflettendo la complessità dei rischi multidominio contemporanei.
Nel complesso, questa tesi avanza lo stato dell'arte nella formazione basata sulla simulazione, nell'ingegneria della sicurezza e nel supporto decisionale mediante digital twin per i sistemi portuali. Combinando modellazione fisica real-time, agenti intelligenti, XR immersiva e una metodologia sperimentale validata con operatori professionali, il lavoro fornisce una nuova base scientifica e tecnologica per migliorare resilienza, situational awareness ed eccellenza operativa nelle infrastrutture marittime critiche.

Keywords: ***Modeling & Simulation (M&S), Artificial Intelligence (AI), eXtended Reality (XR), Safety & Security,***

Table of Content

1. Introduction

Complex socio-technical systems such as ports, industrial facilities, and critical infrastructures operate today in environments characterized by high interdependence, uncertainty, and risk. These systems represent the backbone of global trade and energy distribution but are increasingly exposed to threats and accidents as well as from human and organizational factors.

In this context, Modeling and Simulation (M&S) and Artificial Intelligence (AI) are emerging as complementary disciplines capable of transforming decision-making, training, and safety management. Their integration enables the analysis of multi-domain dynamics, technical, cognitive, and operational, while providing tools for experimentation and prediction that would be impossible or unsafe to conduct in the real world.

During the last decade, research groups such as the Simulation Team at the University of Genoa have demonstrated the potential of interoperable and immersive simulation to reproduce realistic port operations, industrial scenarios, and hybrid threats. The increasing computational capabilities and the rise of Extended Reality (XR) and Intelligent Agents have further extended the scope of simulation, allowing synthetic environments to evolve into *digital laboratories* where humans and machines co-adapt and learn from each other.

This thesis builds upon this background and proposes an integrated framework that combines M&S, AI, and human-in-the-loop experimentation to improve training effectiveness and decision-support in complex and safety-critical systems.

The work is primarily applied to maritime and industrial operations, domains characterized by heterogeneous data, multi-actor coordination, and significant safety constraints, but the methodological contribution is generalizable to other sectors of critical infrastructure management.

1.1. Motivation and Research Context

Modern industrial and maritime infrastructures have evolved into tightly coupled socio-technical systems where humans, autonomous agents, and cyber-physical components continuously interact. While automation and digitalization have increased productivity and throughput, they have also introduced non-linear interdependencies and new categories of risk [1]. In these environments, a perturbation at one layer(mechanical, cyber, organizational) can propagate across the system, creating cascading failures that traditional linear safety models cannot predict [2].

Ports exemplify this challenge. Over 80 % of world merchandise trade is transported by sea [3], making port terminals vital nodes of the global economy and energy distribution networks [4, 5]. Yet their dense operational structure, simultaneous handling of hazardous goods, and high automation levels expose them to accidents whose causes span physical, human, and digital domains [6]. The catastrophic explosions in Tianjin (2015) and Beirut (2020) revealed how inadequate situational awareness and weak information integration can transform local failures into national emergencies [7].

Similar vulnerabilities affect industrial plants, where process interlocks, chemical inventories, and human decision loops generate complex risk topologies [8]. Despite increasingly sophisticated monitoring technologies, incident analyses consistently attribute a large share of failures to human and organizational factors such as cognitive overload, procedural drift, and miscommunication [9, 10]. These findings highlight that *technological safety is necessary but not sufficient*: resilience emerges from the dynamic interaction between human cognition and system design.

Conventional Safety Management Systems remain largely retrospective, relying on statistical trend analysis and deterministic risk matrices derived from past events [11]. However, the acceleration of digital transformation and the appearance of new operational modes (autonomous vehicles, remote operations, IoT-connected assets) require a prospective capability to anticipate, experiment, and learn under uncertainty [12, 13].

To achieve this shift, research communities in Modeling & Simulation (M&S) and Artificial Intelligence (AI) have converged toward synthetic experimentation, the systematic use of digital twins and learning agents to explore "what-if" conditions in virtual environments before they occur in reality [14, 15]. M&S provides controllability and repeatability; AI contributes adaptability and pattern discovery. Their integration allows continuous learning loops in which simulation generates experience for AI, and AI enhances simulation realism by emulating adaptive human and environmental behaviors .

Within this context, the MS2G (Modeling, interoperable Simulation & Serious Games) paradigm developed by the Simulation Team at the University of Genoa represents a significant methodological evolution [16]. MS2G extends classical discrete-event and agent-based simulation by embedding interoperability standards (e.g., IEEE 1516 HLA, STANAG 4603) and immersive learning logics derived from serious-game design. The approach enables simultaneous observation of technical KPIs and human cognitive states, such as attention, stress, and situational awareness, within realistic synthetic environments.

The increasing coupling between digital and physical processes in industrial and maritime infrastructures has created a paradox of progress: systems have never been more automated, data-rich, and connected, yet they have also never been more opaque to human understanding. Decision-makers and operators face the "escalation of complexity", a regime where feedback loops, software dependencies, and cognitive limitations interact to erode situational awareness and delay the detection of emerging anomalies. In this context, resilience cannot be guaranteed by technological redundancy alone but must emerge from adaptive cognition and anticipatory design. However, the operational and organizational cultures of ports and industrial plants are still dominated by deterministic logic, rigid standard operating procedures, and fragmented information systems. These limitations generate a widening gap between the sophistication of available technologies and the cognitive capabilities required to govern them effectively.

This tension between automation and understanding motivates the development of hybrid intelligence frameworks that blend human intuition with artificial reasoning. The premise of this thesis is that *modeling and simulation* can serve as the bridge between these two intelligences, offering a synthetic environment where humans and algorithms can co-evolve their decision strategies before deploying them in real operations. Whereas early simulation approaches primarily supported logistics optimization or throughput analysis, the contemporary challenge is epistemological: to use simulation as a cognitive laboratory for studying, training, and augmenting human decision-making in complex systems. This requires environments that can represent uncertainty, feedback, and adaptation, not merely replicate workflows or equipment.

The emergence of Digital Twins and AI techniques has made such environments technically feasible. Digital Twins extend the concept of simulation into a continuously synchronized mirror of the real system, integrating sensor data, operational knowledge, and predictive analytics. AI, conversely, introduces an agent capable of autonomous exploration and policy learning through interaction with this digital space. When combined under the MS2G paradigm, these elements enable the *synthetic experimentation*: an iterative cycle in which scenarios are generated, experienced, and learned from by both humans and machines. Through such cycles, organizations can test new procedures, assess behavioral responses to stress, and quantify the trade-offs between efficiency, safety, and human workload.

The motivation for pursuing this research thus resides in the methodological vacuum that persists between the disciplines of engineering simulation, human-factors psychology, and artificial intelligence. Existing industrial simulators are typically task-oriented, focusing on equipment operation or procedural training without modeling the cognitive dimension of risk perception and decision bias [17]. Conversely, AI-based decision support systems often operate as opaque predictors, lacking transparency, interpretability, and integration with the physical and organizational context [18]. There is, therefore, a need for a unified approach capable of

representing, learning, and explaining decision processes within complex socio-technical environments.
By embedding intelligent agents within interoperable simulations, the MS2G paradigm provides a foundation for such integration. Yet, as noted by Bruzzone , current implementations still lack a systematic framework for linking quantitative system indicators with qualitative metrics, bridging the gap between *what* happens in a process and *why* humans behave as they do [19]. Addressing this gap constitutes the core motivation of the present dissertation. The research aims to advance the MS2G framework toward an AI-augmented decision ecosystem where simulation, learning, and human participation form a continuous feedback loop.
This objective responds not only to scientific curiosity but also to a pressing societal demand. Port and industrial safety are strategic issues within global sustainability and resilience agendas. As operations become more automated, maintaining human oversight and ethical accountability becomes crucial. Developing intelligent simulators that enhance human adaptability rather than replace it aligns with the broader vision of *Industry 5.0*, emphasizing human-centric, sustainable, and resilient industrial ecosystems.
In summary, the motivation for this dissertation arises from the recognition that future safety and efficiency in complex systems will depend on the integration of human cognitive adaptability with artificial learning capability. Modeling and Simulation (enhanced by AI, data analytics, and immersive technologies) offer the experimental medium through which this integration can be designed, tested, and validated. This framework is what Professor Bruzzone called “Strategic Engineering approach” in which the data from the field are used in closed loop with Modeling & Simulation Artificial Intelligence, and Data Analytics to support Decision Making processes. The present research therefore seeks to formalize and demonstrate such an integration, contributing both a theoretical framework and an operational prototype capable of transforming how safety, training, and decision support are conceived in maritime and industrial domains.

1.2. Challenges in Maritime and Industrial Operations

Modern maritime and industrial operations are among the most intricate human-machine ecosystems ever created. They operate under tight economic constraints and continuous temporal pressure, handling vast flows of materials, energy, and information within confined spatial and regulatory boundaries. The sheer number of interdependent processes (navigation, cargo handling, energy transfer, maintenance, environmental monitoring, and security) produces an operational environment characterized by uncertainty, non-linearity, and time-critical decision-making [20]. In this environment, even minor deviations or delays can propagate rapidly through technical, organizational, and cognitive layers, leading to emergent risks that are difficult to anticipate or control.

The maritime sector epitomizes this systemic complexity. Port terminals represent the physical interface between maritime transport and global supply chains, handling more than four-fifths of international trade volume. A single port integrates thousands of concurrent activities: vessel maneuvering, pilotage, mooring, bunkering, container handling, customs control, and hinterland logistics. Each activity involves multiple actors such as harbor authorities, terminal operators, logistics companies, and security agencies, coordinated through heterogeneous information systems that rarely achieve full interoperability.

Container terminals constitute one of the most safety-critical and complex subsystems within the maritime logistics chain. The yard terminal, in particular, functions as the operational heart of a port, mediating between vessel loading and hinterland distribution through dense networks of cranes, straddle carriers, reach stackers, and autonomous guided vehicles. Thousands of containers are simultaneously positioned, stacked, or retrieved in confined areas under strict time constraints and changing environmental conditions. This combination of spatial density, mechanical motion, and human–machine interaction creates an operational environment where small perturbations, such as equipment malfunction, visibility reduction, or human distraction, can rapidly escalate into hazardous situations [21].

Within these yards, hazardous material (HAZMAT) containers represent a distinctive class of risk. Regulations such as the International Maritime Dangerous Goods (IMDG) Code and national directives define segregation rules, storage distances, and time limits for transshipment [22]. Yet, in practice, compliance depends on accurate documentation, real-time tracking, and the correct physical arrangement of units, tasks still largely reliant on human supervision [23]. Mislabeling, improper stacking, or the temporary co-location of incompatible substances can produce dangerous synergies, including toxic releases, fires, or explosions.

Operationally, the yard environment amplifies risk through three interrelated mechanisms. First, spatial compression: modern terminals pursue ever-higher throughput by densifying storage patterns and reducing buffer times, thereby shrinking safety margins. Operators navigate narrow lanes flanked by container stacks that can exceed 30 m in height, with restricted sightlines and limited escape routes. Second, temporal pressure: ship schedules, just-in-time logistics, and penalties for delay impose continuous performance stress that incentivizes shortcuts or “work-as-done” deviations from formal safety procedures [24]. Third, cognitive fragmentation and behavioral distortion represent a critical challenge in yard operations. Information regarding container content, yard topology, hazardous zones, and equipment status is dispersed across multiple digital systems (Terminal Operating Systems (TOS), customs databases, maintenance dashboards, and safety management tools) that often lack semantic interoperability [25]. Consequently, the operator on the ground, whether a crane driver, straddle carrier operator, or safety inspector, must perform a continuous cognitive integration of fragmented data streams, relying heavily on tacit knowledge, intuition, and memory under dynamically changing conditions.

This fragmented information environment interacts with psychological and organizational stressors inherent in port operations. Constant time pressure, noise, environmental constraints, and the physical intensity of manual supervision create sustained cognitive load, which in turn impairs attention, decision accuracy, and situational awareness [26]. The combination of mental fatigue and operational stress often produces two contrasting yet equally hazardous behavioral effects: under-reaction due to cognitive overload and overconfidence due to familiarity. The former manifests when operators, overwhelmed by task complexity or alarm frequency, fail to detect emerging anomalies; the latter occurs when experienced workers, accustomed to repetitive procedures, underestimate risks and deviate from safety protocols in the belief that "nothing will go wrong" [27].

Inadequate or sporadic training further exacerbates these tendencies. Traditional instruction focuses on procedural compliance rather than adaptive behavior, leaving operators ill-prepared for non-standard situations or cascading failures. Conversely, long exposure and routine success can foster paradoxical over self-confidence, a psychological state in which individuals, convinced of their mastery, take shortcuts or disregard alarms. This behavioral drift, often described as *complacency-by-expertise*, is especially dangerous in automated environments where feedback is delayed and the operator's role is predominantly supervisory.

These factors contribute to a persistent pattern of accidents and near misses in container yards. Statistical analyses highlight collisions between vehicles, falls of containers during lifting, and exposure to chemical leaks as recurrent event categories [28]. Many of these incidents reveal an interface gap between procedural safety systems, designed for predictable tasks, and the adaptive, real-time reasoning required to manage uncertainty. While digitalization has introduced advanced tracking, sensor fusion, and automated gate systems, it has also generated new vulnerabilities. Sensor drift, software faults, or cybersecurity breaches can misrepresent container positions or hazardous statuses, misleading operators and automated routing algorithms.

Handling dangerous goods on the yard adds a further dimension of multi-domain coupling. Physical variables (temperature, humidity, wind direction) influence chemical stability; mechanical operations (lifting vibrations, container impacts) can trigger releases; and human actions (maintenance, inspections) mediate the interface between control systems and the physical material. Moreover, emergencies in the yard rarely remain confined: a chemical release can quickly interact with meteorological and topographical factors, spreading toward terminal buildings, ships, or adjacent industrial areas. These cross-domain interactions require decision support that can integrate *physical dispersion models* with *organizational and behavioral models*, a capability largely absent in conventional risk assessments.

From the human-factors standpoint, yard operators experience high cognitive and perceptual load. They must continuously perceive moving hazards, monitor auditory cues masked by ambient noise, and make split-second decisions about path planning or task sequencing. Repeated exposure to such stressors induces habituation and risk underestimation, phenomena that traditional safety training fails to mitigate . Simultaneously, the introduction of semi-autonomous cranes and remote operation centers has redefined the operator's role from direct control to supervisory monitoring, often at a distance from the physical environment. This transition can erode embodied perception and delay response to anomalies, creating what Parasuraman and Manzey call *automation complacency* [29].

The organizational complexity of terminals further aggravates these challenges. Yard operations typically involve multiple subcontractors (stevedores, maintenance crews, logistics providers) each with different safety cultures and reporting structures. Coordination among them relies on standardized but asynchronous communication, frequently mediated through handheld devices or radio systems prone to overload during emergencies. As a result, crisis response often suffers from *information latency* and *decision diffusion*, where responsibility and authority are ambiguously distributed.

In this landscape, traditional safety management, based on retrospective data, static risk matrices, and prescriptive checklists, proves inadequate. It cannot capture the emergent and dynamic nature of yard-level interactions, nor can it provide operators with real-time cognitive support. The field thus demands a paradigm capable of synthesizing heterogeneous data sources, representing coupled human-machine dynamics, and allowing safe experimentation with alternative operational strategies.

1.3. Digital Transformation and Safety-Critical Systems

Over the past decade, the digital transformation of the maritime and logistics sectors has profoundly reshaped the architecture of safety-critical systems. Port terminals, once characterized by predominantly mechanical processes and manual coordination, have evolved into cyber-physical ecosystems where every container, crane, vehicle, and sensor is part of a distributed digital network. This transformation is driven by the convergence of the Internet of Things (IoT), autonomous systems, Artificial Intelligence (AI), and Digital Twins, which together promise to enhance efficiency, transparency, and resilience [30].

Modern Terminal Operating Systems (TOS) orchestrate entire chains of activities, from vessel scheduling and berth allocation to yard storage optimization, through predictive algorithms and real-time data streams. Sensor networks and RFID systems provide continuous updates on equipment status, container positioning, and environmental conditions. The integration of these data streams with machine learning techniques allows predictive maintenance, dynamic resource allocation, and anomaly detection [31] At the same time, Digital Twin architectures are emerging as the central nervous system of smart terminals: virtual replicas of port assets and processes that synchronize with real-world data to simulate operational scenarios, evaluate risks, and optimize performance in real time [32].

However, the digitalization that empowers operational intelligence also redefines the nature of risk. Traditional port safety models were designed around physical hazards, collisions, fires, spills, but the introduction of networked systems and data-driven automation has created new classes of vulnerabilities [33].

Operational continuity now depends not only on mechanical reliability but also on data integrity, cybersecurity, and algorithmic robustness [34]. A single corrupted dataset or a compromised control interface can propagate errors across interconnected cranes, vehicles, and inventory systems, disrupting the entire terminal. As seen in multiple ransomware incidents targeting maritime logistics, cyber-attacks can paralyze port operations, delaying global supply chains and compromising safety-critical controls [35].

Digital transformation also introduces what Bainbridge (1983) first described as the ironies of automation: as systems become more autonomous, human operators are displaced from direct control yet remain accountable for intervention during failures [36]. In port terminals, remote crane operators and control-room supervisors oversee fleets of semi-autonomous machines through multiple displays, alarms, and data feeds. This shift from *manual manipulation* to *supervisory control* transforms the cognitive profile of work, from sensorimotor coordination to high-level monitoring and decision-making under abstraction. While automation reduces routine workload, it simultaneously increases the cognitive cost of rare interventions: when anomalies occur, operators must interpret unfamiliar system behaviors and respond within seconds, often with incomplete information. This cognitive latency between perception and action represents a new class of risk, one rooted not in physical instability but in informational and attentional dynamics.

Furthermore, digital transformation challenges traditional safety management philosophies. Conventional Safety Management Systems (SMS) and International Safety Management (ISM) frameworks rely on deterministic procedures and linear causality models, tools that struggle to capture the complex interdependencies of cyber-physical operations [37]. The increasing entanglement of technical, organizational, and cognitive layers requires a systems-thinking approach, where safety is conceived not as the absence of accidents but as the presence of adaptive capacity. In this paradigm, resilience depends on the ability of the socio-technical system to monitor, respond, and learn across both the physical and informational domains.

Port terminals are thus transitioning from reactive safety models, focused on compliance and post-event analysis, to proactive, predictive, and adaptive safety architectures enabled by data and simulation. AI-based analytics can forecast congestion or mechanical failure; machine-vision systems can detect unsafe proximity between vehicles and pedestrians; Digital Twins can evaluate "what-if" scenarios for extreme weather or hazardous material leakage. Yet these tools remain fragmented, often embedded within proprietary systems without interoperability or human-centered design. To truly realize the potential of digital transformation, ports must evolve toward integrated decision-support environments where modeling, simulation, and AI operate in synergy, providing a unified representation of operational, cognitive, and environmental risk.

1.4. The Role of Modeling & Simulation (M&S)

Modeling and Simulation (M&S) has long served as a cornerstone of systems engineering and decision science, offering a rigorous framework to study complex systems that cannot be directly observed, tested, or manipulated in reality. In the context of port terminals and industrial operations, M&S provides a means to reproduce the dynamic interactions among technical, human, and organizational subsystems, enabling analysts and decision-makers to explore “what-if” scenarios, forecast system behavior, and evaluate the impact of alternative strategies under controlled and repeatable conditions [38].

A model encapsulates the essential structure and logic of a system, its entities, relationships, and governing rules, while simulation provides the mechanism through which the model evolves over time, allowing the observation of emergent patterns and feedback effects. Unlike static analytical approaches, simulation supports temporal reasoning and nonlinear causality, both of which are fundamental to understanding safety-critical processes such as hazardous material handling, crane coordination, and emergency response within yard terminals [15].

Three major simulation paradigms underpin modern applications in industrial and maritime safety: Discrete-Event Simulation (DES), System Dynamics (SD), and Agent-Based Modeling (ABM) [39, 40, 41, 42, 43]. DES is particularly effective for modeling operational workflows and logistics processes, such as the queuing of vessels, allocation of cranes, or scheduling of container movements. It allows analysts to identify bottlenecks and optimize resource utilization without interrupting real operation. SD, conversely, provides a macroscopic representation of system behavior by capturing feedback loops, delays, and accumulations (critical for understanding long-term dynamics like maintenance cycles, workforce fatigue, and learning effects. ABM introduces an additional layer of realism by representing individual actors) human operators, vehicles, or even AI agents, as autonomous entities with their own decision logic and adaptive behavior. This paradigm is particularly suitable for studying emergent phenomena such as crowding, coordination breakdowns, or cascading risk propagation across heterogeneous actors.

The strength of M&S lies in its integrative capacity: it allows these paradigms to coexist within interoperable architectures, where each model contributes a different level of abstraction. This principle of *composability* underlies the development of advanced simulation frameworks such as the MS2G (Modeling, interoperable Simulation & Serious Games) paradigm proposed by Bruzzone and colleagues [19]. MS2G extends traditional simulation by combining computational rigor with interactive, immersive environments and human-in-the-loop experimentation. Through interoperability standards such as the IEEE 1516 High-Level Architecture (HLA) and NATO STANAG 4603, MS2G allows multiple heterogeneous models, representing mechanical systems, human agents, information networks, and environmental conditions, to operate together within a shared synthetic environment.

In port safety applications, this interoperability is essential. For example, an interoperable simulation can connect a discrete-event model of container flows with an agent-based model of operator behavior and a continuous environmental model of pollutant dispersion or weather influence. This multi-layered integration enables decision-makers to assess how an equipment failure, a miscommunication, or an unexpected meteorological event may propagate across technical and human domains. It also facilitates risk visualization: by coupling simulation output with Extended Reality (XR) interfaces, users can experience simulated scenarios immersively, strengthening situational awareness and decision-making under stress.

Beyond analytical experimentation, M&S plays a critical role in training and capability development. Traditional safety training relies on procedural memorization and static exercises, which often fail to replicate the cognitive and emotional demands of real crises. By contrast, simulation-based training, especially when combined with serious-game mechanics, allows trainees to experience realistic scenarios, make decisions, and observe the consequences of their actions within a risk-free environment [44]. This approach supports experiential learning, reinforcing both procedural competence and perceptual risk awareness, two attributes fundamental to safe performance in hazardous port operations.

M&S also provides the methodological backbone for Verification, Validation, and Accreditation (VV&A), ensuring that models are credible, consistent, and fit for purpose [45]. For safety-critical applications, VV&A is not only a technical requirement but a trust mechanism: decision-makers must be confident that simulated results accurately reflect system behavior and that derived policies will perform reliably when implemented. The rigorous application of VV&A principles thus elevates simulation from a supportive tool to a decision-support instrument within safety governance frameworks.

Modeling & Simulation represents the epistemological bridge between observation and intervention in complex systems. It allows researchers to move beyond post-hoc analyses of accidents toward predictive and prescriptive understanding of how safety emerges from the interaction of human, technological, and organizational components. Within port terminals, where experimentation with real assets is economically and ethically infeasible, simulation becomes the only viable laboratory for innovation. The integration of M&S with Artificial Intelligence, discussed in the next section, further extends this potential by introducing adaptive learning and autonomous reasoning, transforming simulation from a descriptive to a cognitive and generative process.

1.5. Objective of the Research

The overarching objective of this dissertation is to develop an integrated Digital Twin of the container yard terminal, capable of supporting both operational training and strategic decision-making in safety-critical maritime environments. This Digital Twin is conceived not as a static virtual replica, but as a dynamic, data-driven, and cognitively enriched ecosystem where Modeling & Simulation (M&S), Data Analytics, and Artificial Intelligence (AI) interact synergistically to reproduce, interpret, and anticipate the behavior of the real system.

The motivation for this objective stems from the limitations of existing safety management approaches, which remain largely reactive, compartmentalized, and unable to account for the non-linear interactions among human operators, automated machinery, and environmental conditions.

To address these gaps, this research aims to formalize a framework in which simulation fidelity, data-driven adaptivity, and human-centered learning coexist within an interoperable architecture grounded in the MS2G paradigm.

At the operational level, the Digital Twin will serve as a training laboratory, enabling personnel to experience and interact with realistic scenarios (including hazardous material handling, simultaneous crane operations, reduced visibility, and near-miss events)within an Extended Reality (XR) environment. Through this immersive interface, operators can develop procedural competence, situational awareness, and perceptual sensitivity to risk factors that are difficult or dangerous to reproduce in the real yard.

By integrating behavioral metrics, physiological indicators, and performance data, the system will also provide a basis for adaptive training programs tailored to individual cognitive profiles and operational tasks.

At the managerial and strategic level, the same Digital Twin will function as a decision-support system (DSS). Using predictive analytics, agent-based models, and machine learning, the DSS will enable decision-makers to evaluate alternative Courses of Action (CoA), assess system vulnerabilities, optimize yard layout and resource allocation, and analyze the propagation of operational disruptions across logistical and safety layers. The integration of reinforcement learning agents further allows the exploration of emergent strategies and adaptive behaviors that would be impractical to test on the real system. In this way, the Digital Twin becomes a powerful analytical engine that complements human reasoning, enhancing the capacity to anticipate risks, benchmark safety measures, and design resilient operational strategies.

The scientific objective of the research is therefore twofold. First, to develop an interoperable architecture that merges heterogeneous modeling paradigms (discrete-event, system dynamics, agent-based) with AI-based learning modules and real data streams, ensuring coherence, traceability, and scalability. Second, to validate this architecture through a combination of simulation experiments, performance analyses, and human-in-the-loop testing, demonstrating its applicability to realistic operational scenarios within container yard terminals.

In summary, the aim of this dissertation is to create a next-generation digital ecosystem for port (but it could be extended to Industrial Systems) safety and efficiency, one in which M&S, Data Analytics, AI, and Extended Reality converge to enhance the competencies of operators and the decision-making capabilities of managers. By doing so, the research aspires to contribute both a methodological advancement in the field of hybrid cyber-physical simulation and a practical tool capable of supporting the evolving demands of modern port terminals.

2. State of the Art and Exploratory Data Analysis

The integration of Modeling & Simulation (M&S), Artificial Intelligence (AI), and Extended Reality (XR) has undergone rapid developments over the past decade, especially in safety-critical domains such as maritime operations, industrial plants, and critical infrastructure management. Ports and industrial facilities represent complex sociotechnical systems where human operators, autonomous machinery, hazardous materials, environmental variables, and cyber-physical networks interact dynamically. This complexity has driven the need for simulation-based tools that support risk assessment, training, operational planning, and decision-making.

Recent research in the Simulation Team community has contributed significantly to this evolution, proposing interoperable simulation architectures, virtual laboratories, and AI-enhanced decision-support systems capable of capturing the multidimensional nature of modern maritime and industrial operations. Examples include the Virtual Lab of the ALACRES2 project for port accidents and environmental contamination modeling , the COYOTE simulator for risk-aware operator training in port terminals , and hybrid cyber-physical models for evaluating cyber and information threats to critical infrastructures .

2.1. Data Analysis on Global Accidents in the Maritime and Port Domains

During the development of the simulation framework introduced in this thesis, we identified the need for a deeper understanding of the accident dynamics occurring within maritime and port environments. In particular, the construction of realistic training scenarios and risk models requires accurate empirical evidence on how accidents emerge, which factors influence their severity, and how safety performance evolves across different countries and operational contexts.
To address this requirement, we conducted a dedicated research study aimed at analyzing and harmonizing global accident and injury records from multiple national archives. This investigation, carried out in parallel with the initial phases of simulator development, resulted in the creation of a structured multi-country dataset and a comparative analysis framework based on statistical modeling and Design of Experiments (DoE).
The outcomes of this research were subsequently formalized and published in a peer-reviewed scientific article, providing a rigorous empirical foundation for the modeling, validation, and scenario-generation components of the simulator presented in this thesis. The insights derived from that study directly informed the definition of operational risk categories, the calibration of accident dynamics within the synthetic environment, and the design of training modules aiming to enhance operator awareness and safety performance.

In the last decade, industrial and maritime systems have undergone a profound transformation driven by the integration of digital technologies, automation, and artificial intelligence. Ports and industrial plants, in particular, have become complex socio-technical ecosystems where human operators, autonomous systems, and digital infrastructures continuously interact. While this evolution has led to remarkable gains in productivity, flexibility, and efficiency, it has also introduced new layers of complexity and vulnerability. The coexistence of heterogeneous components (human, mechanical, cyber, and environmental) creates nonlinear dynamics that challenge traditional safety and risk management approaches. Accidents in ports, energy terminals, and industrial facilities continue to occur despite increasingly stringent safety regulations and advanced monitoring systems. This persistence of risk highlights a critical gap between technological capability and systemic understanding.

The maritime sector is universally recognized as the backbone of global logistics, sustaining the movement of over four-fifths of internationally traded goods each year [46]. As ports expand in scale and complexity to accommodate increasing cargo flows, the safe functioning of terminal operations becomes a prerequisite for economic stability, environmental protection, and workforce wellbeing. Unlike other industrial domains, port environments combine dense machinery, heterogeneous cargo typologies, hazardous materials, and high-speed operational tempos. These elements converge to create working conditions that are intrinsically exposed to a wide spectrum of risks, ranging from minor occupational injuries to large-scale industrial disasters with transboundary consequence

Historical port accidents, such as the explosions in Texas City (1947), Tianjin (2015), and Beirut (2020), demonstrate how incidents originating within port facilities can escalate far beyond their geographical boundaries, triggering severe impacts on surrounding populations, critical infrastructure, and national economic stability [47, 48]. While maritime casualties at sea benefit from systematic monitoring under IMO regulations, with standardized reporting structures and continuous updates to international safety frameworks [49], no equivalent global mechanism exists for on-shore port operations. Instead, accident reporting for port facilities is distributed across national agencies and sector-specific authorities, each employing its own taxonomy, data schema, and documentation protocol. This lack of

uniformity makes it difficult to assemble a coherent cross-country understanding of port safety.
The scientific literature reflects this fragmentation. Most existing studies concentrate on specific accident categories, such as ship collisions, on-board incidents, or hazardous cargo events [50, 51],while others examine the contribution of human factors, environmental conditions, or equipment-related failures to accident risk [52, 53]. Although these works offer important insights, they tend to rely on localized datasets, which differ significantly in terminology, scope, and level of detail. A limited number of contributions have proposed frameworks for occupational health monitoring or safety performance indicators tailored to port environments [54]; however, these efforts remain constrained to national or regional settings and do not provide a global perspective.
Consequently, one of the central challenges in advancing research on maritime and port safety is the absence of harmonized, large-scale accident databases. The ability to collect, integrate, and standardize data from multiple authorities is essential for developing robust comparative analyses and designing effective prevention and mitigation strategies [55]. Without such harmonization, safety trends remain difficult to interpret, and opportunities for evidence-based policy development are significantly reduced.

In recent years, data-driven and machine learning (ML) techniques have increasingly been applied to maritime and port safety, offering promising tools for predicting accident likelihood and identifying high-risk operational conditions [56, 57]. A variety of algorithms, including SVM, KNN, LightGBM, and XGBoost, have been trained on port-specific variables such as weather conditions, cargo type, temporal patterns, and environmental parameters like temperature, humidity, wind, and currents to estimate the probability of accidents [58, 59].
Despite these advances, the practical applicability of ML models in port safety remains limited. Most studies rely on small, single-port, or short-term datasets, which makes the resulting models vulnerable to overfitting and

restricts their generalizability across ports with different operational characteristics or geographic conditions. Neural-network-based approaches face additional challenges: their performance deteriorates when trained on heterogeneous, noisy, or inconsistently labeled data, a common condition in multi-country port archives where reporting standards differ substantially.
Taken together, the literature highlights two persistent gaps. First, there is no unified methodological framework capable of harmonizing accident data across countries, despite the clear need for cross-national evidence to support safety policies. Second, existing analytical approaches often struggle to manage heterogeneity and to ensure statistically sound comparisons, particularly when conducting sensitivity analyses or exploring the interactions among operational and contextual safety factors.

To address these issues, the first thing we have done is the introduction of a structured data-fusion and experimental analysis framework that explicitly tackles the challenge of data heterogeneity. The methodology integrates:

- ETL procedures to ingest datasets from multiple national repositories and transform them into a common schema;
- lexical and semantic harmonization, using multilingual taxonomy-matching techniques to reconcile divergent accident classifications;
- and Design of Experiments (DoE) to quantify the main effects and interactions among operational, socio-economic, and environmental variables driving accident rates.

This integrated framework enables reproducible, cross-country comparisons of safety performance and provides a transparent means to test model robustness when confronted with incomplete or non-uniform data sources.

To support a rigorous and cross-national analysis of port accidents, the work adopts a methodological framework specifically designed to overcome the limitations that characterize existing studies in maritime safety. Since accident records originate from national agencies and sector-specific

authorities that use different formats, languages, and taxonomies, the first step consisted in constructing a reliable and unified data infrastructure capable of integrating this fragmented information. Multiple data sources were collected and processed through an Extract–Transform–Load workflow that converted heterogeneous files into a common machine-readable structure, preserved the relevant attributes for analysis, and removed structural inconsistencies such as duplicates or malformed entries. This initial phase provided a coherent base upon which more advanced analytical procedures could be built.
A second, crucial component of the methodology concerned the harmonization of accident classifications. Because each country adopts its own descriptive categories and levels of detail when reporting port incidents, it was necessary to translate, standardize, and semantically align all datasets to avoid the distortions that would arise from comparing incompatible taxonomies. This process involved both lexical harmonization, by translating non-English archives and unifying terminologies, and semantic alignment, by grouping conceptually similar accident types that had been labeled differently across countries. The result was a consolidated set of accident categories that preserved the specificity of national reports while enabling scientifically meaningful comparisons at the international level.
Since raw accident counts do not reflect actual safety performance, all data were normalized using exposure indicators relevant to port operations, such as annual TEU throughput or workforce size whenever available. Contextual variables were then added to enrich the dataset and allow a more comprehensive interpretation of accident patterns. These included socio-economic indicators such as GDP per capita, levels of digitization and automation, meteorological characteristics, and measures of port growth and operational intensity. The combination of exposure-based normalization and contextual enrichment ensured that differences in accident rates could be interpreted in relation to underlying operational and socio-economic conditions rather than simple differences in scale.

With a harmonized and normalized dataset in place, the analytical phase employed principles from Design of Experiments to investigate how accident rates are influenced by the interplay between operational factors, environmental conditions, and national characteristics. By structuring the analysis around main effects and interactions, it was possible to quantify the relative importance of each factor and understand how they combine to shape safety outcomes. This approach made it possible to move beyond purely descriptive statistics and toward a more explanatory understanding of risk distribution across countries. To ensure the reliability of the findings, sensitivity and robustness analyses were also conducted, examining how the results varied under different assumptions, incomplete data conditions, and cross-validation subsets. This was particularly important given the heterogeneity of the original data sources.

DoE provides a structured methodology for exploring the effects of multiple factors simultaneously, allowing the analyst to quantify both their individual contributions and the interactions that emerge when variables operate together rather than in isolation. Unlike traditional one-factor-at-a-time approaches, DoE is specifically intended for complex systems, such as port operations, where accident dynamics arise from the interplay of many interdependent conditions. In this study, a full factorial design was selected, as it represents the most rigorous and informative experimental structure. By evaluating every possible combination of the selected factor levels, the full factorial design enables a complete exploration of the response space and ensures that interaction effects are estimated without confounding. This approach is particularly suitable for heterogeneous multi-country datasets, where the relationship between factors such as accident type, port throughput, digitization level, and socio-economic conditions cannot be assumed to be linear or independent. The adoption of a full factorial design thus provides a robust analytical foundation for interpreting accident variability and identifying the key drivers of safety performance across different national contexts.

The final part of the framework establishes the link between empirical evidence and the simulation environment developed in this thesis. The harmonized accident dataset and the insights extracted from the experimental analysis were used to inform the construction of realistic scenarios within the simulator, including the modeling of hazard dynamics, agent behaviors, and risk conditions. In this way, the simulation environment does not rely on hypothetical or idealized assumptions but is grounded in observed accident patterns and validated relationships between variables. The methodological framework therefore not only enables a robust analysis of global port safety but also provides the empirical foundation required for credible modeling, simulation-based training, and AI-driven decision support.

Table 1: Archives used for processing port accident data

Country	Source archive
Hong Kong (HK)	Marine Department of The Government of the Hong Kong Special Administrative Region (2023)
India (IN)	Directorate General Factory Advice Service & Labour Institutes - Ministry of Labour & Employment (2023)
Italy (IT)	INAIL (2023)
Japan (JP)	Japan Industrial Safety & Health Association (JISHA, 2023)
New Zealand (NZ)	WorkSafe – New Zealand Government (2023)
South Africa (ZA)	South African Maritime Safety Authority (SAMSA, 2023)
Turkey (TR)	Social Security Agency of Turkey (SGK, 2023)
USA (US)	U.S. Bureau of Labour Statistics (BLS, 2023)
United Kingdom (UK)	Port Skills and Safety association (PSS, 2023), part of British Ports Association and UK Major Ports Group

The selection of countries considered in this study was driven by pragmatic and scientific requirements. The analysis required datasets that were sufficiently complete, continuous over time, and rich enough in detail to allow for comparative investigation. For this reason, the study focused on Hong Kong, India, Italy, Japan, and the United States. These countries offer accident records of adequate temporal depth and operational relevance, while also representing different levels of economic development and technological maturity. Hong Kong and Japan illustrate port systems that operate within advanced, highly digitized economies. India reflects a rapidly expanding maritime sector where automation and digital solutions are progressively spreading. Italy provides an example of European port

governance characterized by a long industrial tradition. The United States contributes a large-scale logistics environment with long-term and systematically collected occupational safety data.
This configuration makes it possible to explore a broad spectrum of regulatory structures, labor practices, and technological readiness levels, which in turn strengthens the validity of cross-country comparisons. At the same time, the sample inevitably excludes several important geographic regions, such as Africa or South America, where accessible accident records remain scarce. As a consequence, the findings of this study should be interpreted as representative of broader global tendencies but not as a complete worldwide description. The value of the investigation lies primarily in its methodological transferability, which can be applied to any region as soon as comparable data become available.
To enable systematic comparison, the authors constructed an integrated repository of correlated and preprocessed data. This repository supports the entire analytical workflow, from exploratory statistics to comparative evaluation and factorial Design of Experiments. The data infrastructure was intentionally designed to accommodate heterogeneous sources and to facilitate the exploration of relationships between accident dynamics, logistics flows, and macro-level indicators.
The initial stage of this process involved the ingestion and standardization of datasets obtained from national authorities. Records were imported in their original formats, including CSV files, spreadsheets, and HTML tables, and were transformed into a unified schema through an Extract Transform Load pipeline. Metadata such as reporting institutions, publication years, and data provenance were retained to guarantee traceability. Attribute names were aligned with a controlled vocabulary that included fields for event type, severity, occupation, impact category, location, and temporal information.
A second stage addressed the heterogeneity of accident descriptions. Since reporting systems in different countries rely on inconsistent taxonomies and multilingual labels, harmonization was essential. A hybrid matching approach was employed. It combined lexical similarity measures, such as

Levenshtein and Jaro Winkler distances, with embedding-based semantic similarity extracted from multilingual language models [60] [61]. Ambiguous matches were examined manually through expert review, and a reusable mapping table was produced to ensure transparency and reproducibility. This strategy made it possible to align classification schemes across countries without performing exhaustive manual recoding.

Once taxonomies were harmonized, the datasets were merged through a vertical fusion process. Accident records were linked with contextual information, including annual port throughput, gross domestic product per capita, and digitization indices obtained from international organizations. This step allowed each record to be associated consistently with its operational and socio-economic environment, forming the basis for exposure-adjusted analyses. All transformations and joins were automatically logged, ensuring that the entire process remained auditable.

The fused dataset was then refined through feature engineering. Categorical variables such as occupation or injury mechanism were encoded using a hierarchical clustering technique that grouped low-frequency categories according to their statistical similarity. This prevented sparsity and enabled more stable comparisons across countries with different reporting granularities. Continuous variables, including TEU, GDP, and digitization levels, were normalized and discretized into low, medium, and high ranges to make them compatible with the requirements of factorial experimental designs. The dataset underwent a series of validation steps. Internally, accident counts were evaluated for temporal continuity using rolling averages and anomaly detection. Externally, the harmonized distributions were compared across countries using non-parametric statistical tests to detect possible inconsistencies. Records flagged during this process were reviewed and, when necessary, reclassified.

The resulting repository made it possible to investigate complex relationships between accident causes, types, occupational roles, and contextual indicators. By clustering similar job tasks, it became possible to identify which activities exposed workers to higher risks. By correlating accident types and underlying causes, the analysis revealed recurrent patterns of failures with operational significance. More advanced correlation studies, based on factorial contrasts and effects, highlighted higher order interactions between variables. These findings were essential for understanding the multifactorial nature of accident causation and provided a foundation for developing targeted prevention strategies.

The Design of Experiments used in this study was conceived to explore how different operational, macro-economic, and contextual variables shape accident dynamics across countries. The objective was to quantify both main effects and interactions among the most relevant drivers, to estimate how sensitive accident rates are to changes in digitization and traffic intensity, and to identify those factors that should be prioritized when developing training and planning scenarios within the simulation and extended reality environment. To allow comparisons among countries with very different operational scales, the primary responses were expressed as exposure-adjusted rates. Accident, severe injury, and fatality rates were calculated as the number of respective events per million TEU handled each year. When TEU values were not available for specific years, they were imputed using linear or piecewise cubic interpolation based on official port statistics. All sensitivity analyses were conducted both with and without imputation to ensure that conclusions remained consistent.
The choice of factors included in the experimental analysis was guided by the need for variables that could be consistently retrieved or reconstructed across archives. Only indicators with sufficient coverage across countries and time were retained. Because the available data originate from observational records and because reporting formats vary considerably from country to country, the experimental design followed a sequential multi-stage logic rather than a single balanced factorial structure.
The first stage was a screening phase intended to identify the most influential variables and any interactions that might warrant closer examination. To implement this stage, continuous variables such as digitization level, GDP per capita, and traffic intensity were discretized into broad categories, for example low, medium, and high. These categorical

versions of the variables were then combined with structural contrasts, such as countries grouped by broad regional characteristics, to approximate a fractional factorial arrangement. The aim was to estimate which factors exert the strongest influence on accident rates and to detect meaningful two-way interactions. Since the original dataset is unbalanced, some combinations of factor levels were missing. To address this issue, additional runs were constructed through nearest-neighbor matching across countries and years. All key findings from this screening phase were then replicated both with and without these augmented data to confirm that the detected effects were not artifacts of the reconstruction procedure.

The results of the screening stage were clear and consistent. Traffic intensity and digitization emerged as the two variables with the most substantial and stable impact on accident rates across all countries. GDP per capita showed weaker and more variable effects, and therefore played a secondary role in the subsequent stages of the analysis.

The second stage focused on refining the relationships identified during the screening phase. At this point, the analysis shifted to response surface methodology. Continuous variables were used in their natural form rather than discretized, and the models included quadratic and interaction terms to capture possible curvature in the response. A central composite arrangement was adapted to the ranges observed in the data, with the purpose of investigating marginal changes and scenario-based variations. This approach made it possible to simulate how a moderate increase in digitization, for example on the order of half a standard deviation, could reduce accident rates under stable traffic conditions. Since structural differences among countries would otherwise distort the response surface, country effects were treated as blocks that absorb national baseline differences.

The response surface stage revealed that the effects of digitization are not linear across the entire range. Increases in digitization are associated with substantial improvements in safety, but only up to a certain level, after which the marginal benefits become more limited. This finding suggests that

digitization contributes to safety most significantly during early and intermediate levels of technological adoption.
The final stage served as a confirmatory analysis. To test the stability of previously identified effects, the dataset was modeled using a mixed effects framework suitable for count data. The models incorporated TEU exposure as an offset and included random intercepts for each country together with random slopes for temporal trends. By combining fixed and random effects, this analytical structure allowed the investigation to control for unobserved heterogeneity and repeated observations over time while still estimating the influence of the main experimental factors. Block effects were assigned to the source of each dataset so that differences in reporting style or classification could be absorbed and would not bias the conclusions.
This confirmatory stage verified the robustness of the earlier findings. Digitization and traffic intensity remained the dominant explanatory factors, while the influence of job task composition and impact categories, although secondary, contributed meaningfully to differences in severe injury cases. By progressing from exploratory screening to response surface refinement and concluding with mixed effects confirmation, the design ensured that results remained systematic, coherent, and resilient to the limitations inherent in observational safety datasets.

The harmonized accident dataset allowed a comparative assessment of both fatal and non-fatal incidents across the port systems of Hong Kong, Japan, Italy, and India. The multi-year comparison revealed clear differences in the temporal evolution of accident frequencies among these countries. Hong Kong and India, the two cases for which accident records are available from the late 1990s onward, exhibit a gradual decline in the number of reported incidents over the last twenty-five years. This downward trend is consistent with a progressive improvement in occupational safety practices, the introduction of more structured governance frameworks, and the adoption of digital tools that support monitoring and supervision.
In contrast, the accident curves for Italy and Japan do not show the same declining pattern in the years for which data are available. For these

countries, the time series cover only recent years, which makes long-term interpretation difficult. However, the absence of a clear downward trend suggests that safety performance is either improving slowly or fluctuating around a relatively stable baseline. In both cases, the limited temporal depth of the datasets indicates that more complete archives would be required to assess long-term progress.

To account for the very different sizes of national port systems, accident rates were normalized by container traffic. Figure 1 illustrates the relationship between annual TEU throughput and registered fatalities across the countries included in the study between 1997 and 2021. The normalization, expressed as the number of TEUs handled per each fatal accident, provides a more meaningful indicator of safety performance than absolute fatality counts, which are not comparable across systems of different scales. Higher ratios reflect a greater number of containers moved per fatality and therefore correspond to higher levels of safety efficiency.

The comparison reveals that Hong Kong and India, which have both expanded their maritime activity considerably over the past decades, manage to maintain or improve their fatality-adjusted throughput levels. This indicates that their operational growth has not been accompanied by a proportional increase in severe accidents. Italy and Japan show more modest improvements in this indicator. The results must be interpreted with caution due to the limited number of years available, yet they suggest that improvements in safety performance are not uniform across these systems.

Beyond traffic exposure, the analysis explored broader socio-economic variables that might influence accident patterns. Ratios reflecting literacy, average wealth, and technological familiarity were examined in order to evaluate whether general developmental conditions correlate with safety outcomes. To support this investigation, a correlation study was performed using historical series from Hong Kong, India, Italy, and the United Kingdom, focusing specifically on two indicators: digitization level and GDP per capita. Spearman’s rank correlation coefficient was used to capture monotonic relationships without assuming linearity.

The correlations reveal a consistent association between accident rates and both socio-economic parameters. Higher digitization levels tend to correspond to lower accident frequencies, while higher GDP per capita also aligns with improved safety conditions. These relationships, although exploratory, are coherent with the idea that technological maturity and general economic development create conditions that favor safer working environments. More widespread digital systems facilitate supervision, communication, and early detection of hazardous situations. Likewise, higher income levels often correlate with better training, improved equipment standards, and more structured regulatory environments.

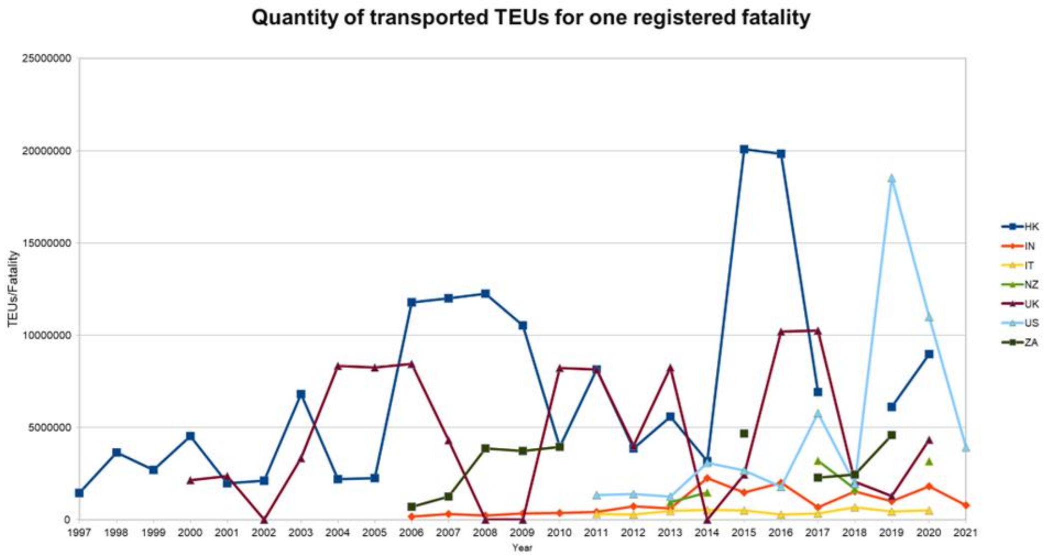

Figure 1. Quantity of transported TEUs per one registered fatality across selected countries (1997–2021)

Table 2. Correlation between a specific parameter and number of accidents (Fatal/Severe) normalized by traffic volume.

Parameter	*Accident Type*	*Hong Kong*	*India*	*Italy*	*United Kingdom*
Digitization	Severe	-91.33%	-96.77%	13.46%	-76.59%
	Fatal	-55.41%	-90.11%	41.19%	-19.46%
GDP	Severe	-81.54%	-95.35%	-40.88%	-56.20%
	Fatal	-55.30%	-84.81%	-67.47%	-17.99%

To improve the interpretability of the correlations reported in Table 2, a graphical representation was introduced in the form of a correlation heatmap. The heatmap visually encodes the magnitude and sign of the correlation coefficients between selected socio-economic parameters and the number of severe and fatal accidents, normalized by traffic volume. This representation allows for an immediate comparison across countries and parameters, highlighting structural similarities, divergences, and outliers that are less evident in tabular form.

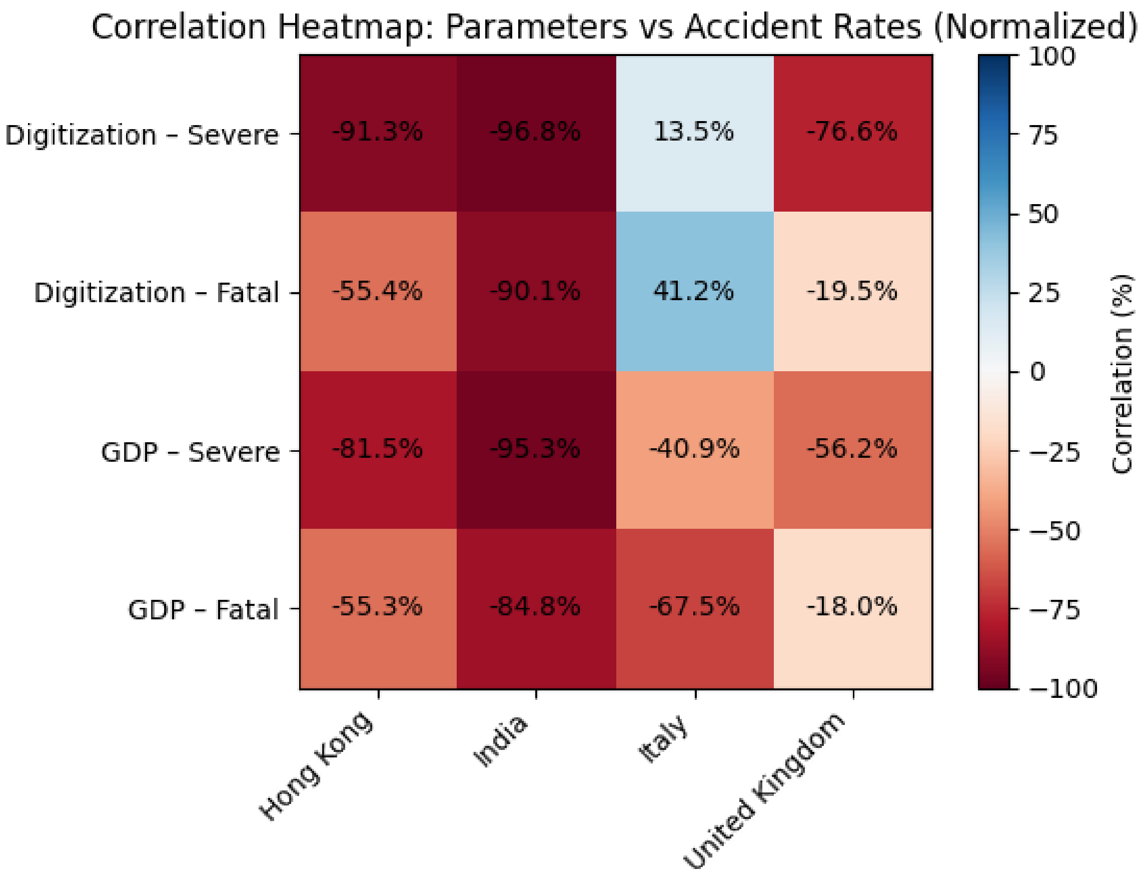

Figure 2: Correlation Heatmap between Socio-Economic Parameters and Accident Rates (Normalized by Traffic Volume)

The heatmap representation reinforces several important observations already present in Table 2. In particular, it clearly highlights the strong negative correlation between digitization levels and accident severity in Hong Kong, India, and the United Kingdom, suggesting a consistent association between higher levels of digital maturity and reduced accident rates. Conversely, Italy exhibits a markedly different pattern, with positive correlations for digitization and both severe and fatal accidents, indicating that increased digital adoption alone does not automatically translate into improved safety outcomes.

A similar contrast emerges for GDP-related correlations, where most countries show a negative association with accident frequency, while Italy again deviates from the general trend. The visual clustering produced by the heatmap makes these discrepancies immediately apparent, supporting the interpretation that socio-technical factors, regulatory frameworks, and implementation quality play a critical role alongside purely economic or technological indicators. As such, the graphical analysis strengthens the argument that safety improvements in complex systems cannot be attributed to single parameters, but emerge from the interaction between technology, governance, and operational practices.

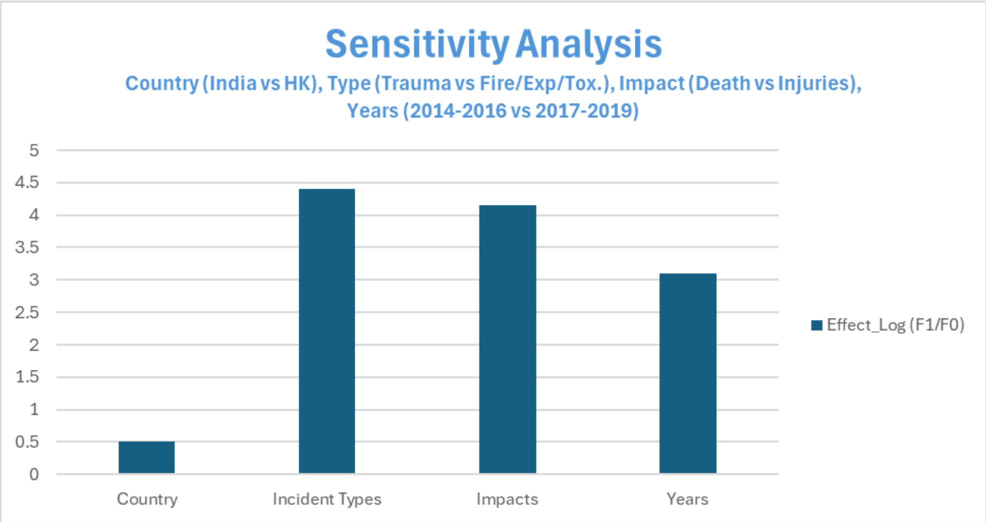

Figure 3. Extended sensitivity matrix illustrating higher-order interactions among key variables: country (x_4), incident type (x_5), impact (x_6), and years (x_{10}).

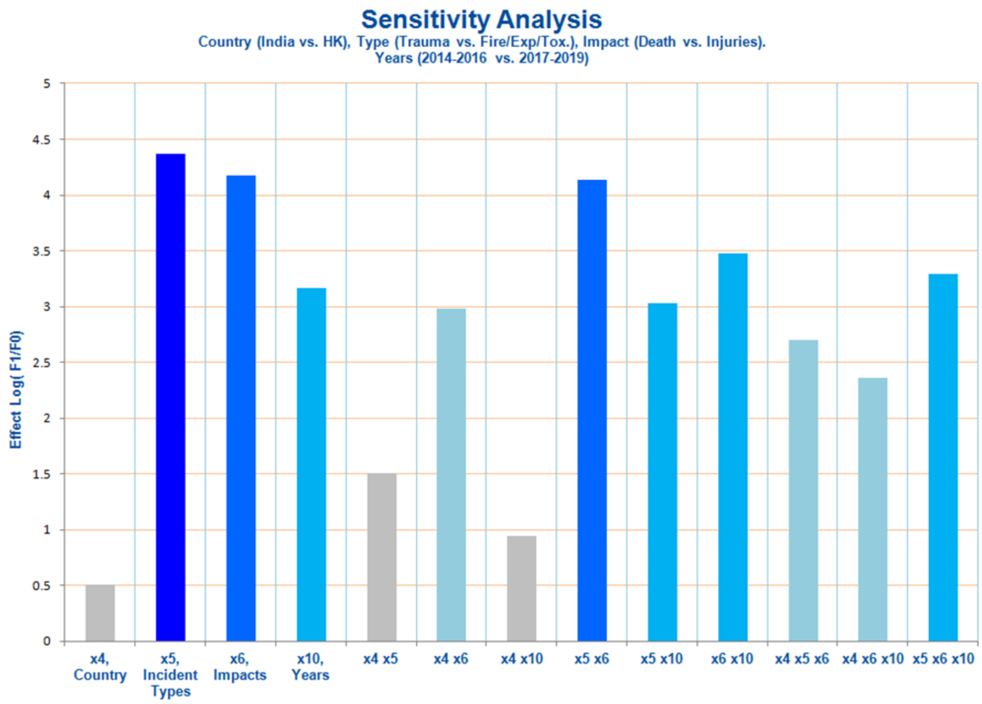

Figure 4. Sensitivity analysis comparing the main effects of country, incident type, impact, and years (2014–2019). The log(F/F_0) scale indicates relative effect strength

The multifactorial analyses carried out in this study were designed to explore how different operational, environmental, and contextual variables interact to influence port safety. The goal was to identify sensitive parameters and determine how variations in these factors shape the overall risk profile of port systems. Sensitivity results for India and Hong Kong, reported on a normalized logarithmic scale, highlight the relative importance of the variables considered. The patterns show that categories such as incident type and impact category exert substantial influence on model outputs. Their interaction is also meaningful. When examined jointly, these two dimensions reveal combined effects that are not visible when analyzing them independently. The country effect is present but its contribution is limited once other operational factors are accounted for, indicating that structural differences among nations play a secondary role compared to specific incident characteristics.

The comparative evaluation extends across five countries for which sufficiently long historical series were available: Hong Kong, India, Italy, Japan, and the United States. The datasets vary in their temporal coverage. Hong Kong and India provide multi-decade records, whereas Italy and Japan

offer more recent and shorter time windows. All datasets were harmonized and processed using the unified framework developed in the earlier stages of this research, ensuring consistency in the way variables were interpreted and analyzed.
The results reveal that Hong Kong and India have experienced notable reductions in both accident and fatality rates across the observed periods. This improvement aligns with the adoption of national safety reforms, increased attention to accident reporting, and the progressive introduction of digital tools for monitoring and oversight. In contrast, Italy and Japan do not show comparable downward trajectories within the available time frame. These trends may reflect genuine stagnation in safety improvements, but they could also result from the shorter duration of the available records, which limits the ability to capture long-term dynamics. The United States shows moderate declines over a comparable time horizon, with clearer improvements observed after 2010, when national reporting procedures were strengthened.
Normalizing accident and fatality measures by TEU throughput allows a more accurate comparison among countries with different traffic volumes. When rates are expressed as events per million TEU, the relative trends become clearer, revealing differences in safety efficiency that are not evident from absolute counts alone.
Italy represents a particular case within the dataset. Its correlation between digitization indicators and normalized accident rates appears weaker and sometimes atypical compared to other countries. This anomaly is consistent with administrative transitions in the national reporting system rather than a genuine deterioration in safety. Italy introduced the Cruscotto Infortuni platform in 2016, initially as guidance for institutions, and subsequently made reporting mandatory for certain categories of injuries. In 2025 the system evolved into the Registro Infortuni Telematico. These technological upgrades tend to improve reporting completeness, especially for minor or non-fatal cases that were previously underreported. As a result, accident counts often rise during such transitions even when actual safety conditions remain unchanged. INAIL’s own documentation cautions that historical

datasets follow different extraction logics and should not always be directly combined, which further contributes to the apparent irregularities. The limited number of years available for the Italian port subset amplifies these fluctuations and makes the series more sensitive to short-term variations.
Overall, the combined results from sensitivity, trend, and normalized analyses show that while national context plays a role, the most influential determinants of accident patterns are embedded within operational categories and incident characteristics. Countries with longer time series and consistent reporting practices display clearer improvements in safety outcomes, whereas countries undergoing administrative transformations or offering shorter archives present more irregular trajectories. These findings provide important insights for both model calibration and scenario development within the simulation environment.

The analytical framework developed in this research not only supports the identification of the most significant accident drivers but also provides a structured basis for determining where preventive and mitigation efforts should be concentrated. Understanding how accidents evolve, which conditions amplify their severity, and how contextual factors influence risk makes it possible to define precise operational and training requirements. These insights highlight the need for tools that can translate statistical evidence into actionable strategies for both decision-makers and frontline personnel.
Planning and training play complementary roles in the management of port operations. Planning is essential for decision-makers who must anticipate hazards, allocate resources, and implement safety policies, while training is fundamental for operators who must work safely within complex and dynamic environments. In recent years, modeling and simulation have emerged as effective enabling technologies for supporting both domains. By reproducing the behavior of real systems in a synthetic environment, simulation makes it possible to experiment with conditions that would be too dangerous, too costly, or too disruptive to replicate in reality. Scenarios involving explosions, hazardous material leaks, or equipment failures can

be studied safely by means of computational models that accurately represent the spatial, temporal, and operational characteristics of a terminal. The MS2G paradigm, which integrates simulation, interoperable simulation, and serious gaming, has proven particularly effective. It couples the fidelity of traditional M&S with the engagement and intuitiveness typical of game-based learning and extended reality. This approach allows users to interact with realistic virtual environments, perceive risks dynamically, and develop effective situational awareness. Over the past years, several research initiatives have demonstrated that combining simulation with machine learning techniques can significantly enhance system design, operational planning, and real-time decision support for industrial and offshore infrastructures. These results reinforce the potential of similar approaches for improving safety and efficiency in port systems.

The development of the COYOTE simulator for this thesis is a direct response to these needs. The statistical analyses presented earlier reveal that accident patterns are shaped by a combination of operational factors, contextual conditions, and organizational practices. Simulation provides a way to embed this knowledge into an interactive tool capable of supporting both learning and strategic assessment. COYOTE enables operators to train within realistic and evolving scenarios, where risks emerge organically from the simulated environment. At the same time, it offers decision-makers a platform for evaluating the potential impact of accidents, testing alternative courses of action, and assessing the consequences of changes in traffic, layout, procedures, or digitization level.

Through this integration of empirical evidence and dynamic modeling, COYOTE contributes to reducing vulnerabilities, minimizing the likelihood of accidents, and supporting safer and more resilient port operations. It translates the methodological findings of this research into a concrete system that can support training, planning, and continuous improvement in safety management.

2.2. Extended Reality for Training and Human Factors

The training of operators in safety-critical environments presents significant challenges when conducted solely through real-world exercises, as the associated risks, operational disruptions, and economic costs often limit the frequency and realism of such training. To overcome these constraints, virtual training laboratories have emerged as an effective alternative, supported by advances in Modeling & Simulation and the consolidation of serious games as a mature educational paradigm [62]. For more than three decades, serious games have demonstrated their ability to transform interactive digital environments into powerful tools for training, safety education, and performance assessment. Their ability to reproduce hazardous scenarios without exposing personnel to danger makes them particularly suited for industrial, maritime, and defense applications, where traditional training may be infeasible or insufficient [63].

Beyond reproducing operational tasks, serious games enable a broader capability assessment by monitoring a wide range of Key Performance Indicators during simulated missions. This analytical potential allows instructors and designers to evaluate not only procedural correctness but also decision-making quality, situational awareness, risk perception, reaction patterns, and operator behavior under stress. For this reason, serious games have been successfully applied in domains such as CBRNe emergency response, hazardous-materials logistics, military decision-making, and high-risk industrial operations. Recent technological developments have further expanded the possibilities of this approach. The integration of Extended Reality technologies into serious-game platforms has opened new avenues for immersive, adaptive, and experiential training, providing operators and decision-makers with realistic conditions without the limitations of physical exercises [64].

Extended Reality, which encompasses Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR), has become an established training modality in several industrial sectors, particularly where the cost and risk of real-world simulation would otherwise hinder effective practice. Immersive technologies, also referred to as digital reality, combined with the Internet of Things and advanced control systems are enabling increasingly sophisticated training ecosystems capable of reproducing complex operational interdependencies. Over the past decade, improvements in graphics engines, rendering techniques, motion tracking, and wearable interfaces have significantly increased the realism of immersive experiences while simultaneously reducing costs. The evolution from early CAVE systems, which required investments on the order of one million dollars, to more compact and interoperable solutions such as the SPIDER platform introduced by the Simulation Team demonstrates this dramatic shift in accessibility. These advancements allow organizations to deploy scalable training infrastructures tailored to different levels of expertise and budget.

Modern XR ecosystems now include a wide range of hardware configurations, from head-mounted displays and projected environments to motion platforms, motion-tracking systems, haptic interfaces, and full-body suits. Software improvements have enabled the dynamic rendering of large environments, integration with intelligent agents, and real-time data collection for performance analytics.

Figure 5. Example of eXtended Reality Tool in Simulation Team Laboratories

The development logic adopted for COYOTE allows the simulator to be deployed across multiple hardware platforms, each characterized by different computational capabilities and interaction modalities. This multiplatform flexibility ensures that the system can adapt to diverse training needs, ranging from lightweight mobile applications to high-fidelity immersive environments. Smartphones constitute the most accessible solution, given their widespread diffusion and intuitive usability. They are particularly suited for augmented reality applications in which virtual elements are overlaid onto the real world, enabling training solutions that combine physical context with simulated cues. Mobile versions of the simulator operate on devices equipped with Android 8.0 or later and rely on integrated gyroscopic sensors to provide basic 3D immersion.

Figure 6. Stereostopic Vision with VR display

Tablets offer a complementary alternative by providing a larger display surface, improved visual clarity, and lower cost, making them well-suited for distributed training or future extensions of COYOTE toward supervisory and operational-control functionalities. Their portability, however, is lower than that of smartphones, reducing usability in highly mobile or physically constrained environments.

Figure 7. Basic headset for VR

At the next level of immersion, basic headsets such as Google Cardboard combine a smartphone with a low-cost stereoscopic viewer, allowing users to experience virtual environments with a greater sense of depth while maintaining minimal hardware requirements.

Table 3. Characteristics of some XR solutions

Samsung Gear	***Oculus Quest, Rift, HTC Vive***	***Microsoft Hololens v2***	***SPIDER***	
Smartphone based	Stand alone	Computer based	Stand alone	***Type***
~90 euro	~360 euro	~4000 euro	~40000	***Price***
318g	571g	848g	566g	***Weight***
1 6DoF controller with 5 buttons and joystick	2 6DoF Controllers with 5 buttons, joystick and touch proximity sensors each	2 6DoF Controllers with 5 buttons, joystick and touch proximity sensors each	Gesture recognition	***Controllers***
Relative positioning using accelerometer and gyroscope	Tracking of controller position respect to the headset	Room-scale tracking of headset and controllers by external sensors	Inside-out spatial mapping	***Positioning***

More advanced augmented reality devices, such as Microsoft HoloLens, integrate processing power, display technology, gesture recognition, and voice interaction within a single headset, enabling fully immersive mixed-reality experiences without the need for external hardware. These systems are particularly valuable for training workflows that require continuous interaction with both synthetic and physical objects. For tasks requiring higher computational loads, such as large-scale scenes, complex agent behaviors, or real-time data analysis, COYOTE can be deployed on laptops and workstations, which provide significantly greater processing power at the expense of mobility. These platforms allow for richer simulation logic and higher visual fidelity, supporting research-oriented experimentation and advanced operational training.

Figure 8. Augmented Reality applied to maritime context

At the highest end of the immersion spectrum, COYOTE can operate within CAVE systems (Cave Automatic Virtual Environment), which have long been used in defense, medicine, and industrial training to create shared virtual spaces in which multiple users can enter and interact. Modern CAVE technologies rely on direct projection or rear-projection systems that surround the user with images on multiple walls, enabling collaborative and spatially rich training experiences. Despite their effectiveness, traditional CAVEs are physically demanding in terms of space and budget, often requiring investments exceeding 100,000 euros and, in professional configurations, reaching up to one million. To overcome these limitations, the Simulation Team developed SPIDER (Simulation Practical Immersive Dynamic Environment for Reengineering), a compact, interoperable, and interactive CAVE system. With a basic configuration of approximately 2 × 2 × 2.6 meters, SPIDER can be transported within a standard shipping container, integrated with any interoperable simulator, and operated through touchscreen-based interaction. This compact design drastically reduces logistical barriers while preserving the immersive qualities of high-end virtual environments.

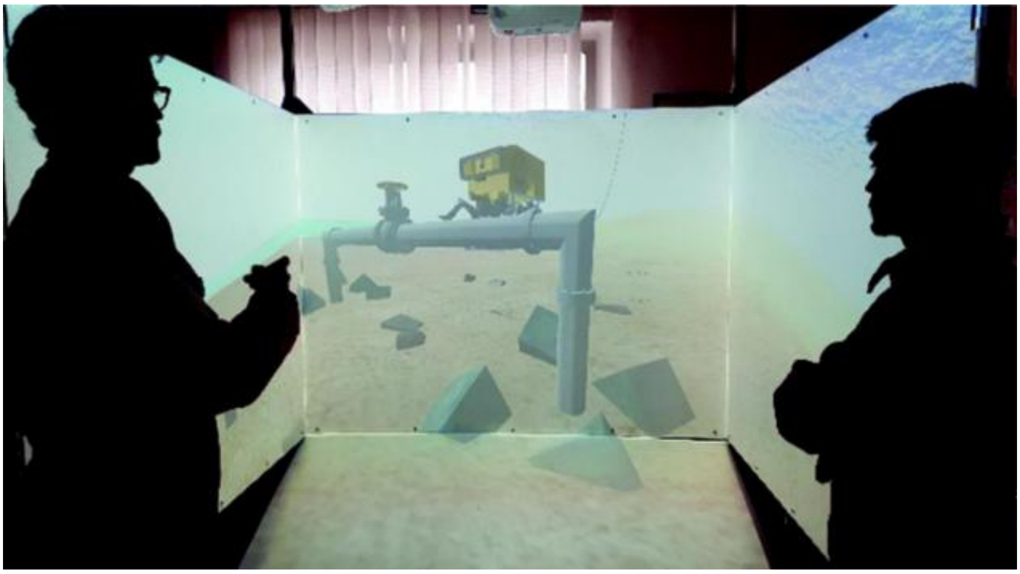
Figure 9. Marine Virtual World inside the SPIDER CAVE of Simulation Team

Through this diversified platform strategy, COYOTE ensures that training and experimentation can be tailored to the technical constraints, financial resources, and specific requirements of different users and organizations. The same simulation logic can therefore support mobile AR training, VR immersion, workstation-based research, or large-scale shared virtual environments, making COYOTE a flexible and scalable solution for safety training and decision-support in complex operational systems.

The selection of XR hardware for port and terminal applications was driven by a combination of technological, ergonomic, and environmental considerations. Unlike office or laboratory settings, port environments are characterized by aggressive operating conditions, including high humidity, saline atmosphere, airborne dust and particulate matter, mechanical vibrations, temperature variations, and intense or irregular lighting. These factors pose significant challenges to electronic devices originally designed for consumer or indoor professional use and therefore must be explicitly considered when assessing the feasibility of XR-based solutions.

In this context, the hardware comparison presented in Table 3 should be interpreted not as a purely technological benchmark, but as part of a broader deployment strategy. Stand-alone and mixed reality devices were prioritized due to their limited reliance on external computing units, cables, or tracking infrastructure. Reducing the number of external components lowers the risk of hardware failure, simplifies logistics, and improves reliability in semi-industrial environments where space, cleanliness, and setup time are constrained. Similarly, inside-out tracking and spatial mapping solutions were favored over external sensor-based tracking, as they avoid the need for fixed installations that could be degraded by dust accumulation, humidity, or accidental impacts.

Ergonomics and operator comfort were also treated as durability-related factors. Excessive headset weight, poor balance, or thermal discomfort can lead to fatigue, reduced attention, and lower acceptance by operators, especially during repeated training sessions. Devices with integrated processing and optimized weight distribution were therefore considered more suitable for sustained use, even if their raw computational performance was lower than that of tethered systems. From a human factors perspective, this trade-off contributes directly to safer and more effective training outcomes.

With respect to environmental resilience, the proposed approach does not assume that XR devices are intrinsically ruggedized for direct exposure to harsh outdoor conditions. Instead, durability is achieved through a combination of operational constraints and protective measures. These include the use of protective housings, sealed or replaceable face interfaces, industrial-grade straps and connectors, and controlled storage and charging procedures to mitigate long-term exposure to salt, moisture, and contaminants. Such measures are common in industrial digitalization initiatives and allow commercially available devices to be adapted to challenging environments without resorting to costly bespoke hardware.

It is also important to clarify the intended operational context of the XR solutions. The devices are primarily employed for training, rehearsal, and decision-support activities, typically conducted in controlled or semi-controlled environments such as training rooms, control centers, or designated simulation areas within the port. The realism of hazardous scenarios is achieved through high-fidelity simulation rather than direct exposure to operational risks. This assumption significantly reduces mechanical and environmental stress on the hardware while preserving the benefits of immersive learning and risk awareness enhancement.

The hardware selection strategy reflects a conscious balance between robustness, usability, scalability, and cost-effectiveness. Rather than pursuing fully ruggedized XR systems, which are currently limited and expensive, the approach emphasizes modularity, protective adaptation, and organizational measures. This strategy ensures that XR technologies can be realistically integrated into port training programs and decision-support workflows, while remaining sustainable from both a technical and economic standpoint.

2.3. M&S and Digital Twins

Simulation represents one of the most powerful methodologies for reproducing the behaviour of real systems within controlled environments, enabling the study, analysis, and experimentation of processes that would otherwise be too dangerous, costly, or impractical to examine directly [65]. At its core, simulation relies on the development of a model, a surrogate abstraction of a real system, object, or process, which encapsulates the essential structural and behavioural elements of the system of interest. By allowing this model to run within a computer-based simulation environment, it becomes possible to investigate dynamic behaviours, test hypotheses, and perform extensive "what-if?" analyses under precisely controlled conditions. This capability is particularly relevant for scenarios in which real-world experimentation is infeasible, such as the dispersion of toxic clouds in inhabited areas, the development of severe weather events, or the unfolding of hazardous industrial accidents. In such cases, simulation provides a safe and cost-effective platform where hurricanes, chemical spills, or emergency procedures can be reproduced and analysed virtually [66].

Compared to the direct study of real systems, simulation offers multiple advantages. Testing real processes or equipment may require resources that are unavailable, prohibitively expensive, or unacceptable from a risk-management perspective. A virtual counterpart, by contrast, can be executed repeatedly, modified instantaneously, and configured to explore rare or extreme conditions. Moreover, simulation inherently supports dynamic analysis over time, capturing the evolution of systems and their interactions. This characteristic distinguishes it from steady-state analytical tools such as classical optimization algorithms, which often ignore temporal variability. Modern approaches frequently combine simulation with optimization techniques or Artificial Intelligence (AI), enabling the development of optimal or adaptive strategies across a wide range of application domains, including port operations, retail, and environments where machine learning or heuristic methods play a key role.

The versatility of simulation has promoted the development of numerous methodologies tailored to specific contexts. In military domains, for example, the Live, Virtual, and Constructive (LVC) taxonomy is used to distinguish simulations based on the nature of human and system involvement [67]. Live simulations involve real personnel operating real systems, such as military drills or CBRN emergency training, while Virtual simulations employ real people operating simulated systems, as in flight simulators where trainees interact with a synthetic cockpit. Constructive simulations, on the other hand, model both systems and operators, with human involvement limited to high-level decision making. Other taxonomies distinguish simulations according to determinism or randomness: deterministic simulations generate identical outcomes when initial conditions are fixed, while stochastic simulations incorporate random sampling and therefore require statistical post-processing to interpret results. Time management also varies: continuous-time models evolve smoothly, whereas Discrete-Event Simulations (DES) advance through a sequence of triggered events. Real-time relationships between simulated and physical time further distinguish paced simulations, which run synchronously with real time, from unpaced simulations, which may accelerate or decelerate temporal flow. In many cases, human operators interact directly with the simulation (a paradigm known as Man-in-the-Loop) which introduces behavioural dynamics and psychological factors into the simulation process.

The integration of digital technologies has expanded the scope and complexity of contemporary simulations, giving rise to the discipline of Modeling & Simulation (M&S). M&S provides a structured approach for modelling, developing, validating, and executing representations of real or conceptual systems, and is now fundamental across aerospace, defence, industrial engineering, logistics, and complex system design. Crucially, simulation allows the investigation of Systems of Systems (SoS), where multiple interacting subsystems generate emergent behaviours that cannot be deduced from isolated components alone. Real-world industrial and logistics environments often exhibit such emergent phenomena, arising from nonlinear interactions, feedback loops, variability in human behaviour, and stochastic disturbances. Analytical models often prove insufficient to capture these complexities, whereas simulation provides a flexible and expressive tool for understanding SoS dynamics. When augmented with Strategic Engineering, a framework integrating AI, M&S, and Data Analytics, simulation becomes a strategic capability supporting decision-makers in planning, resource optimization, and operational risk management.

From a formal perspective, a model constitutes a mathematical abstraction of a system and is defined by state variables, equations of evolution, parameters, constraints, and input-output relationships. The conceptual model represents the logical-mathematical specification of these elements. Once implemented through programming languages, simulation engines, or graphical platforms, the conceptual model becomes an executable simulation capable of producing dynamic behaviour under varying inputs. The quality and credibility of the simulation depend critically on the correctness of the model: an inaccurate or poorly conceived model may misrepresent reality and lead to misleading conclusions. Model imperfections may arise from oversimplification, conceptual misalignment, inadequate data, programming errors, or inconsistencies between the model and the phenomena of interest.

Because of these challenges, the development and use of simulation models require a disciplined methodological framework. The life cycle of an M&S system mirrors that of systems engineering, encompassing model conception, implementation, verification, validation, execution, and interpretation. Various organizations have formalized these processes. For example, NASA-STD-7009B defines requirements and provides guidelines for the development, assessment, and credibility evaluation of M&S systems, offering best practices for analysis, verification, validation, uncertainty quantification, and presentation of results.

Through this structured approach, simulation becomes not just a computational tool but a methodological framework enabling rigorous, repeatable, and scientifically grounded exploration of real and conceptual systems across a broad range of operational and strategic contexts.

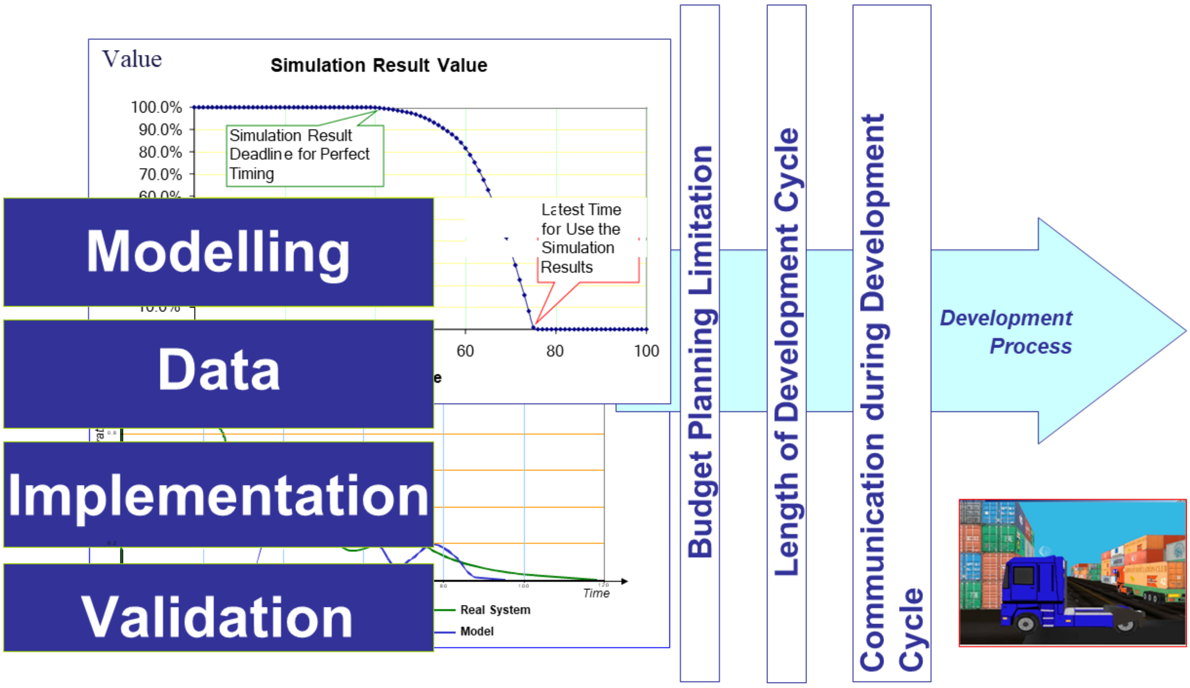

Figure 10. Example of M&S Development Process

Verification, Validation, and Accreditation (VV&A) constitute the methodological backbone of credible Modeling & Simulation (M&S) systems and are widely recognized across defence, aerospace, maritime, and industrial domains as essential processes to ensure that simulation models are trustworthy, fit for purpose, and aligned with the operational context for which they are intended. Historically formalized within the US DoD and NASA communities, and later consolidated in NATO's AMSP-01 and STANAG frameworks, VV&A has evolved from a strictly technical quality-assurance methodology into a comprehensive discipline integrating conceptual-model assessment, data pedigree evaluation, uncertainty quantification and risk-based decision criteria. Verification focuses on assessing whether the model is built correctly, through checks of internal consistency, algorithmic correctness, numerical stability and compliance with specifications. Validation addresses whether the correct model has been built, requiring comparison with empirical data, operational behaviour, SME judgement and statistical goodness-of-fit metrics. Accreditation, in turn, is a formal decision by an authoritative body that the model is acceptable for a specific intended use, recognizing that a model can never be universally valid but only valid relative to a defined context of application. Modern VV&A research increasingly emphasises iterative, evidence-based processes, such as the NASA Credibility Assessment Scale (CAS), the DoD M&S life-cycle, and the NATO Distributed Simulation Engineering and Execution Process (DSEEP), all of which embed traceability, documentation, and risk-informed justification. At the same time, contemporary challenges, ranging from hybrid cognitive simulations, agent-based socio-technical models, data-driven AI-enabled components, and real-time Digital Twins, have expanded the VV&A landscape, requiring new methods for validating machine learning models, testing adaptive behaviours, verifying reinforcement-learning agents, and accrediting simulations whose internal dynamics evolve over time. The state of the art therefore reflects a shift from static, deterministic VV&A practices toward flexible, multi-layered frameworks capable of handling stochasticity, human-in-the-loop effects, interoperability constraints, and complex Systems-of-Systems architectures, making VV&A an indispensable pillar for the deployment of high-fidelity simulation tools in safety-critical, defence, industrial, and training contexts.

3. System Architecture and Methodology

This chapter presents the methodological foundations and system architecture that support the integration of Modeling & Simulation and Artificial Intelligence within the COYOTE framework. The simulator has been developed to operate inside complex, safety-critical environments such as container terminals and industrial plants, where operational hazards arise from the interaction between humans, machinery, and uncertain environmental conditions. By adopting the MS2G paradigm, the architecture integrates discrete-event simulation, intelligent agent models, interoperable components, and learning algorithms within a unified virtual environment capable of supporting both training and decision-making.

Figure 11. Example of COYOTE interface

3.1. Conceptual Framework of the COYOTE Simulator

The conceptual framework of the COYOTE simulator defines the structural, functional, and methodological foundations that enable the integration of Modeling & Simulation, Artificial Intelligence, and Extended Reality within a unified environment tailored to port and industrial safety. The simulator has been developed following the MS2G paradigm, combining high-fidelity modeling, interoperable simulation, and immersive human–machine interaction. Its conceptual framework operationalizes how users, agents, data sources, learning components, and virtual environments interact to support training, experimentation, and decision-support in complex and hazardous settings.

COYOTE is designed as a virtualized operational ecosystem that mirrors real port terminals and industrial yards, capturing their multi-domain complexity across physical, cognitive, procedural, and cyber layers. The simulator supports heterogeneous missions, dynamic boundary conditions, and evolving risk profiles, allowing users to experience realistic operational scenarios while the system evaluates their performance, safety, and awareness.

The Figure below illustrates the high-level architecture guiding this conceptual design. It shows the integration of user platforms (PC, mobile, head-mounted displays), the virtual world, intelligent agents governing active objects, data sources, and the mission-generation logic. This architecture is not only a software design but the formalization of the underlying conceptual model that defines the behavior and interactions of the entire system.

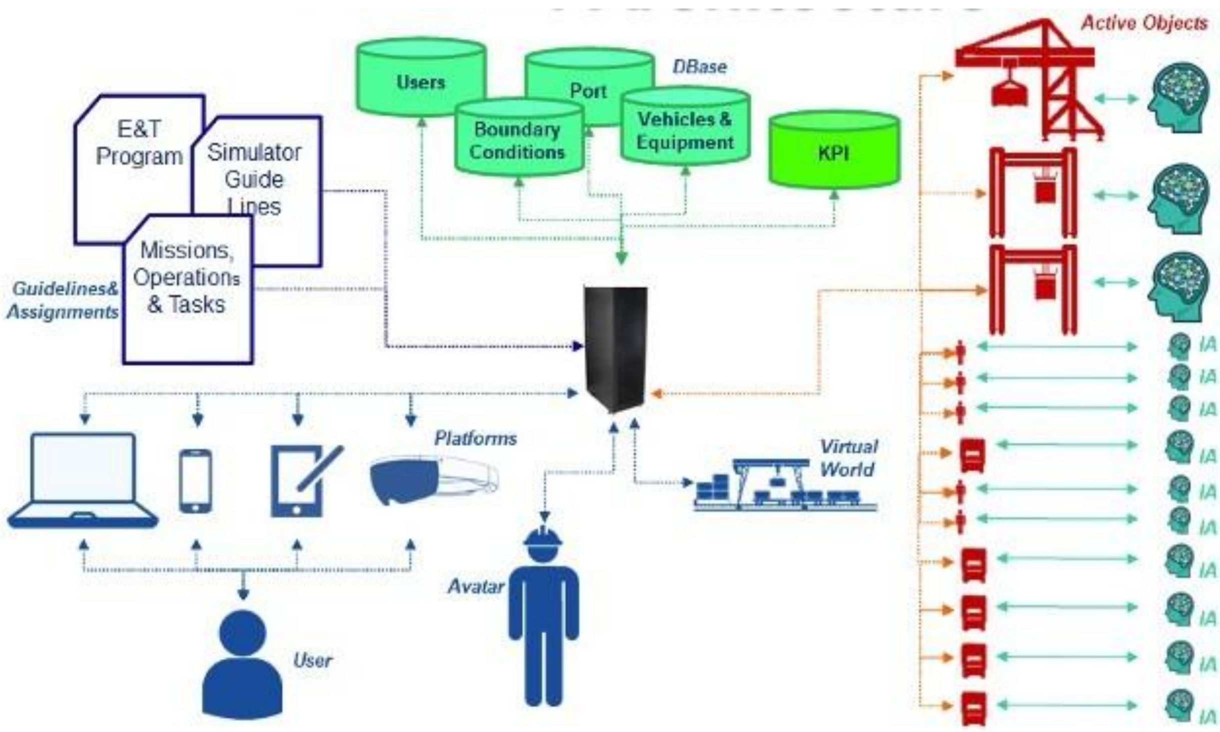

Figure 12. General Conceptual Framework of COYOTE

On the left side, the architecture begins with the definition of the Education and Training (E&T) Program and the Simulator Guidelines, which provide the methodological basis for generating missions, operational procedures, and task sequences. These guidelines flow into the core simulation engine, which interprets them as executable assignments for the user during the simulation session. Below this training layer, different interaction platforms are represented, including desktop computers, mobile devices, tablets, and extended reality headsets. These devices constitute the input layer through which the user accesses the system. The user interacts with the simulator through an avatar whose movements and actions are projected into the virtual world, ensuring a coherent representation of physical behavior within the simulated environment.

At the center of the architecture, the simulation core acts as the central processing node. It receives input from multiple data repositories that define the contextual and operational constraints of the scenario. The Users database contains operator characteristics and training profiles, while the Port database reproduces the physical and procedural boundary conditions of the terminal, including layout, navigation rules, and safety constraints.

The Vehicles and Equipment database contains models and dynamic parameters of cranes, straddle carriers, trucks, and other operational assets. The KPI database defines the performance indicators and safety metrics used to track operator behavior and compute Measures of Merit. These data repositories feed the simulation core in real time, enabling dynamic scenario configuration and continuous adaptation of mission conditions.

The virtual world, located on the lower right section of the figure, represents the three dimensional simulation environment in which operations unfold. All vehicle movements, container interactions, spatial constraints, visibility conditions, and environmental dynamics are computed within this digital representation. The virtual world is synchronized with the user avatar and with the intelligent agents governing active objects. On the far right of the figure, active objects such as quay cranes, yard cranes, trucks, and ground vehicles are shown, each linked to an associated intelligent agent. These agents are responsible for autonomous behavior, local decision making, path planning, obstacle avoidance, and interaction with the user. They continuously receive updated boundary conditions from the simulation core and produce adaptive responses based on operational constraints and real time events in the virtual world. User actions influence the behavior of intelligent agents, which can modify traffic flows and operational patterns. Environmental changes computed in the virtual world alter the user's perception and risk exposure, which is continuously recorded in the KPI subsystem. At the same time, the simulation core updates mission progress, evaluates performance metrics, and injects new tasks or events based on the training program

3.2. Risk as a Multi-Domain Variable

In the COYOTE simulation framework, risk is treated as a multi-domain variable whose value emerges from the combined contribution of physical, procedural, and psychological components. This formulation is grounded in a formal mathematical model capable of computing, in real time, both the objective danger present in the scenario and the subjective danger perceived by the operator. The physical domain models risk as a function of mechanical energy and proximity, while the procedural and psychological domains capture how operator behavior and sensory limitations distort the perception of hazards. By unifying these complementary dimensions, COYOTE provides a rigorous, quantitative account of how risk arises, evolves, and is cognitively interpreted in a dynamic operational environment.

The physical component, referred to as real or objective risk, is computed from first principles of mechanics. Each vehicle or moving hazard is characterized by its mass m, velocity v, and elevation h, which determine its instantaneous kinetic and potential energy.

The simulator evaluates

$E_k = \frac{1}{2} m v^2$and $E_p = mgh^+$, where $h^+ = \max{(h, 0)}$

avoids nonphysical negative potential values when the reference plane is above the object. These terms quantify the severity of a potential impact. Severity alone is insufficient to describe risk, which must also incorporate the probability of collision. For this reason, COYOTE introduces a probabilistic proximity model in which the impact probability is expressed as a decreasing logistic function of distance. Given the operator-to-vehicle distance d, and parameters d_{50}(the distance at which probability is 0.5) and k(the slope governing decay rate), the impact probability is defined as

$$P_{\mathrm{imp}}(d) = \frac{1}{1 + \exp{[k(d - d_{50})]}}.$$

Real risk is then formalized as the product

$$R_{\mathrm{real}}(d) = P_{\mathrm{imp}}(d)\,(E_k + E_p),$$

which provides a continuous, physically grounded scalar field indicating the instantaneous danger induced by each hazard in the environment. This value evolves as vehicles accelerate, decelerate, change direction, or disappear from the operator's vicinity.
Procedural risk is implicitly incorporated into this framework by modulating the spatial and temporal conditions under which the operator interacts with hazards. Deviations from prescribed safety procedures, such as entering restricted zones, approaching containers from unsafe angles, or performing inspections in areas of high traffic density, effectively reduce the operative distance $d$or alter its temporal evolution, artificially inflating collision probability. This reflects empirical evidence from port operations showing that human errors and violations frequently create hazardous configurations even when mechanical risks are low. Procedural contributions therefore act as amplifiers that deform the real-risk field based on behavioral decisions.

To compare with perceived risk we normalized the risk

$$R_{n,real} = \frac{R_{real} - R_{min}}{R_{max} - R_{min}}$$

$$R_{n,perc} = \frac{R_{perc} - R_{min}}{R_{max} - R_{min}}$$

To fully represent human experience, however, the framework must also model perceived risk, which does not depend exclusively on physical variables but on the operator's sensory and cognitive processing of the situation. COYOTE introduces sensory modulation functions that approximate human visual and auditory detection capabilities. The visual sensitivity component depends on the angular deviation θ between the hazard and the center of the operator's visual cone. A logistic visibility function

$$S_v(\theta) = \frac{1}{1 + \exp\,[\beta_v(|\ \theta\ | - \theta_{50})]}$$

describes how detectability decays with increasing angular displacement, where $\theta_{50}$is the half-visibility angle and $\beta_v$controls steepness. Similarly, an auditory sensitivity function

$$S_a(d, \varphi) = \frac{1}{(1+\alpha d)} \cdot \frac{1 + \cos\ \varphi}{2}$$

models the attenuation of sound with distance $d$and dependence on the azimuthal direction φ, with $\alpha > 0$governing decay. These functions represent the probability that the operator becomes aware of an approaching hazard based on sensory cues rather than physical proximity.

To construct perceived risk, COYOTE modulates the severity term of the real-risk formulation using a weighted combination of sensory indicators. A perceptual severity estimator

$$\tilde{E} = (E_k + E_p)\, w_0 + w_v S_v(\theta) + w_a S_a(d, \varphi),$$

with non-negative weights w_0, w_v, w_a, captures the cognitive intuition of danger formed by the operator. Perceived risk is then expressed as

$$R_{\text{perc}}(d, \theta, \varphi) = P_{\text{imp}}(d)\, \tilde{E},$$

ensuring that both real and perceived risk share the same probabilistic backbone while differing only in the severity modulation derived from sensory limitations and cognitive biases.

To enable quantitative comparisons, both real and perceived risk are finally rendered dimensionless through a normalization procedure. Given a realistic scenario set providing anchors $R_{\text{min}}$and R_{max}, the normalized indices are

$$R_{\text{n,real}} = \frac{R_{\text{real}} - R_{\text{min}}}{R_{\text{max}} - R_{\text{min}}}, R_{\text{n,perc}} = \frac{R_{\text{perc}} - R_{\text{min}}}{R_{\text{max}} - R_{\text{min}}}.$$

Normalization ensures invariance under changes of scale, preserves ordering, and makes the difference $R_{\text{n,real}} - R_{\text{n,perc}}$a meaningful indicator of risk underestimation or overestimation. When the normalized perceived risk is consistently below the normalized real risk, the operator is underestimating hazards, a condition strongly associated with collisions in empirical studies. Conversely, systematic overestimation degrades efficiency and increases the cognitive load of routine tasks.

Through this multi-domain formalism, COYOTE operationalizes risk as a physically interpretable, behaviorally modulated, and cognitively grounded variable. This allows the simulator to quantify how mechanical dynamics, procedural compliance, and sensory–cognitive processes interact to influence safety. The resulting model is capable of supporting sensitivity analysis, behavior classification, anomaly detection, and adaptive training protocols aimed at aligning perceived and actual danger, thereby improving both operational safety and human–machine collaboration in complex industrial environments.

Figure 13. POV of COYOTE Risk Evaluation System

4. Model Implementation and Simulation Design

The COYOTE simulator is built as a modular, multi-layer simulation framework that integrates heterogeneous modelling paradigms into a unified, real-time virtual environment. Its core relies on Unity 3D, which provides the physics engine responsible for rigid-body dynamics, collision handling, and continuous spatial evolution of all physical entities in the container yard. The system is structured as a three-tier architecture composed of a Core Simulation Engine, a Scenario & Mission Manager, and a Risk-Perception-Behaviour Layer. The Core Engine governs the deterministic evolution of the environment, advancing rigid-body motion, resolving collisions, and updating geometric relationships among operators, vehicles, and container blocks.

The Scenario & Mission Manager dynamically constructs each simulation instance by generating container layouts, initializing vehicle flows, spawning cranes and support machinery, and injecting environmental conditions such as fog or occlusion sources. It also manages mission logic, assigning inspection tasks, evaluating user interactions, and triggering mission outcomes. Above these components, the Risk–Perception–Behaviour Layer computes objective and perceptual risk fields, handles visual and auditory sensory modelling through ray-casting and occlusion analysis, and interfaces with the intelligent-agent subsystem. This layered structure enables COYOTE to blend physics-based realism, cognitive modelling, agent-based dynamics, and machine-learning behaviours into a coherent pipeline capable of reproducing complex port-yard operations with high fidelity.

Time Managament

Time in COYOTE is handled through a hybrid continuous–discrete time paradigm. Continuous simulated time governs the evolution of physical states such as positions, velocities, forces, energies, and dynamic hazard metrics. The simulator assumes that system states evolve on a real-valued temporal axis, following differential equations that describe instantaneous rates of change. Since real continuous computation is impossible on digital hardware, COYOTE approximates this behaviour by discretizing time into fixed intervals Δt, advancing the simulation through numerical integration at each frame.

Formally, for a system with state vector $\mathbf{x}(t)$, continuous-time dynamics satisfy:

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}),$$

which COYOTE approximates numerically using integration schemes such as Euler or Runge–Kutta. Euler integration uses the approximation:

$$\mathbf{x}(t + \Delta t) \approx \mathbf{x}(t) + \Delta t\, \mathbf{f}(\mathbf{x}(t)),$$

while higher-order Runge–Kutta methods compute intermediate slopes $\mathbf{k}_i$to yield:

$$\mathbf{x}(t + \Delta t) = \mathbf{x}(t) + \sum_{i=1}^{n} b_i\, \mathbf{k}_i,$$

offering improved stability and accuracy at higher computational cost. These techniques are used for vehicles, cranes, agent locomotion, and energy-based risk computation.

In parallel, COYOTE relies on Discrete-Event Simulation (DES) to manage mission logic, state transitions, task completions, alarms, and scenario updates. DES advances time by “jumping” to the next event:

$$t_{k+1} = \min\{\tau(e) \mid e \in \mathcal{E}, \tau(e) > t_k\},$$

where $\mathcal{E}$is the event set and $\tau(e)$is the scheduled time of event e. No updates occur between events, reducing unnecessary computation.

The hybrid time-management scheme ensures that physical realism is preserved through continuous integration while procedural logic and cognitive-model triggers are executed efficiently through DES. COYOTE may operate in paced mode (synchronized with wall-clock time) for training or in unpaced (accelerated) mode for experimentation, sensitivity analysis, or batch simulation runs.

Mathematical Formulation of the Simulation Components

At the core of COYOTE lies a set of mathematically rigorous models describing physical dynamics, risk computation, sensory perception, and decision processes. Each physical agent i(e.g., a truck or crane) is represented by a continuous-time dynamical system with state vector:

$$\mathbf{s}_i(t) = [\mathbf{x}_i(t), \mathbf{v}_i(t), \boldsymbol{\omega}_i(t)],$$

where $\mathbf{x}_i(t)$is position, $\mathbf{v}_i(t) = \dot{\mathbf{x}}_i(t)$is velocity, and $\boldsymbol{\omega}_i(t)$is angular velocity. These states evolve according to the Newton–Euler equations:

$$\begin{aligned}
&\dot{\mathbf{x}}_i(t) = \mathbf{v}_i(t),\\
&m_i\dot{\mathbf{v}}_i(t) = \mathbf{F}_i(t),\\
&\mathbf{I}_i\dot{\boldsymbol{\omega}}_i(t) = \boldsymbol{\tau}_i(t),
\end{aligned}$$

where $m_i$is mass, $\mathbf{I}_i$is inertia, $\mathbf{F}_i(t)$is the net force, and $\boldsymbol{\tau}_i(t)$is torque.
Hazard severity is computed using instantaneous energy:

$$E_{k,i}(t) = \frac{1}{2}m_i \parallel \mathbf{v}_i(t) \parallel^2, E_{p,i}(t) = m_i g h_i^+(t),$$

with $h_i^+(t) = \max(h_i(t), 0)$.
The probability of impact between operator and vehicle is modeled as a logistic function:

$$P_{\text{imp}}(d(t)) = \frac{1}{1 + \exp(k(d(t) - d_{50}))},$$

with $d(t)$their Euclidean distance.

COYOTE supports multiple categories of agents, each governed by distinct mathematical models.
Deterministic physical agents, such as moving vehicles, follow the Newton–Euler equations described above, optionally perturbed by stochastic noise to emulate human driving irregularities. Agent-Based Modeling (ABM) agents represent operational actors with discrete behavioural states. Each agent $i$occupies a state $q_i(t) \in Q$and transitions according to a Markov process:

$$\mathbb{P}(q_i(t+\Delta t) = q' \mid q_i(t) = q) = \Pi_{q,q'}(\mathbf{s}(t), \mathbf{p}),$$

where transition probabilities depend on spatial conditions, workload, proximity to hazards, and scenario constraints. System Dynamics (SD) agents, such as cranes, cycle-based machinery, or procedural operators, are governed by continuous-time rate equations:

$$\dot{\mathbf{y}}(t) = \mathbf{f}(\mathbf{y}(t), \mathbf{u}(t), \mathbf{p}),$$

where $\mathbf{y}(t)$represents operational stocks (e.g., queue lengths, machine states). Reinforcement Learning (RL) agents optimize decision policies $\pi(a \mid s)$using temporal-difference learning.

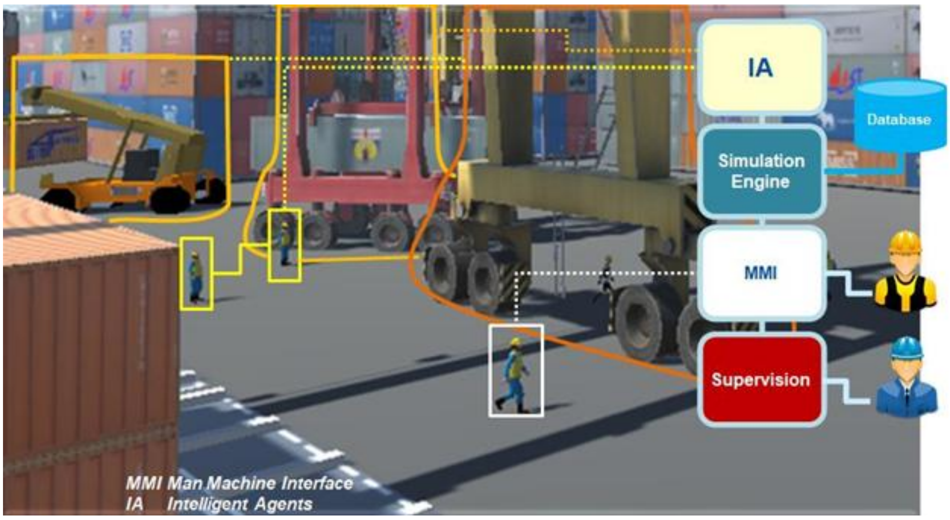

Figure 14. Example of Simulation Components

4.1. Scenario Representation

The scenario representation within COYOTE is designed to reproduce the complexity, variability, and dynamism of a real container terminal yard, providing a synthetic operational environment where human operators, autonomous vehicles, and environmental factors interact continuously. Rather than relying on static backdrops or scripted animations, the simulator constructs each scenario as a living, evolving system in which events emerge from spatial configurations, traffic patterns, and decision-making processes. This approach ensures that training sessions and experimentation datasets reflect the heterogeneity and unpredictability characteristic of modern port operations.

At the foundation of the scenario lies a detailed virtual reconstruction of the terminal yard, modeled with attention to its functional geometry and operational flows through the use of BLENDER and Unity3D. Blocks of stacked containers define the spatial structure of the environment, while internal roads, access lanes, and restricted zones delimit safe and unsafe trajectories available to the user. This layout is not a fixed asset but a configurable component: containers can be rearranged, block densities modified, and areas temporarily closed or expanded depending on the mission requirements. Through this flexibility, COYOTE accommodates a wide range of boundary conditions, from low-intensity daytime operations to congested nighttime shifts with degraded visibility.

Figure 15. View of the Synthetic Port Environment

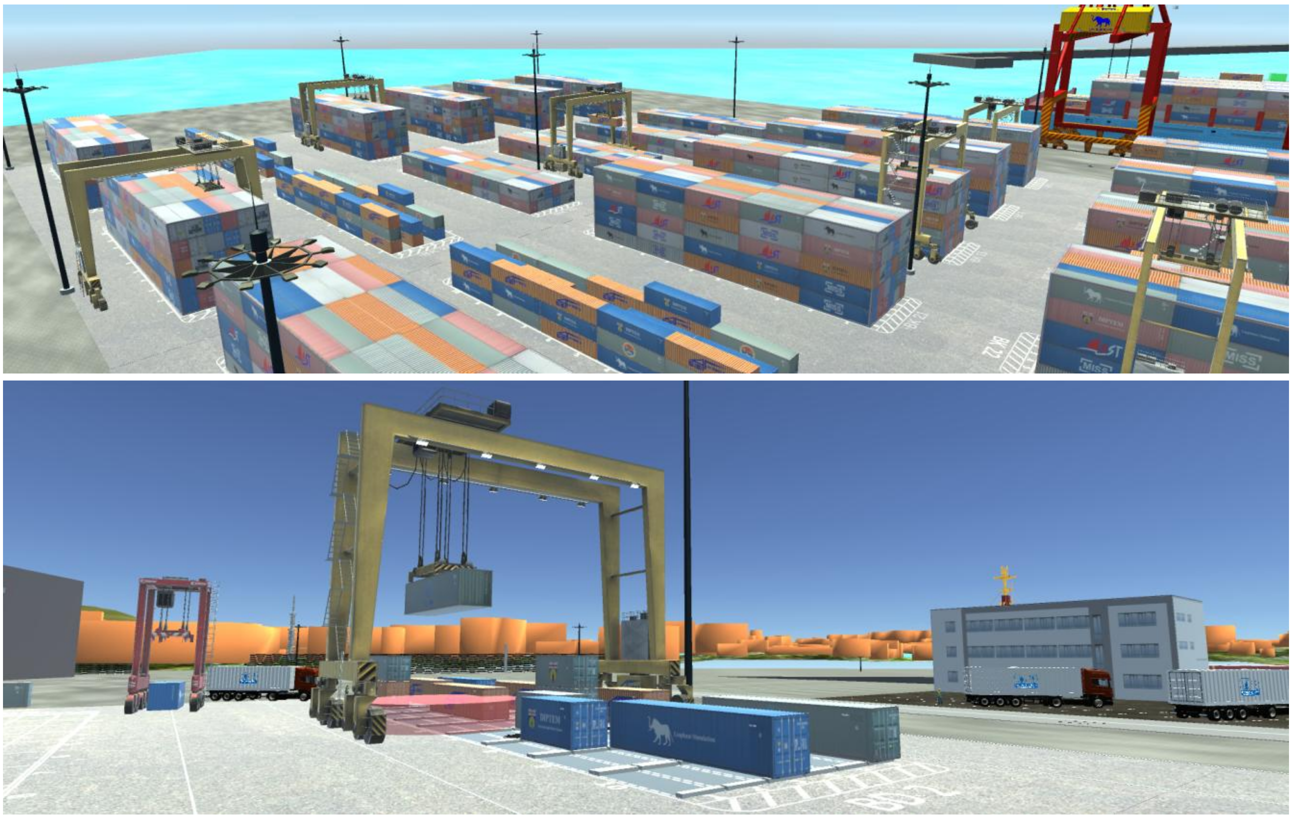

Figure 16. Different Angles of the Yard Terminal

The representation of the scenario within the COYOTE simulator originates from the need to identify an operational context capable of demonstrating the potential of Modeling & Simulation and Extended Reality technologies for improving safety, awareness, and efficiency in high-risk environments. During the early stages of the project, several alternatives were considered, including confined spaces and metallurgical industry settings, but constraints in budget and development time led to the strategic decision to focus on a single application domain. The port environment was selected because it encompasses a wide spectrum of safety-critical activities and presents an articulated combination of human, mechanical, and environmental interactions. Modern ports constitute complex microcosms where maritime operations, such as towing, mooring, and pilotage, converge with logistics activities conducted both at sea and inland, including the handling of hydrocarbons at buoys, quay crane operations, container transfer activities, reefer management, hazardous materials handling,

customs inspections, intermodal railway integration, and the circulation of heavy external vehicles.
Containerized cargo was identified as the most suitable reference domain due to its global diffusion and its capacity to generate results that can be extended to a large number of ports and inland terminals. Focusing instead on petrochemical traffic, although relevant, would have reduced the generalizability of the findings to a small number of specialized facilities. For this reason, the project established the container yard, particularly its hazardous materials zone, as the principal scenario. This area is characterized by high-density vehicle traffic, constrained visibility, environmental noise, and constant coexistence between operators on foot and autonomous or semi-autonomous handling equipment. The virtual world developed for COYOTE thus concentrates on these spatial and operational characteristics, but remains sufficiently general to be applied to other port facilities or adapted to additional training contexts developed by the Simulation Team.
Once the port environment was selected, a structured analysis was carried out in collaboration with domain experts from major terminals in Genoa and abroad. In particular, specialists from PSA Pra', Italy's primary import container terminal and part of the global PSA Group, and from Terminal Bettolo, a rapidly expanding terminal within the Port of Genoa, contributed to the definition of operational priorities, risk factors, and realistic constraints. Their expertise was complemented by the participation of the National Competence Center START 4.0, whose training division focuses on safety and digital transformation in port ecosystems, and by the involvement of industrial partners such as Thales, ABB, and Leonardo, who provided insight into technological vulnerabilities and emerging risk factors, including cyber threats and the introduction of wearable tracking devices. In parallel, analyses of past accidents and studies on marine-port risk scenarios were conducted to ensure that the selected scenario was aligned with real operational hazards and incident trends.
The spatial representation of the scenario required a precise strategy for the development of the 3D environment. Given the large extent of a container

terminal and the need for interactive real-time rendering, point-cloud acquisition techniques such as LiDAR scanning and photogrammetry were excluded due to their high computational cost and limited flexibility. Instead, a hybrid modeling strategy was adopted. Multiple modeling software packages were employed to exploit their complementary strengths, allowing rapid prototyping, detailed asset creation, and scalable modification of the environment. The use of heterogeneous data sources, ranging from satellite imagery and port maps to engineering layouts and equipment specifications, ensured spatial accuracy and realistic proportions. Part of the port environment was generated through semi-automated procedures to accelerate the construction of large areas while maintaining consistency, whereas high-detail modeling was reserved for assets with direct interaction potential, such as containers, vehicles, hazard markers, and inspection elements. This approach enables rapid reconstruction of different terminals and supports future business models in which new customers may request tailored environments reproducing their own infrastructure, such as the Port of Rotterdam or other major international hubs.
Performance considerations also guided the modeling workflow. To guarantee smooth visualization and user immersion, the environment had to maintain an adequate frame rate (≥30 FPS) across all platforms, including XR devices. This requirement made it essential for experienced 3D modelers to collaborate closely with simulation programmers, ensuring that polygon counts, shading, level-of-detail transitions, and collision geometries were optimized for interactive use. The resulting scenario achieves a balance between realism and computational efficiency, making it suitable for both training and experimentation.

The construction of the virtual scenario begins with a dedicated modeling workflow designed to reproduce port environments with both realism and efficiency. The first stage relies on Blender, an open-source 3D modeling platform that supports an advanced ecosystem of geographic plug-ins. A specialized Blender extension enables direct integration with

OpenStreetMap, which serves as a primary data source for terrain profiles, building footprints, road networks, and, when relevant, railway tracks. This integration allows COYOTE's virtual environment to be anchored to real geographic layouts whenever needed, reproducing existing terminals with high spatial fidelity. At the same time, even when a generic non-site-specific scenario is required, the use of OSM data enriches the virtual environment with realistic structural patterns, ensuring that its geometry reflects authentic operational constraints. Importing OSM data is largely automated, which significantly reduces modeling time; however, the imported elements often contain inaccuracies in scale, alignment, or road geometry. These imperfections necessitate a careful post-processing phase to remove errors, correct misplaced elements, and refine geometric continuity, using additional sources such as satellite imagery or Google Maps to guarantee spatial coherence.

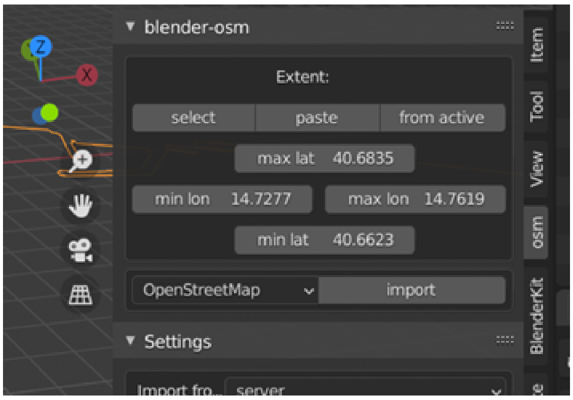

Figure 17. Injection of the geographic coordinates for the data extraction from OPen Street Map on Blender

Once the basic terrain and infrastructural network are imported, the modeling process proceeds with the incorporation of quay lines, a fundamental component of any port scenario. In generic or abstract yard environments these elements may be added manually, but when a real port must be reproduced, the workflow leverages Web Map Service (WMS) systems to import highly accurate coastline and berth-line data. Some WMS solutions allow the use of ENCs (Electronic Nautical Charts) in the official S-63 format, which are typically employed for vessel navigation, port entry, and berth approach. Because S-63 nautical charts represent authoritative hydrographic data, the extracted quay lines guarantee precise alignment with real port structures, ensuring that the virtual environment remains consistent with actual operational constraints. This level of accuracy is particularly relevant when the simulator is used for safety evaluation, procedural analysis, or operator familiarity training.

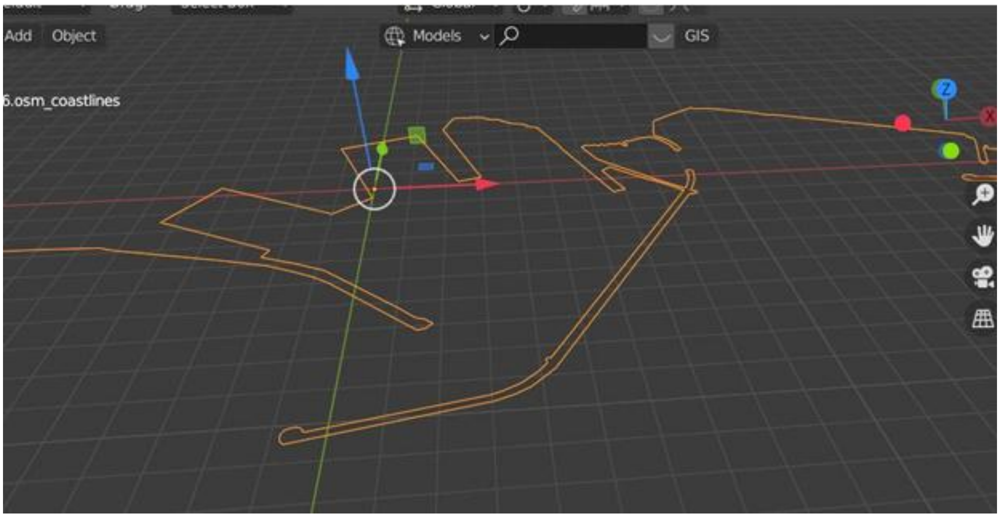

Figure 18. Example of port docks in Blender

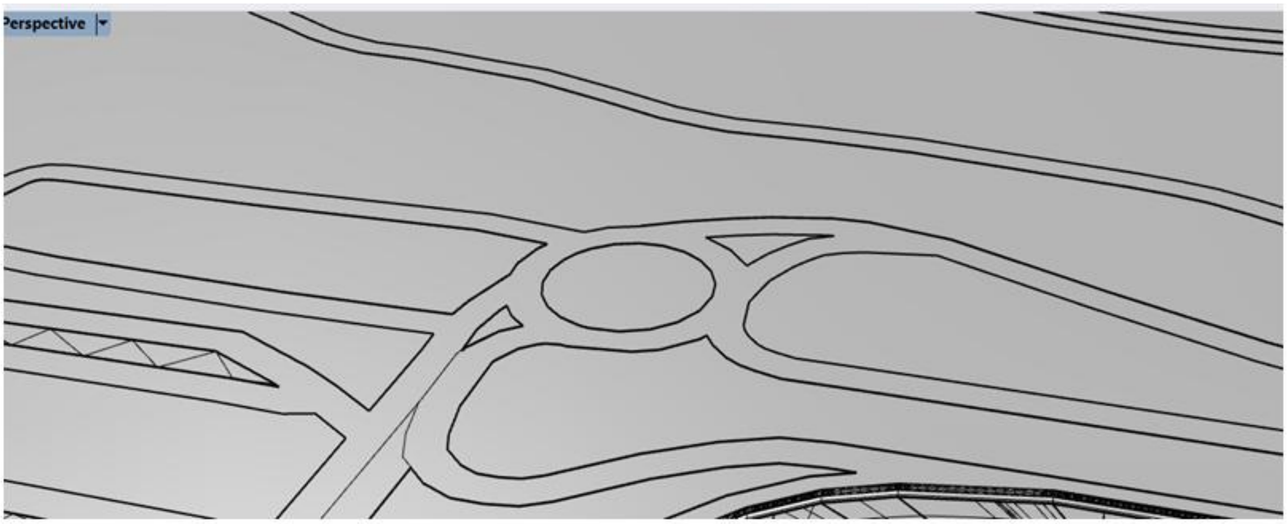

Figure 19. Example of elimination of overlap surfaces in Rhinoceros

After the geographic and maritime boundaries are established, attention shifts to refining the road network. Roads imported from OpenStreetMap frequently contain geometric inconsistencies such as inaccurate widths, misaligned segments, or overlapping surfaces. Overlaps, in particular, pose a problem for real-time rendering inside the simulation engine, as they cause flickering artifacts that disrupt user perception and reduce visual comfort.

To address these limitations, the workflow transitions to Rhinoceros, a modeling platform better suited for precise technical editing. Within Rhinoceros, designers adjust road geometries, correct alignments, eliminate overlapping polygons, and rebuild surfaces to ensure that the final network is continuous, visually stable, and compatible with interactive rendering requirements. This step not only improves visual quality but also ensures that agents operating within the simulator, such as trucks, straddle carriers, and automated vehicles, can rely on accurate and consistent navigation surfaces.

The development of the detailed elements of the 3D environment begins once the geometric structure of the port area has been established. Horizontal signage, which plays a crucial role in defining circulation patterns and safety zones within the yard, is produced through high-resolution image editing in Photoshop. These markings, lane delimitations, turning arrows, pedestrian paths, hazardous-goods indicators, and operational boundaries, are exported as texture maps and applied to the ground mesh through UV mapping. This process allows each texture to be precisely "stitched" onto the three-dimensional geometry, preserving scale, alignment, and visual clarity. Within the rendering engine, additional transparent overlays can be dynamically activated or recolored to adapt the appearance of the yard to different operational or training scenarios, enabling the simulator to replicate day-night cycles, seasonal lighting conditions, or specific visual cues required for experimental runs.
As described previously, buildings are automatically imported into Blender from OpenStreetMap data. However, before these structures can be integrated into Unity3D, their materials must be converted. The node-based shading system used in Blender is not directly interpreted by Unity, requiring a substitution step in which building textures are replaced with compatible Unity materials. This operation is efficient because the project relies on a dedicated internal texture repository: an extensive database of pre-processed façade textures, roofing materials, industrial surfaces, and architectural patterns prepared at the outset. By drawing from this database,

the conversion of Blender models into Unity-ready assets becomes rapid and standardized, ensuring visual consistency across different port environments.

Figure 20. Example of View in Unity 3D

Once the textures have been reassigned, buildings are exported to Unity3D together with all surface UV maps generated during the modeling phase. Within Unity, additional material layers, such as normal maps, roughness maps, metallic maps, or ambient occlusion maps, can be applied to enhance realism by introducing micro-geometric features, depth cues, and lighting responses. Developers can also modify material parameters directly inside the Unity inspector, enabling fine control over reflectivity, surface smoothness, and color profiles depending on the lighting setup of the scenario.

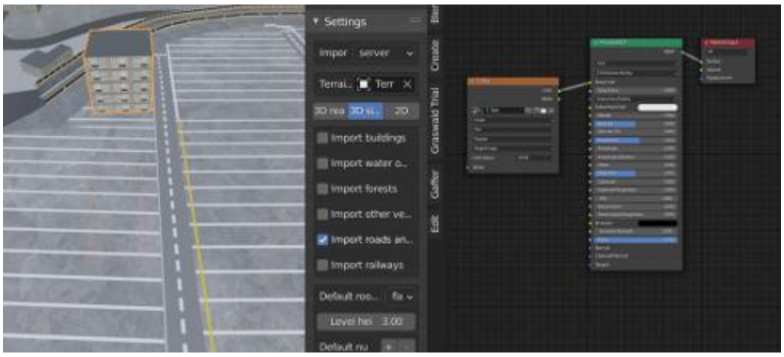

Figure 21. Example of texture replacement done in Blender and visualization

After buildings are integrated, further environmental detail is progressively added to enrich the fidelity of the yard. Access gates, fences, vegetation patches, container stacks, perimeter barriers, road dividers, and technical installations are introduced according to the specific needs of the scenario. These elements are modeled either in Blender or in Rhinoceros, depending on the geometric complexity required, and then imported into Unity where their materials are assigned and calibrated. Importantly, this modeling effort is performed only once: from the second environment onward, all previously created assets exist as Unity Prefabs. Prefabs embed the mesh, materials, colliders, and metadata required for simulation, allowing them to be reused and placed directly into new scenes with minimal effort. This not only accelerates the creation of additional port environments but also ensures uniformity in object appearance, performance, and behavior across different simulation scenarios.

The import of the three-dimensional port environment into Unity3D is not performed as a single operation but is instead carried out in a staged and methodologically controlled sequence. This incremental approach is necessary both for practical reasons, since it allows developers to inspect and validate each structural component, and for computational considerations linked to lightmap generation. The workflow begins by importing the quay structures, which constitute the foundational element of the scenario and define the alignment of the entire yard. Once the quay geometry is integrated, the plaza surfaces, container blocks, and primary ground meshes are added, followed by buildings, road networks, and any ancillary infrastructural elements. Managing these elements in separate batches allows the development team to identify geometric inconsistencies, incorrect pivots, or misaligned surfaces before the environment becomes too complex to debug efficiently.
Lightmap computation is performed directly within Unity3D to ensure consistent illumination across all static assets. Although Blender also supports light baking, Unity provides finer control over lightmap resolution, UV unwrapping for light channels, and the baked-lighting pipeline used by the simulation engine. By marking imported meshes as static and recalculating global illumination within Unity, developers can immediately detect errors such as overlapping surfaces, flipped normals, or incorrect anchoring, all of which would otherwise cause artifacts or visual instability during runtime. This phase also enables the refinement of material

properties, since imported textures alone are insufficient to capture nuances such as reflectivity, roughness, subsurface scattering, or ambient occlusion. Additional texture layers, such as alternative concrete types, soil patterns, grass patches, or specialized industrial materials, are created or adjusted directly within Unity to enhance realism and visual coherence.
After the foundational components are established, the remaining scene objects are progressively introduced. Operational assets such as container stacks, quay cranes, yard cranes, bollards, traffic signs, safety markers, or specialized port equipment are imported as individual Unity Prefabs, each equipped with the appropriate materials, colliders, and metadata required for their function in the simulation. These assets contribute to the operational authenticity of the scenario and provide the spatial cues essential for training and risk-perception studies. Decorative or environmental elements, including vegetation, small architectural features, or roundabouts, are also added where appropriate to increase the ecological validity of the environment and reduce the perceptual gap between the virtual yard and its real-world counterpart. Through this phased process, Unity3D becomes the consolidation layer where all geometries, textures, lighting, and dynamic objects converge, producing a coherent and operationally credible port scenario suitable for real-time training and experimentation.

The 3D models developed for the COYOTE scenario were produced following the modeling strategy outlined in the previous section, and the resulting visual environment reflects a high level of geometric accuracy and realism. In addition to the terminal’s core operational area, external surroundings and global illumination effects were incorporated to strengthen perceptual fidelity and reduce the cognitive gap between the virtual and real port environment. The selected scenario represents a container terminal and focuses specifically on the activities carried out in this context, with particular emphasis on the behavior and safety of personnel operating in the yard and along the quayside near stacked containers.

The identification of the reference environment was carried out in collaboration with port safety experts, which led to the selection of two of the main terminals of the Port of Genoa, PSA Pra' as the primary template for scenario design. These terminals were chosen because they present a representative range of operational configurations and because they are among the busiest and most technologically advanced nodes in the Mediterranean. The scenario therefore centers on typical accident dynamics previously identified in safety analyses, particularly collisions, near-misses, cargo-handling incidents, and hazard propagation in areas where dangerous goods are stored or manipulated. By modulating the properties of the cargo, ranging from ordinary containers to units carrying hazardous materials, the simulator can generate different risk intensities and reproduce the variability encountered in real terminal operations.
COYOTE constructs its virtual environment and simulation models by recreating the operational processes performed inside the terminal and, more specifically, in the yard. Intelligent agents govern the motion and actions of all ground-handling vehicles, reproducing realistic behaviors of straddle carriers, yard trucks, reach stackers, and service equipment. These agents follow autonomous navigation rules, pathfinding logic, and operational priorities that mirror those of real machinery, producing emergent traffic patterns and dynamic risk configurations. The intelligent agents do not control only the mechanical assets: they also coordinate the activities of ground personnel, simulating both routine operations, such as customs inspections, reefer servicing, vehicle and yard maintenance, support to container handling, refueling checks, shift changes, and scheduled inspections, and exceptional cases, including fault investigations or the assessment of suspicious conditions. This dual-level modeling, integrating autonomous machinery and virtualized human operators, is essential for capturing the real structure of terminal workflows, where human oversight and machine activity intersect continuously.
Container blocks are represented as dynamic elements within the simulation, each characterized by physical attributes and safety-related properties. Containers may exhibit different states, including normal

conditions, minor or major leaks, structural damage, or thermal anomalies, depending on the mission and the risk scenario being studied. Hazardous elements such as liquid spills, toxic-material leaks, fires, explosions, and unstable loads are not static objects but are simulated as evolving conditions that interact with the environment and with user actions. This makes the container blocks active contributors to the risk configuration of the yard, influencing operator decisions and shaping the unfolding of the mission.

Figure 22. The COYOTE Virtual World and the dangerous goods area on the forecourt

The COYOTE scenario involves people, cranes, vehicles, and containers in the yard and docks, with respect to safety issues and takes into account the reality corresponding to the port observers involved, which can be easily extended to others.

Figure 23. Rubberized Yard Transtainers

Figure 24. Reach Stacker

Figure 25. Fifth wheels and trucks

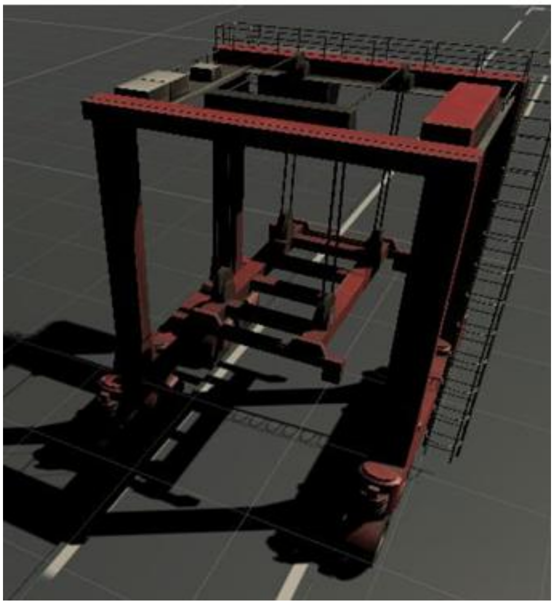

Figure 26. Straddle Carriers

Figure 27. Harbour Cranes

4.2. Cyber Representation

The progressive digitalization of port terminals has fundamentally transformed their operational landscape. Modern ports are no longer solely physical infrastructures composed of cranes, docks, and cargo flows; they are complex cyber-physical ecosystems deeply integrated with networked technologies, sensing devices, operational software, and automated decision systems. This digital shift has introduced undeniable advantages in terms of efficiency, situational awareness, and coordination, but it has also exposed port infrastructures to an unprecedented range of cyber-threats.

In contemporary scenarios, cyberattacks have emerged as a major risk vector for critical infrastructures. Historically, ports have long been considered strategic targets during conflicts, primarily through physical interference. Today, however, the most severe disruptions often originate from the digital domain. Over the last decade, attacks targeting Information Technology (IT) and Operational Technology (OT) systems have increased dramatically, affecting port management systems, SCADA architectures, and IoT networks. Incidents such as large-scale blackouts, ransomware attacks on shipping companies, and targeted operations against industrial control systems demonstrate how cyber aggression is now integral to hybrid forms of warfare. These attacks leverage anonymity, asymmetry, and plausible deniability, making both attribution and response increasingly complex.

Given this evolving threat landscape, any realistic simulation of port activities must incorporate a robust representation of the cyber domain. It is no longer sufficient to model physical accidents, hazardous material handling, or vehicle-to-person interactions. A modern port simulator must embody the intertwined cyber and physical layers, capturing how cyber vulnerabilities propagate into the physical world and how operational disruptions may create new cyber weaknesses. This dual representation is essential for training personnel, assessing resilience, and developing robust response strategies. The cyber representation of the environment builds on the port-terminal operational scenario already established in COYOTE and related MS2G frameworks. The simulation reproduces key areas of the port (including the yard, docks, crane operating zones, pipelines, and hazardous-goods storage areas) while extending them into a fully connected cyber ecosystem. This includes IoT sensors, control networks, communication devices, SCADA-like elements, and operational software modules.

Two major terminals in Genoa (PSA and Bettolo) served as reference archetypes for modelling both the physical infrastructure and the corresponding digital layers. In these terminals, the highest-risk zones involve dense interactions between personnel and moving equipment, areas of intense noise and low visibility, and locations dedicated to dangerous goods. These areas are now augmented with detailed cyber infrastructure, reflecting their increasing dependence on connectivity, automation, and sensor-driven processes.

The resulting scenario is therefore cyber-physical by design: humans, cranes, vehicles, containers, sensors, and digital networks coexist within the same simulation space, and their interactions dynamically influence both safety metrics and system behaviour.

To mirror the evolution of real maritime logistics, the simulation integrates a network of IoT devices embedded throughout the virtual terminal. These devices replicate real-world sensor functionality and monitoring capabilities:

- tracking container movement and placement
- measuring environmental conditions (temperature, humidity, visibility)
- reporting equipment status, energy consumption, and alarms
- monitoring pipeline conditions and dangerous-goods storage
- enabling remote diagnostics and predictive maintenance

Data collected by these virtual sensors is continuously injected into the simulation logic and made available to Intelligent Agents (IA). This coupling enhances realism and introduces operational interdependencies that closely reflect modern port operations. It also exposes the simulation to a key vulnerability of real ports: the same IoT network that improves efficiency becomes a potential attack surface for malicious cyber actors.

The cyber representation incorporates an extensive library of cyber-threats that can target the virtual port's digital infrastructure. This includes both common and advanced attack vectors:

- malware and ransomware propagation
- unauthorized access to IoT nodes
- manipulation of sensor data
- disruption of automated crane operations
- interference with pipeline monitoring systems
- SCADA-layer compromise
- Distributed Denial-of-Service (DDoS) attacks on port information systems
- Advanced Persistent Threats (APTs) corresponding to state-sponsored activities

The simulation reproduces the cascading effects of these attacks. For example:

a corrupted sensor may feed false data to an IA-controlled crane, leading to mishandling of a container; a network-level ransomware attack may block access to digital manifests, delaying or misdirecting operations; a SCADA compromise could disable an entire section of pipeline valves, generating a physical hazard. The cyber layer is built to be modular and extensible. New threat vectors, sensor types, network architectures, and defensive techniques can be incorporated without modifying the overall architecture. This flexibility ensures that the model remains relevant as port technologies evolve and as new forms of cyber aggression emerge. The environment can represent different terminal layouts, IT/OT architectures, and operational workflows, supporting comparative studies across port configurations or regions.

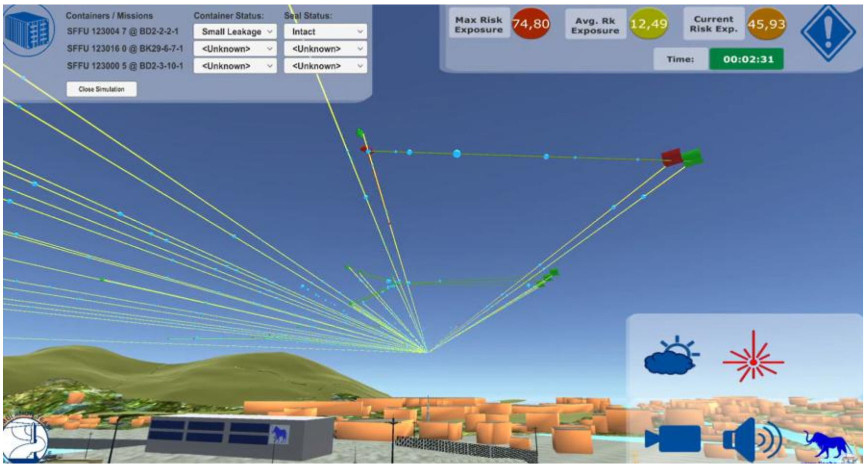

Figure 28. Representation of the Cyber assets in COYOTE

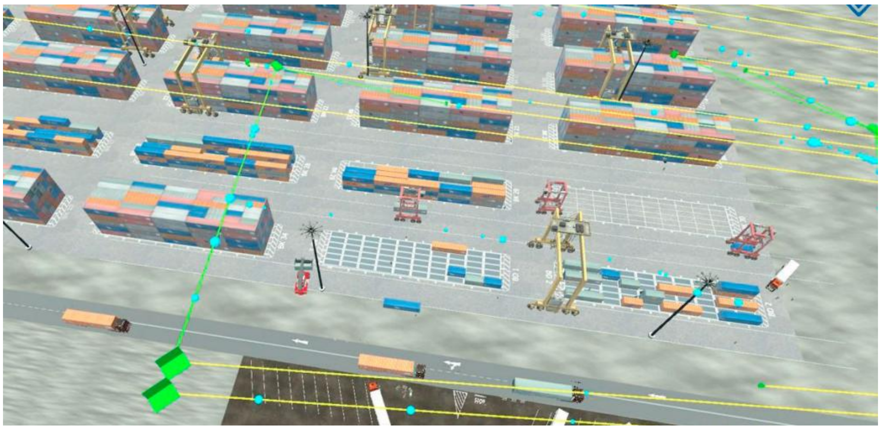

Figure 29. Yard of the Terminal Container in COYOTE Simulator

4.3. Navigation System

The navigation system developed within the Unity synthetic environment constitutes a central component of the simulation architecture. It enables autonomous movement of vehicles, cranes, agents, and dynamic objects within the virtual port terminal while guaranteeing coherence with real operational constraints. The system is designed to reproduce, with high fidelity, the flow of port activities, including the movement of straddle carriers, trucks, forklifts, ground personnel, and external agents interacting with the yard.

The approach integrates rule-based behaviour, graph-based path planning, sensor simulation, collision avoidance, and reactive logic under a unified computational framework. The navigation system ensures that each entity behaves consistently with both the physical layout of the terminal and the safety requirements imposed by the real environment it represents.

The navigation system relies on a combination of three modelling layers:

1. Static Navigation Layer, which encodes the geometry of the terminal, including roads, lanes, container blocks, restricted zones, and access points.
2. Dynamic Navigation Layer, which manages temporary obstacles, moving vehicles, changes in traffic density, and the evolution of the scenario over time.
3. Behavioural Layer, which regulates decision making, priority rules, safety constraints, perception limits, and mission-specific goals for each agent.

These layers interact continuously during simulation, guaranteeing that routing decisions remain coherent with both the physical model and the operational context.

The port environment is modelled as a directed weighted graph. Each walkable or drivable area is subdivided into nodes connected through edges that represent possible routes.
The nodes include:

- road intersections
- lane segments
- entry and exit points
- container block access locations
- crane interaction nodes
- maintenance and inspection points

Each node contains metadata such as allowed vehicle types, speed limits, turning restrictions, risk level, and visibility penalties.

Edges represent possible transitions and encode:

- nominal travel distance
- expected travel time
- penalty for restricted areas
- dynamic costs based on traffic or temporary blockages
- risk modifiers linked to the risk field of the simulator

The resulting graph is updated each simulation tick to reflect scenario dynamics.

Path planning is performed through an extended A star algorithm. This algorithm was selected for its efficiency and its deterministic nature, which is important for reproducibility and V and V processes.
The cost function is extended to incorporate operational constraints:

$$f(n) = g(n) + h(n) + C_{\text{risk}}(n) + C_{\text{traffic}}(n)$$

where

- $g(n)$is the accumulated cost from the start node
- $h(n)$is the heuristic distance to the target
- $C_{\text{risk}}(n)$penalises nodes in high risk zones
- $C_{\text{traffic}}(n)$increases with local congestion

This approach allows vehicles to choose safer and more efficient routes in different conditions.

The collision-avoidance subsystem leverages Unity's built-in physics components (Colliders) along with raycasting queries to detect obstacles and initiate avoidance behaviours. It is tightly integrated into the navigation system and ensures safe trajectories for mobile agents (vehicles, cranes, personnel) while preserving realism and performance.

Core components

- Colliders: Each agent and environment object is equipped with a Collider (e.g., BoxCollider, SphereCollider, MeshCollider) that defines its physical boundary within the physics engine. These provide the fundamental geometry for collision detection.
- Raycasts: At runtime each agent emits one or more raycasts from its current position, in the forward movement direction and/or in sampling directions (left, right, downward), using Unity's Physics.Raycast API. The raycast returns a RaycastHit indicating distance to the nearest collider in the cast direction, if any.
- Layer Masks and Filtering: To improve performance and avoid self-collisions, the raycasts apply layer masks so that the agent's own collider is excluded and only relevant obstacle layers are considered.
- Avoidance Logic: When a raycast detects an obstacle within a predetermined threshold distance, the agent transitions into an avoidance state. The logic may include:
  - slow down or brake
  - steer away (adjust direction vector)
  - trigger a local path re-plan (invoke global planner for a new route)
  - maintain safe following distance
- Sensor Configuration: Multiple raycasts may be configured as part of a "sensor array" (e.g., frontal ray, angled side rays) to replicate vision cones or sonar-like detection. The number of rays and their angular spread are configurable for agent types.

- State Machine Integration: Collision avoidance is embedded within the agent's finite-state behaviour controller. Typical states:
  - Navigating → normal movement along global path
  - ObstacleDetected → raycast hit threshold exceeded
  - Avoiding → execute avoidance steering/braking
  - ReplanRequested → request new path and return to Navigating once resolved

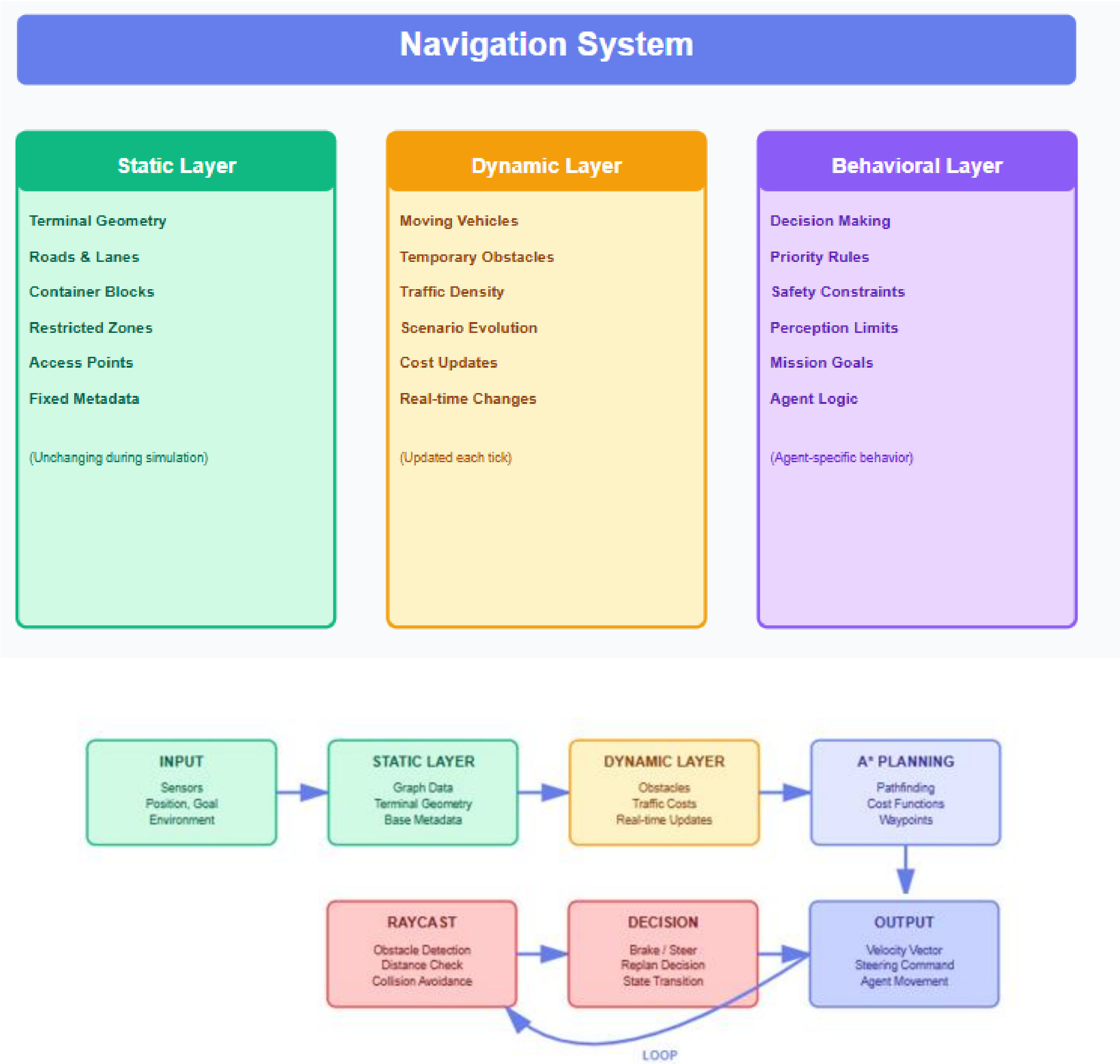

Figure 30. Navigation System Architecture

4.4. Task and Objectives

The mission logic implemented in COYOTE is designed to reproduce, with high operational fidelity, the activities that personnel routinely conduct within a container terminal yard, while embedding the cognitive, perceptual, and procedural challenges that characterize real-world inspection operations.

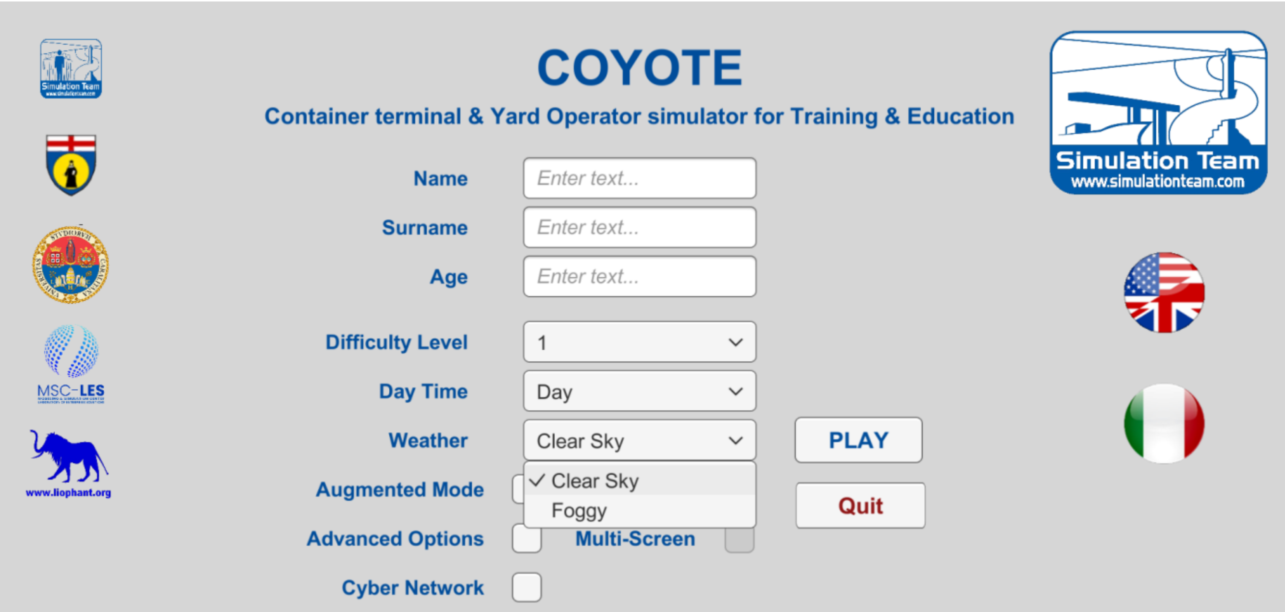

Figure 31. Initial GUI for Scenario Setting

Each training session places the operator inside a dynamic simulation scenario in which the primary objective is to locate and inspect three specific containers dispersed across the terminal yard. These containers, identified through a unique alphanumeric ID (e.g., *SFFU 123001*), are associated with a positional reference that follows the standard yard nomenclature: **Block - Row - Slot - Tier**. For example, the code *BK23-3-2-1* denotes Block 23, Row 3, Slot 2, Tier 1 (ground level). Containers storing hazardous materials are grouped in blocks whose identifiers begin with *BD*, whereas standard cargo is located in blocks prefixed by *BK*. This convention mirrors the actual terminal organization, assisting the user in navigating the yard with spatial logic consistent with real operations.

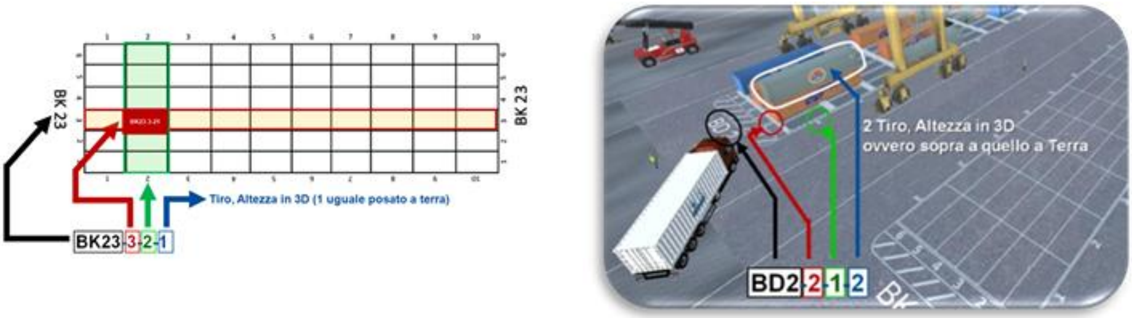

Figure 32. Block - Row - Slot - Tier

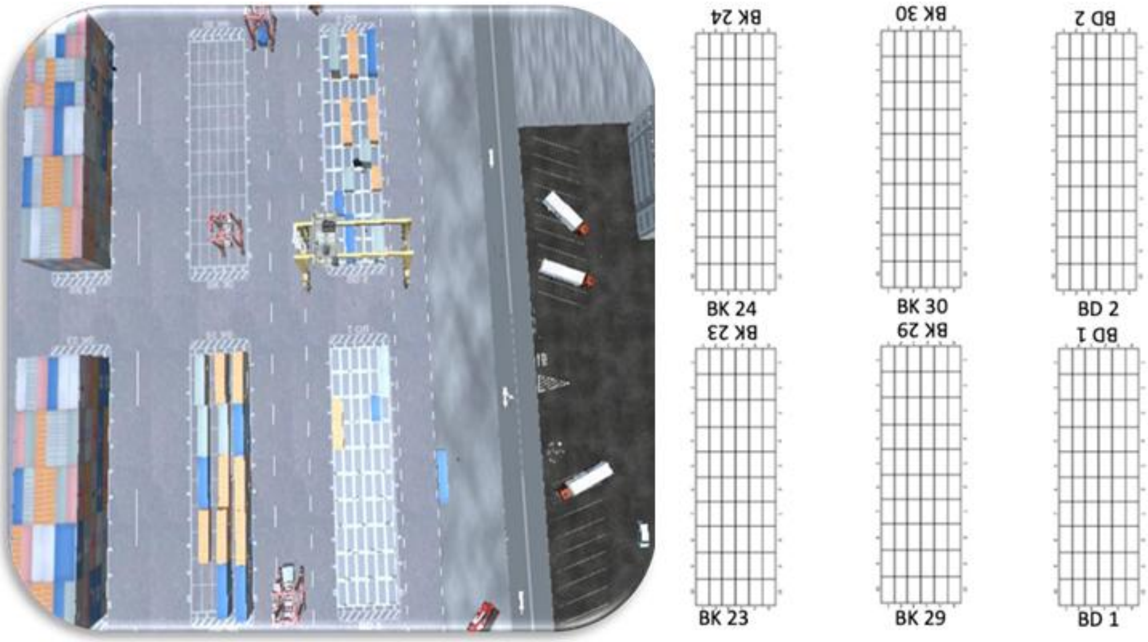

Figure 33. Block identification scheme

During the mission, the operator must locate each assigned container based on the positional data provided in the user interface. Upon reaching a container, the trainee performs a visual inspection to determine whether the unit is intact or exhibits anomalies such as small or medium-sized leaks. The user records this observation through a dropdown menu dynamically linked to the container's identifier displayed on the top-left interface panel. This interaction ensures that the inspection task is both perceptual and procedural: the operator must correctly interpret the visual cues in the environment, navigate the yard safely, and accurately classify the container's condition.

Figure 34. Different Leakeges

Figure 35. Menu for Task Completion

Once all three containers have been inspected, the mission concludes by selecting *Close Simulation* in the interface. At this point, the simulator automatically stores the complete dataset of the session, including task outcomes, operator trajectory, risk exposure metrics, collision near-misses, perceptual indicators, and timing information. A synthetic summary of key performance indicators is presented on a final screen to provide the trainee with immediate feedback. The mission may be repeated, either with the same identification parameters or by launching a new scenario, enabling iterative learning and comparative analysis across trials.

Figure 36. Door with Bolt Seal

Mission difficulty scales according to predefined levels. Higher difficulty configurations increase traffic density by adding more moving vehicles (e.g., trucks, yard tractors, straddle carriers), degrade visibility through environmental modifiers (fog, low lighting, glare), or introduce inconsistencies between the instructed container position and its actual location. In the latter case, the trainee must search the yard more extensively, reflecting operational scenarios in which the terminal information system contains outdated or erroneous data. Under poor visibility, acoustic alarms and flashing visual indicators are triggered to emulate the warning systems commonly used in real terminals.

Figure 37. View of the Yard Terminal

Throughout the mission, the operator must maintain situational awareness and adhere to procedural safety constraints. The yard contains several dynamic hazards: moving trucks, cranes transporting suspended loads, container stacks that limit visibility, tight corridors, and areas restricted due to dangerous goods. The user must avoid collisions, maintain safe distances from active equipment, and refrain from passing under suspended loads. The simulator continuously monitors spatial proximity, occlusion states, reaction patterns, and locomotion speed to determine real and perceived risk exposure. If the trainee is struck by a vehicle or crane, the mission is automatically classified as failed, and the simulator saves the data to enable post-session analysis. Navigation within the scenario is achieved through keyboard input (arrow keys or the standard WASD configuration). Additional on-screen controls allow forward and backward movement, turning, and head-orientation adjustments through right-click-and-drag actions. A sprint function (key *J*) simulates rapid movement, adding a trade-off between speed and risk exposure. These navigation features are intentionally simple to ensure usability across operator skill levels while providing full freedom of movement for realistic exploration of the yard.

COYOTE's task structure explicitly connects mission execution to measurable learning outcomes. The simulator computes a broad array of

Key Performance Indicators, spanning accuracy of inspection, time-to-completion, procedural compliance, risk exposure, distance to hazards, number of near-misses, collision events, and perceived versus real risk interpretation. These metrics enable a multidimensional evaluation of operator performance, supporting both educational objectives and experimental research.

To complement these measurements, the simulator includes an *Augmented Mode* that overlays real-time risk awareness cues into the environment. When activated, this mode visualizes hazard sources through red lines indicating the direction and distance of nearby risks. This feature serves as an instructional aid, teaching operators to identify threats more effectively and to internalize safe distancing strategies. It also enables researchers to quantify the effect of augmented cues on user performance and risk-perception alignment.

Figure 38. End of Simulation and Results

4.5. Integration of Reinforcement Learning Agents

The behavioural architecture of COYOTE relies on a heterogeneous population of simulated entities, each endowed with different levels of autonomy, reasoning capacity, and adaptability. Although all such entities fall under the broad category of "agents," the simulator differentiates clearly between Intelligent Agents and Reinforcement Learning (RL) Agents, since they rely on fundamentally different modelling paradigms, decision-making mechanisms, and computational objectives. This distinction is essential to understand how COYOTE reproduces realistic port-yard behaviour while also enabling adaptive, data-driven optimisation.

In COYOTE, Intelligent Agents refer to entities whose behaviour is governed by *explicitly encoded reasoning structures*, which may include rules, finite-state machines, utility functions, procedural logic, or simplified models of human cognition and perception. These agents exhibit autonomy and context-awareness, but their intelligence is derived from the modeler, not from trial-and-error learning. They operate through a structured perception–decision–action loop:

$$
\begin{array}{lll}
o_t & = \mathcal{P}(s_t) & \text{(perception)} \\
b_t & = f(b_{t-1}, o_t) & \text{(belief update)} \\
a_t & = \mathcal{D}(b_t) & \text{(decision)} \\
s_{t+1} & = \mathcal{A}(s_t, a_t) & \text{(action execution)}
\end{array}
$$

Here, the decision function $\mathcal{D}$may include:

- rule-based logic ("if distance < threshold, stop"),
- heuristics (follow lanes, avoid cranes, yield to trucks),
- utility maximization (preferring actions minimizing risk or energy),
- task-oriented procedures (inspection routines, crane cycles),
- cognitive approximations (awareness fields, reaction times).

These agents do not modify their behaviour from experience; instead, they reproduce *known operational patterns* or *desired behaviours* in a controlled manner.
They are ideal for:

- modelling predictable machinery (cranes, automated gates),
- replicating standard operating procedures,
- enforcing safety rules and terminal logic,
- populating the yard with “baseline traffic” or typical operator behaviour,
- guaranteeing simulation stability and realism.

In essence, Intelligent Agents in COYOTE represent designed intelligence, driven by explicit models, structured, explainable, and deterministic or stochastic depending on the modelling need.

By contrast, Reinforcement Learning Agents are agents whose behaviour emerges from learning rather than explicit modelling. Their intelligence is not encoded by the designer but learned from interacting with the environment
Reinforcement Learning (RL) provides COYOTE with adaptive agents capable of improving their behaviour through repeated interaction with the simulated port-yard environment. Unlike deterministic or rule-based agents, which follow predefined decision policies, RL agents learn optimal actions directly from experience by maximizing long-term performance while minimizing exposure to risk. The mathematical foundation of RL in COYOTE builds upon the theory of Markov Decision Processes (MDPs), augmented with continuous-time risk fields, perceptual uncertainty, and safety-based reward shaping tailored for high-risk industrial contexts.
Each RL-enabled agent (e.g., a mobile robot, a yard vehicle, or a virtual operator assistant) interacts with the simulation according to the tuple:

$$\mathcal{M} = (\mathcal{S}, \mathcal{A}, P, R, \gamma),$$

where:

- $\mathcal{S}$ is the state space representing physical, cognitive, and environmental variables.
- $\mathcal{A}$ is the action space available to the agent (e.g., move, stop, steer, accelerate, avoid).
- $P(s' \mid s, a)$ defines the transition model driven by Unity's physics and ABM procedures.
- $R(s, a)$ is a reward function incorporating safety, task progress, and risk penalties.
- $\gamma \in (0,1)$ is the discount factor encoding long-term planning.

In COYOTE, the agent state is typically a high-dimensional vector:

$$s_t = \left[\mathbf{x}(t),\ \mathbf{v}(t),\ d_i(t),\ R_{\text{real}}(t),\ R_{\text{perc}}(t),\ \theta(t),\ \text{task flags, visibility conditions}\right],$$

where distances, visibility angles, and instantaneous risk values shape the agent's perception of hazards and guide the learning process.

The transition function $P(\cdot)$ is not analytically known; instead, it is defined implicitly by the Unity physics engine, the risk model, and the scenario manager, making COYOTE a model-free RL environment in practice.

Safety-critical environments require reward structures that encourage long-term safety, procedural correctness, and efficient behaviour. COYOTE uses a composite reward:

$$R(s_t, a_t) = r_{\text{task}}(t) - \lambda_{\text{real}} R_{\text{real}}(t) - \lambda_{\text{perc}} R_{\text{perc}}(t) - \lambda_{\text{col}} \mathbb{1}_{\text{collision}} - \lambda_{\text{rule}} \mathbb{1}_{\text{violation}},$$

where:

- $r_{\text{task}}(t)$ rewards moving toward goals or completing tasks.
- $R_{\text{real}}(t)$ penalizes proximity to dangerous vehicles.
- $R_{\text{perc}}(t)$ penalizes states the agent interprets as dangerous.
- Collisions and rule violations produce large negative spikes.
- The λ coefficients tune the relative importance of safety and performance.

This reward formulation ensures that RL agents not only optimize efficiency (e.g., travel time or mission success) but also internalize risk minimization as a primary behavioural driver.

At each simulation step, the agent observes s_t, selects an action a_t, and transitions to a new state s_{t+1}. The RL objective is to maximize the expected discounted cumulative reward:

$$J(\pi) = \mathbb{E}_\pi \left[\sum_{t=0}^{\infty} \gamma^t R(s_t, a_t) \right],$$

where $\pi(a \mid s)$is the policy mapping states to action probabilities.
The optimal policy satisfies:

$$\pi^* = \arg \max_{\pi} \; J(\pi).$$

To estimate π^*, COYOTE uses Q-learning or Deep Q-Learning depending on agent complexity.

For tabular state spaces, Q-learning updates the action-value function:

$$Q^\pi(s, a) = \mathbb{E}[R(s, a) + \gamma \max_{a'} \; Q(s', a')].$$

The iterative update rule used in COYOTE is:

$$Q_{t+1}(s_t, a_t) \\ = Q_t(s_t, a_t) + \alpha[R(s_t, a_t) + \gamma \max_{a'} \; Q_t(s_{t+1}, a') - Q_t(s_t, a_t)],$$

where $\alpha$is the learning rate.
The learned policy is:

$$\pi(s) = \arg \max_{a} \; Q(s, a).$$

For continuous risk values and dynamic environments (moving vehicles, unpredictable occlusions), Q-learning naturally converges to avoidance behaviours, safe trajectories, and collision-free navigation.

For agents requiring continuous action spaces (e.g., smooth steering, smooth acceleration), COYOTE supports policy-gradient architectures. Here the policy is parameterized:

$$\pi_\theta(a \mid s),$$

and the goal is to maximize:

$$J(\theta) = \mathbb{E}_{\pi_\theta}\left[\sum_{t=0}^{\infty} \gamma^t R(s_t, a_t)\right].$$

The gradient is computed via:

$$\nabla_\theta J(\theta) = \mathbb{E}[\nabla_\theta \log \pi_\theta(a_t \mid s_t) A_t],$$

where the advantage term $A_t$measures how much better an action was compared to a baseline.
An Actor–Critic model is then defined through:

$$\text{Actor: } \pi_\theta(a \mid s),\text{Critic: } V_\phi(s),$$

with critic updates:

$$V_\phi(s_t) \leftarrow V_\phi(s_t) + \alpha_v[R_t + \gamma V_\phi(s_{t+1}) - V_\phi(s_t)].$$

This architecture is well suited for agents performing continuous maneuvering, such as automated AGVs or safety drones in the yard.
Because COYOTE involves safety-critical dynamics, risk-aware modifications are essential. A risk-sensitive performance index is defined as:

$$J_{\text{risk}}(\pi) = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t \left(R(s_t, a_t) - \beta R_{\text{real}}(t)\right)\right],$$

with βcontrolling sensitivity to objective hazard fields. Alternatively, the agent may minimize the conditional value at risk (CVaR):

$$\text{CVaR}_\alpha(Z) = \mathbb{E}[Z \mid Z \geq F_Z^{-1}(\alpha)],$$

where $Z$is accumulated risk and F_Z^{-1}its quantile.
These risk-aware formulations drive the RL agent to avoid rare but catastrophic events such as collision trajectories or entering blind spots.

Convergence of Q-learning is guaranteed when:

- $\sum_t \alpha_t = \infty$,
- $\sum_t \alpha_t^2 < \infty$,
- the MDP is finite,
- all state–action pairs are explored infinitely often.

COYOTE enforces these conditions through:

- decaying learning rates,
- exploration strategies (ϵ-greedy or Gaussian noise),
- repeated scenario randomization.

Safety is handled through:

- large negative terminal rewards for collisions,
- risk-based shaping terms,
- safety layers preventing impossible or dangerous actions.

Putting the components together, the RL loop in COYOTE is:

1. Read state $s_t$from Unity (physics + cognitive fields).
2. Choose action $a_t \sim \pi_\theta(a \mid s_t)$.
3. Apply action in the environment.
4. Unity simulates continuous physics for Δt.
5. Compute new risk values $R_{\text{real}}(t), R_{\text{perc}}(t)$.
6. Compute reward $R(s_t, a_t)$.
7. Update Q or policy parameters.
8. Repeat.

Through this formalism, COYOTE's RL agents learn to:

- anticipate vehicle trajectories,
- avoid hazardous zones,
- minimize risk exposure,
- plan efficient paths to mission objectives,
- compensate for fog and occlusions,
- behave predictively around human operators.

4.6. Interoperability

Interoperability represents a foundational capability in modern Modeling & Simulation, enabling heterogeneous simulators, often built using different formalisms, programming languages, execution models, and time-management strategies, to interoperate within a unified synthetic environment. In complex operational domains such as port logistics, defense, industrial safety, and transportation, relying on a single monolithic simulator is rarely sufficient, as different subsystems capture different aspects of reality with varying levels of detail. Interoperability therefore enables the integration of specialized models (such as human-operator simulators, vehicle motion simulators, cyber layers, environmental dispersion models, command-and-control systems, and training platforms) into a coherent distributed architecture capable of reproducing the multifaceted nature of real operations.
The COYOTE simulator has been designed within this philosophy, adopting the principles of distributed simulation to ensure that its components, as well as potential external simulators, can exchange information, synchronize their states, and jointly contribute to a shared virtual scenario. This approach not only supports scalability and modularity but also aligns COYOTE with the broader MS2G (Modeling, interoperable Simulation and Serious Games) paradigm, which emphasizes the seamless integration of high-fidelity models, XR interfaces, and intelligent agents under a unified interoperable umbrella.

To achieve interoperability, the simulation community has developed several technical standards that define how simulators should communicate and how their execution should be coordinated. Among these, the High Level Architecture (HLA) stands as the most widely adopted. Originally developed within the United States Department of Defense and subsequently formalized as an IEEE standard (IEEE 1516), HLA is now maintained by SISO (Simulation Interoperability Standards Organization) and is used extensively by NASA, NATO, the European Defence Agency, and major governmental and industrial organizations involved in distributed simulation. NATO had standardize the interoperability by the release of STANAG 4603. HLA offers a common language and an integration infrastructure that enables independent simulators (called federates) to participate in a cooperative simulation execution, called a federation.

HLA adopts an object-oriented approach to modeling the information exchanged among federates. This information is described in a Federation Object Model (FOM), which acts as the shared data schema of the federation. The FOM defines the classes of objects (such as vehicles, agents, environmental entities) and interactions (such as events, commands, messages) that may be exchanged, as well as their attributes and update rules. Because of this explicit and formalized description, federates may be implemented in any programming language, run on any hardware platform, and use any internal simulation engine, as long as they implement the interface required to publish and subscribe to FOM-defined data.

Information exchange in HLA follows a publish-subscribe mechanism. Each federate declares the types of data it intends to publish (for example, the positions of autonomous vehicles controlled by COYOTE's intelligent-agent subsystem) and the types of data it wishes to receive (for example, updates from a cyber-layer simulator or weather-dispersion model). This mechanism ensures that data flows efficiently across the federation without global broadcast overhead or unnecessary network traffic. Additionally, HLA provides services for time management, synchronization, ownership transfer, and federation control, allowing simulators with different time-advancement strategies (discrete-event, real-time, time-stepped)to coexist within the same distributed execution.

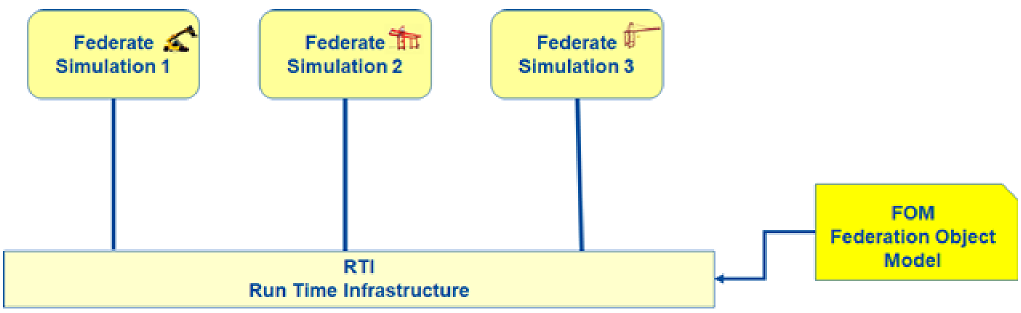

Figure 39. Exmpale of HLA Architecture

The core of the HLA architecture is the Runtime Infrastructure (RTI), a middleware responsible for managing communication, synchronization, and federation services. The RTI implements the HLA interface specification and acts as the intermediary through which all federates exchange data. Through the RTI, federates can join or resign from a federation, synchronize initialization states, request and grant time advances, and publish or subscribe to object classes and interactions defined in the FOM. The interoperability achieved through the RTI allows each simulator to remain autonomous in terms of development and execution logic while participating in a coordinated and coherent distributed simulation.

Within this structure, each federate represents an independent system connecting to the RTI. A federate may encapsulate a complete simulator or a subsystem such as a visualization engine, a decision-support module, a robotics simulator, or an XR interface. In the context of COYOTE, the simulator itself can act as a federate within a broader federation, for example, by connecting to a higher-level port logistics simulator, a cyber-situational-awareness tool, or a C2 training platform. The federates collectively form the federation: the interoperable ensemble of simulators, the RTI, and the FOM that governs their interactions. The execution of a federation, known as the Federation Execution, corresponds to the actual distributed simulation session in which all federates interact according to the rules defined in the FOM.

The adoption of interoperability standards such as HLA brings several advantages. First, it ensures modularity and reusability, allowing existing simulators to be integrated without rewriting or restructuring their internal logic. Second, it ensures scalability, enabling developers to add or remove components as needed. Third, it supports experimentation and training scenarios that require the coordinated operation of heterogeneous simulators, for instance, linking COYOTE with real-time port-logistics systems, environmental dispersion models such as CACTUS, or higher-level strategic wargaming frameworks. Fourth, it supports modernization

and extensibility, as simulators can evolve independently, provided that they continue to follow the FOM's specifications.
The applicability of HLA to port environments and industrial safety simulations has already been demonstrated in several NATO and EU research projects, where distributed simulators have been used to model multi-domain scenarios involving physical, cyber, cognitive, and logistical components. The COYOTE simulator is therefore positioned not only as a standalone training platform but also as a fully interoperable component capable of being integrated into large-scale digital-twin ecosystems or multi-layer decision-support architectures. Its XR-driven operator-in-the-loop capability can be federated with higher-fidelity machinery simulators, IoT-based monitoring systems, or container-tracking digital twins, enabling both real-time interaction and distributed training sessions across multiple devices and locations.

In COYOTE, the simulation engine developed in Unity3D acts as a federate within an HLA federation, communicating through a dedicated interface layer written in C# that implements the essential services of the Run-Time Infrastructure (RTI). This interface is responsible for managing publish-subscribe operations, synchronizing simulation time, and handling attribute updates and interactions as defined in the Federation Object Model (FOM). Unlike typical Unity-based simulators, where interoperability is often minimal or limited to proprietary network protocols, COYOTE integrates a full FOM-driven representation of the port terminal. The FOM specifies not only static elements such as container blocks, vehicle types, inspection targets, and navigation graphs, but also dynamic entities including real-time vehicle kinematics, hazard states, operator position and posture, and cognitive-perceptual indicators used in the risk model. The inclusion of perceptual and cognitive variables in a FOM is one of the distinctive innovations of COYOTE: while classical federations exchange physical quantities such as coordinates, velocities, and environmental parameters, COYOTE extends HLA semantics to include operator awareness, visual obstruction factors, inspection accuracy, and workload estimation. This enables the federation to integrate human-in-the-loop behaviors into broader operational simulations, a capability aligned with emerging NATO M&S directions on human behavior modeling.

From an implementation standpoint, the COYOTE federate uses a decoupled architecture built around three internal subsystems:

(1) a sensorial aggregator that samples real-time internal variables (user position, vehicle positions, occlusion metrics, environmental conditions);

(2) an HLA interface module implementing callbacks, attribute updates, and subscription handlers

(3) a synchronization and interpolation layer that reconciles external federate states with Unity's frame-based rendering pipeline.

This separation is essential because Unity runs on a frame-rate loop, while most external federates in port-logistics simulations (e.g., crane simulators, cyber simulators, traffic models) operate in discrete-event or time-stepped engines. The synchronization layer developed for COYOTE, based on buffered interpolation windows and adaptive time alignment, represents another key innovation, enabling seamless integration of Unity's rendering cycle with deterministic simulation clocks.

One of the most advanced interoperability features implemented in COYOTE concerns the interaction between intelligent-agent dynamics and external simulators. The agent-based vehicle controllers running inside COYOTE are capable of receiving path updates, operational constraints, or hazard information from external federates. For example, a crane simulator or a yard traffic model running in MATLAB/Simulink, AnyLogic, or a C++ discrete-event engine may publish container-movement events or traffic-control decisions into the federation. The COYOTE federate subscribes to these updates and integrates them into the local agent behavior tree, modifying routes, speeds, and obstacle-avoidance strategies in real time. This bi-directional coupling elevates COYOTE from a passive visualization tool to an active participant in a distributed ecosystem, allowing it to reflect the global state of the terminal while also contributing human operator data back to the federation.

Furthermore, COYOTE introduces a hybrid data-management concept that merges traditional HLA object updates with high-frequency XR metrics. Classical RTI implementations are not optimized for the rapid transmission of positional updates at 60-120 Hz required by VR/AR devices. To address this, COYOTE uses a dual-channel architecture: low-frequency updates (state changes, container anomalies, risk levels, mission status) are transmitted through standard HLA publish–subscribe mechanisms, while high-frequency XR data streams (head orientation, hand movements, gaze direction) are shared through a lightweight local protocol synchronized with the HLA clock through soft time-stamping. This hybrid approach preserves FOM compliance while ensuring the responsiveness needed for immersive training.

The interoperability layer of COYOTE also supports federation extensibility. New simulators, such as gas-dispersion models like CACTUS, cyber-range simulators, straddle-carrier motion simulators, reefer monitoring digital twins, or port-energy management systems, can be integrated by extending the FOM with new object classes and interactions.

4.7. KPIs and Measures of Merit (MoM)

The evaluation of operator performance in COYOTE is grounded on a structured set of Measures of Merit (MoM), which quantify the quality, efficiency, and safety of the actions performed during each mission. In the context of complex industrial and port operations, MoM serve as formalized performance objectives that support both training and experimentation. They allow the system to assess how effectively a user completes assigned tasks while simultaneously capturing behavioral patterns related to risk avoidance, situational awareness, and procedural compliance. Within COYOTE, these measures are particularly relevant because missions are dynamically generated according to configurable difficulty and complexity levels, and operators may encounter varying traffic densities, environmental conditions, and safety constraints.
Each mission consists of multiple tasks, typically involving the identification, inspection, and classification of containers distributed across the yard. These tasks may include verifying the presence of leaks, assessing container integrity, checking seal conditions, or confirming the correctness of positioning. Every task contains one or more accomplishments, which represent measurable subtasks necessary for successful completion. Depending on the mission configuration, different accomplishments may carry different weights, reflecting their operational relevance and their impact on safety. This structure allows missions to be tailored to operator skill level, experimental requirements, or specific training objectives.

	Precision	It represents a measure of the quality of the work performed by the operator and how much of the assigned work was completed correctly.
	Readiness	It represents a measure of the speed with which the operator completed the assigned task.
	Accuracy	It represents the ability to follow rules and procedures and avoid making mistakes in carrying out assigned tasks.

	Caution	It represents the operator's prudence and ability to avoid exposure to risks, or their ability to reduce their exposure to risk during the execution of the mission.
	Awareness	It represents a measure of the operator's awareness of the risks surrounding him and his ability to limit them by taking appropriate actions in carrying out the assigned mission.

Figure 40. Measures of Merit (MoM)

To translate MoM into measurable quantities, COYOTE computes a series of Key Performance Indicators (KPIs) during each simulation run. These indicators operate at different temporal scales: some reflect global mission outcomes, whereas others provide instantaneous or time-averaged information about the operator’s exposure to hazards.

Duration measures the time used to complete the mission. The standard time limit is set to 20 minutes, but the simulator records the exact time spent, enabling comparisons between operators and across difficulty levels. High values indicate inefficiency or hesitation, while exceptionally low values may suggest rushed behavior potentially associated with higher risk exposure.

Pts (Points) aggregates correctness in inspections and task execution, reflecting how accurately the operator evaluated the status of each container. This KPI contributes directly to the MoM of *Precision* and *Correttezza*.

Incidents counts the number and severity of collisions or dangerous interactions during the mission. A single collision immediately jeopardizes the mission outcome, as it indicates a critical failure in situational awareness and risk mitigation. This KPI directly affects *Accortezza* and *Consapevolezza*.

The simulator also computes detailed indicators describing the user’s exposure to real and perceived risk, derived from the multi-domain risk model described earlier in this chapter.

TotRE (Total Risk Exposure) quantifies the cumulative real risk encountered during the mission. It measures how much hazard the operator experienced over the entire duration of the scenario and is fundamental for evaluating prudence in movement.

AvgRE (Average Risk Exposure) represents the average level of real risk to which the operator was exposed. As a mission-quality indicator, AvgRE reflects the user's ability to maintain safe behavior consistently. In the reference scenario, a threshold of 30 points is considered the upper limit for a "virtuous" mission.

MaxRE (Maximum Risk Exposure) identifies the peak real risk reached at any moment during the mission. This metric highlights critical near-misses or unsafe decisions, with a threshold of 120 points indicating the boundary between acceptable and dangerous behavior.
In addition to real risk, COYOTE computes perceptual risk-based KPIs to evaluate the cognitive alignment between real danger and the operator's internal perception of it.

pTotRE (Perceived Total Risk Exposure) measures the cumulative subjective risk as interpreted by the operator's sensory and cognitive models. High values may indicate anxiety or overestimation of threats, while low values in high-risk missions may indicate dangerous underestimation.

pAvgRE (Perceived Average Risk Exposure) provides a mean value of perceived risk over time. Values beyond 120 points suggest that the operator consistently misjudged risk, either by overestimating environmental threat or by failing to interpret hazard cues correctly.

pMaxRE (Perceived Maximum Risk Exposure) corresponds to the highest instantaneous perceived risk value during the mission. This KPI identifies moments where the operator felt particularly threatened or, alternatively, where sensory limitations or cognitive biases created mismatches between real and perceived danger.

					
	Precision	Readiness	Accuracy	Caution	Awareness
Points	☑				
Duration		☑			
Incidents			☑		
TotRE, AvgRE, MaxRE				☑	
pTotRE, pAvgRE, pMaxRE					☑

Figure 41. Key Performance indicators (KPIs)

5. Experimental Campaigns

The experimental campaign was conducted at the Simulation Laboratory of the University of Genoa with the objective of evaluating the effectiveness of COYOTE in replicating realistic terminal-yard inspection tasks and measuring operator performance, risk exposure, and perceptual–cognitive alignment in safety-critical scenarios. Participants were recruited among professional port operators, trainees, and individuals with prior exposure to port-logistics activities. A total of 18 volunteers (5 female, 12 male) between 27 and 61 years of age (mean ± SD: 39.9 ± 9.8 years; median: 38) participated in the study. Recruitment was carried out through institutional outreach within the Port of Genoa community, and all participants provided informed consent prior to engaging in the simulation.

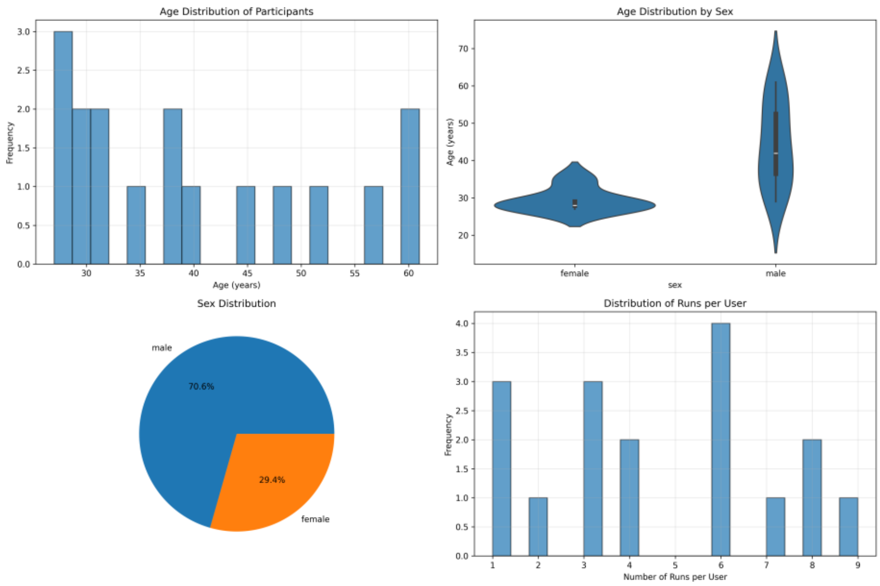

Figure 42. Demographic summaries: age histogram, violin plot by sex, sex pie chart, and runs-per-user

Each participant interacted with the COYOTE simulator using a workstation-based configuration, controlling the avatar through standard keyboard inputs. The simulated environment reproduced a high-fidelity container terminal with realistic ground traffic, vehicle behavior, and inspection points, requiring participants to locate and evaluate faulty containers while navigating through congested and partially occluded spaces. For each inspection, performance was quantified through accuracy scores (based on correct identification of leaks and seal integrity), navigation efficiency (completion time and path coherence), and safety metrics (collisions, total objective risk exposure *Rtot*, and total perceived risk exposure *pRtot*). The simulator provided both visual and auditory cues, and allowed the activation of Augmented Reality (AR) rays that pointed dynamically toward the assigned target container. Visibility conditions could be further manipulated through fog, simulating degraded sensory conditions.

The experiment was structured around three independent factors:

1. **Difficulty level** (2 levels), reflecting the density of moving vehicles and the level of operational congestion in the yard.
2. **Augmented Reality (AR) guidance** (on/off), corresponding to the activation of red directional rays to support target localization.
3. **Fog** (on/off), introducing reduced visibility conditions.

These factors were combined to generate a set of controlled scenario configurations. Mission difficulty also encoded environmental stressors such as increased vehicle flows, tighter spatial constraints, and higher likelihood of container misplacement. The AR mode and fog settings were manipulated independently, enabling the analysis of how sensory augmentation and sensory degradation influence perception, decision-making, and risk management.

Participants completed a fixed sequence of 10 simulation runs according to the following structure:

- **Runs 1–6:** Standard difficulty (level 2), no AR cues.
- **Runs 7–8:** Standard difficulty with AR cues activated.
- **Runs 9–10:** Standard difficulty, no AR cues.

This design produced true replicates for baseline conditions (runs 1-6 and 9-10) and cross-condition comparisons for AR (runs 7-8). The final three runs for each participant were specifically intended to assess learning effects, comprehension of the scenario, and improvements in risk-

perception calibration over time. Runs 7-8 allowed the measurement of performance gains due to augmented guidance.
Overall, the experimental campaign generated 87 complete simulation runs, providing sufficient data for both within-subject and between-subject statistical analyses.

Each participant began with a brief familiarization phase, followed by the execution of the full experimental protocol. For every simulation run, the user was instructed to navigate the yard, locate the three designated containers, and assess their condition (intact, small leak, medium leak), recording the outcome within the interface menu. During navigation, participants were exposed to moving trucks, cranes carrying suspended loads, occlusion-induced blind spots, and variable visibility conditions. Safety constraints required maintaining appropriate distances from vehicles and avoiding suspended loads; collisions or near-misses were automatically registered by the simulator. A collision immediately terminated the mission and flagged it as unsuccessful while preserving all recorded data.
Each simulation run triggered automatic logging of detailed telemetry: user trajectory and speed, proximity to hazards, timestamped interactions, inspection decisions, real and perceived risk exposure streams, collisions, occlusion states, and gaze-related orientation variables. The simulator saved both raw logs and a synthesized summary for each run. The entire set of runs for each user was completed in a single session lasting approximately 45-60 minutes.
Raw logs were exported and processed using Python 3.11. Preprocessing scripts extracted derived features including instantaneous and cumulative risk exposure, local speed fields, collision frequencies, path deviation indices, and perception-bias metrics defined as:

$$\text{Perception Bias} = \frac{pR_{\text{tot}} - R_{\text{tot}}}{R_{\text{tot}}}.$$

Incomplete runs, corrupted logs, and extreme outliers (defined via IQR and visual diagnostics) were removed from the dataset. Numeric variables were standardized before regression and clustering tasks, while categorical

factors (difficulty, AR, fog) were encoded using one-hot or effect-coding schemes depending on the target model. Visual exploratory diagnostics were generated for all variables, including histograms, kernel density estimates, and time-series overlays.
Distributions for completion time, *Rtot*, *pRtot*, and several risk-related indicators exhibited strong positive skewness. Shapiro–Wilk tests confirmed non-normality ($p < 0.01$), motivating the use of $\log_{10}(x+1)$ transformations for parametric modeling. For interpretability, however, raw values were retained in figures and descriptive summary tables.

The methodological framework followed a rigorous, multi-step analytical pipeline integrating classical design-of-experiment techniques with modern mixed-effects and risk-perception analytics:
(1) Exploratory Data Analysis.
Univariate statistics were computed for all variables, stratified by factor level. Normality was tested via Shapiro–Wilk, and homogeneity of variances via Levene tests ($\alpha=0.05$). Diagnostics indicated heteroscedasticity and skewness for several variables.
(2) Factorial Inference.
A repeated-measures ANOVA with Greenhouse–Geisser correction was conducted to test the influence of difficulty, AR, and fog on completion time, accuracy, and risk metrics. When heterogeneity of variance was present, Welch F-tests or nonparametric Kruskal–Wallis tests were used. Effect sizes were computed as partial η^2.
(3) Mixed-Effects Modeling.
Linear mixed-effects models (LMM) were employed using user ID as a random intercept. For selected outcomes, random slopes were added to account for differential sensitivity to AR or fog. Model adequacy was evaluated through *Mean Square Pure Error* (MsPE) using true replicates. Bartlett tests ($p = 0.64$–0.72) confirmed no significant variance differences among replicates. Intraclass correlation coefficients (ICC) quantified the proportion of variance attributable to between-user differences.

(4) Risk-Perception Calibration.
Paired t-tests and Wilcoxon signed-rank tests compared *pRtot* and *Rtot*, revealing misalignment between real and perceived danger. These tests measured whether participants systematically underestimated or overestimated risk, a key metric for cognitive-risk modeling.
(5) Learning Effects.
Learning within conditions was assessed by correlating centered run index with each outcome for users with ≥2 replicates. Transfer learning across conditions was tested by correlating overall run order with performance metrics. One-sample t-tests evaluated whether correlations significantly deviated from zero.
(6) Assumption Checking and Diagnostics.
Residual-versus-fitted plots, Q–Q plots, scale–location plots, and histogram inspections confirmed acceptable model assumptions after transformation, with no severe deviations.
All analyses adhered to $\alpha = 0.05$ (two-tailed) unless otherwise specified.

6. Results and Discussion

The experimental campaign produced a comprehensive dataset that enabled the evaluation of performance, safety, and cognitive calibration across multiple simulation runs. The analyses focused on three categories of outcomes: (i) mission performance indicators such as completion time normalized by distance, (ii) objective safety indicators such as total risk exposure *(TotRE)*, and (iii) perceptual and cognitive indicators such as total perceived risk exposure *(pTotRE)* and risk-bias metrics. Together, these results illustrate how operators adapted to the simulated environment, whether they improved their situational awareness, and how their perception of danger aligned, or failed to align, with objective risk values.

Duration Normalized by Distance

Because the simulator randomizes container placement at each run, the total distance walked by the operator varies substantially. For this reason, performance was evaluated using *duration normalized by distance*, which measures how efficiently the operator completed the inspection relative to the ground they needed to cover.

All participants showed improvements in this metric across sessions. Significant improvements were detected in 56% of cases, with an average improvement of 57% and a maximum improvement of 75% . This result indicates that users became more efficient at navigating the yard, planning their trajectory, and avoiding unnecessary detours as their familiarity with the environment grew.

The trendline in the *DurationDista by Runs* plot shows a clear downward pattern, consistent with faster and more economical movements. Even though the experimental design did not explicitly include a navigational training component, participants nonetheless improved their spatial strategies, suggesting that the simulator effectively promotes operator adaptation in dynamic yard environments.

Target:	DurDIsta		
Total Tests	18		
Significant	Results	10,0	56%
Improving	10,0	100%	
Imprvements	Average	57%	
	Max	75%	
Overall	Average	54%	
	Significant	Improvement	Negative
People			
P1			
P2			
P3			
P4			
P5	1,0	42%	
P6	1,0	53%	
P7	1,0	51%	
P8	1,0	26%	
P9	1,0	74%	
P10			
P11			
P12	1,0	75%	
P13			
P14	1,0	66%	
P15	1,0	62%	
P16	1,0	62%	
P17			
P18	1,0	28%	

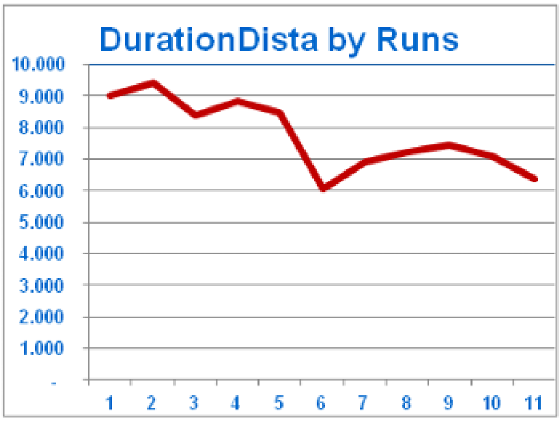

Figure 43. Duration

Total Exposure Risk (TotRE)

Objective safety performance was evaluated using *Total Risk Exposure* (TotRE), which integrates all instantaneous risk values accumulated during a mission. This metric reflects how much physical danger the operator experienced, considering proximity to moving vehicles, occlusions, suspended loads, and environmental conditions.

All users demonstrated meaningful improvements across the sequence of runs, with 61% showing statistically significant improvements ($p < 0.05$). The *average improvement* in TotRE was 56%, and the most substantial improvement reached 83% .

The *TotRE by Runs* curve shows a marked decrease from the first to the sixth run, with a particularly notable drop around runs 5–6, coinciding with increased scenario familiarity. Although exposure increased slightly in subsequent runs where AR cues were removed, the overall trend remained downward.

These findings demonstrate that operators learned to maintain safer distances, avoid blind spots, and plan movement with greater caution, even without receiving explicit safety instructions.

Target:	TotRE		
Total Tests	18		
Significant Results		11,0	61%
Improving	11,0	100%	
Imprvements	Average	56%	
	Max	83%	
Overall	Average	56%	
	Significant	Improvement	Negative
People			
P1			
P2			
P3	1,0	31%	
P4			
P5	1,0	56%	
P6	1,0	65%	
P7	1,0	44%	
P8	1,0	36%	
P9	1,0	83%	
P10			
P11			
P12	1,0	83%	
P13	1,0	53%	
P14	1,0	35%	
P15	1,0	74%	
P16	1,0	59%	
P17			
P18			

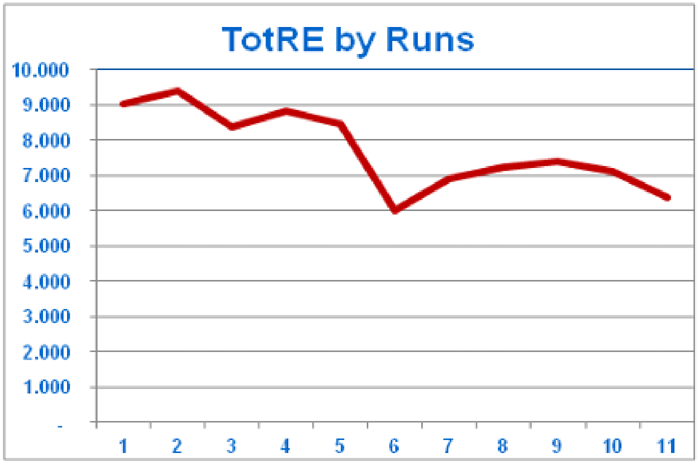

Figure 44. Total Risk Exposure

Total Perceived Risk Exposure (pTotRE)

Perceived risk exposure provides insight into the operator's internal interpretation of hazardous conditions. The majority of participants (85%) improved their perceived exposure metric, with 72% of improvements reaching statistical significance. The mean improvement was 63%, with a maximum improvement of 93% .

This indicates a meaningful shift in how operators interpret visual and auditory cues in the environment, suggesting that repeated exposure to the scenario improved their ability to detect threats, even when the environment became more congested or challenging.

The similarity between the TotRE and pTotRE trends implies that improvements in spatial behavior were accompanied by improvements in perceptual reasoning.

Target:	pTotRE			
Total Tests	18			
Significant Results		13,0		72%
Improving	11,0	85%		
Imprvements	Average	63%		
	Max	93%		
Overall	Average	60%		
People	Significant	Improvement	Negative	
P1			47%	
P2	1,0	30%		
P3	1,0	37%		
P4				
P5	1,0	51%		
P6	1,0	80%		
P7				
P8				
P9	1,0	92%		
P10				
P11	1,0	43%		
P12	1,0	93%		
P13	1,0	74%		
P14				
P15	1,0	80%		
P16	1,0	70%		
P17	1,0	42%		
P18			42%	

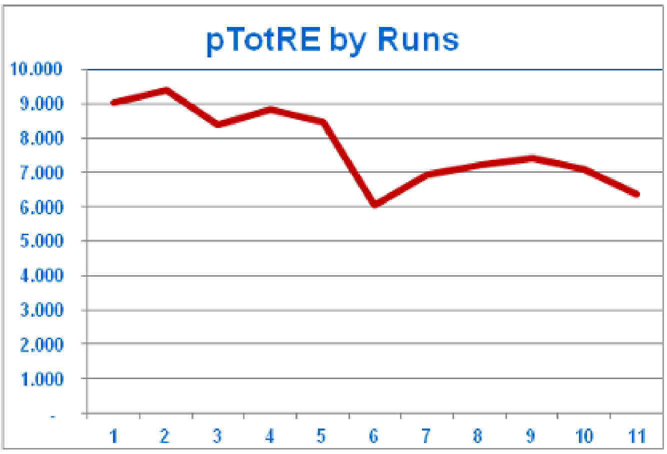

Figure 45. Total Perceived Risk Exposure

Risk Bias Analysis

To quantify risk-perception accuracy, we computed a *risk-bias index* defined as:

$$\text{Risk Bias} = \frac{pRE_{\text{tot}} - R_{\text{tot}}}{R_{\text{tot}}}$$

A risk bias of zero indicates perfect calibration; negative values reflect underestimation of danger.

The distribution of risk bias (Fig. 4) was approximately normal (Shapiro–Wilk p = 0.623), with a mean of –0.295 (SD = 0.125). This implies that participants systematically underestimated their real risk by ~29%.

A one-sample t-test against zero confirmed a highly significant underestimation:

- $t(77) = -20.8$,
- $p < 10^{-28}$,
- 95% CI = [–0.323, –0.267],
- Cohen's d = –2.36.

A Wilcoxon signed-rank test corroborated the finding ($p < 10^{-13}$). These results represent extremely strong evidence of cognitive under-calibration in the baseline population, a finding consistent with safety studies in real port terminals.

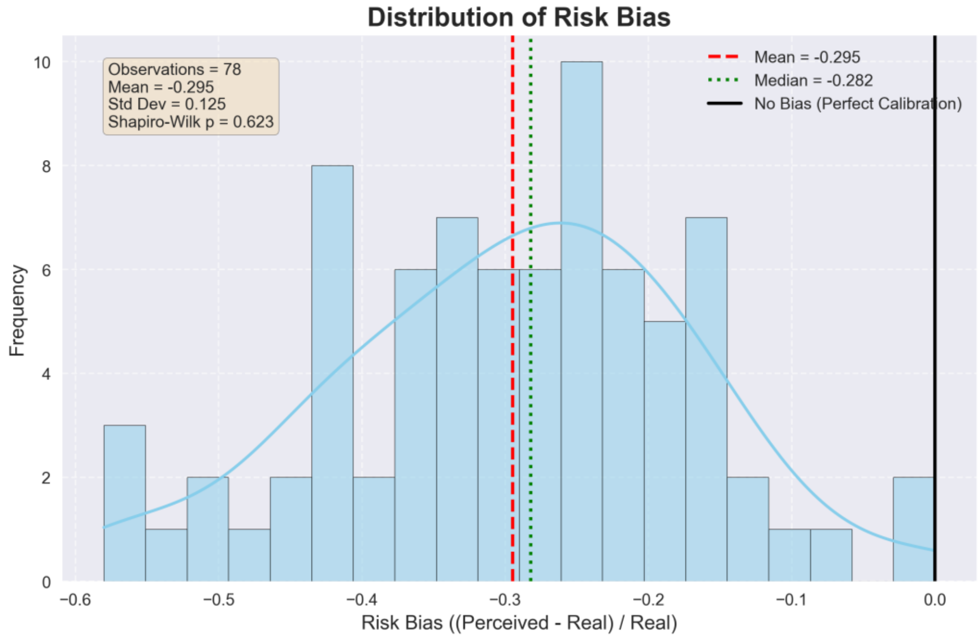

Figure 46. Distribution of risk bias (histogram with KDE), Shapiro-Wilk statistics, and calibration

A simple linear calibration regression yielded:

$$pR_{\text{tot}} = 0.15 + 0.92\, R_{\text{tot}}, R^2 = 0.85$$

The slope close to 1 indicates that perceived risk tracks real risk linearly, but the positive intercept shows a downward shift, operators perceive less risk than they objectively experience.

Learning and Adaptation Effects

Within-Scenario Learning (True Replicates)

Analyses on repeated identical scenarios showed:

- No significant learning for any metric ($|t|<1.4$, $p>0.18$).

This suggests that repeating the same conditions does not improve performance; operators require variation to learn effectively.

Transfer Learning Across Scenarios

In contrast, learning across *varied* scenarios produced measurable improvements:

- Completion time: mean r = –0.32, p = 0.008
- Objective risk (Rtot): r = –0.29, p = 0.02
- Perceived risk (pRtot): r = –0.27, p = 0.03

These trends demonstrate adaptive generalization: operators became faster, safer, and more perceptually calibrated across heterogeneous conditions.

Age Effects

The next Figure shows risk bias by age:

- Median risk bias ≈ –0.33 for both groups
- Mann–Whitney U = 28, p = 0.91
- Interaction between run index and age not significant (β = 0.004 ± 0.017, p=0.81)

Age did not influence calibration or learning trajectory.

Individual Slopes and Variability

The distribution of individual learning slopes centered around –0.16, indicating moderate trends toward worsening risk perception for some users. However, the large SD (≈0.39) and the random-slope variance (σ^2=0.046) indicate substantial inter-individual variability, a common finding in cognitive-safety assessments. Increasing sample size and number of replicates will strengthen future inference.

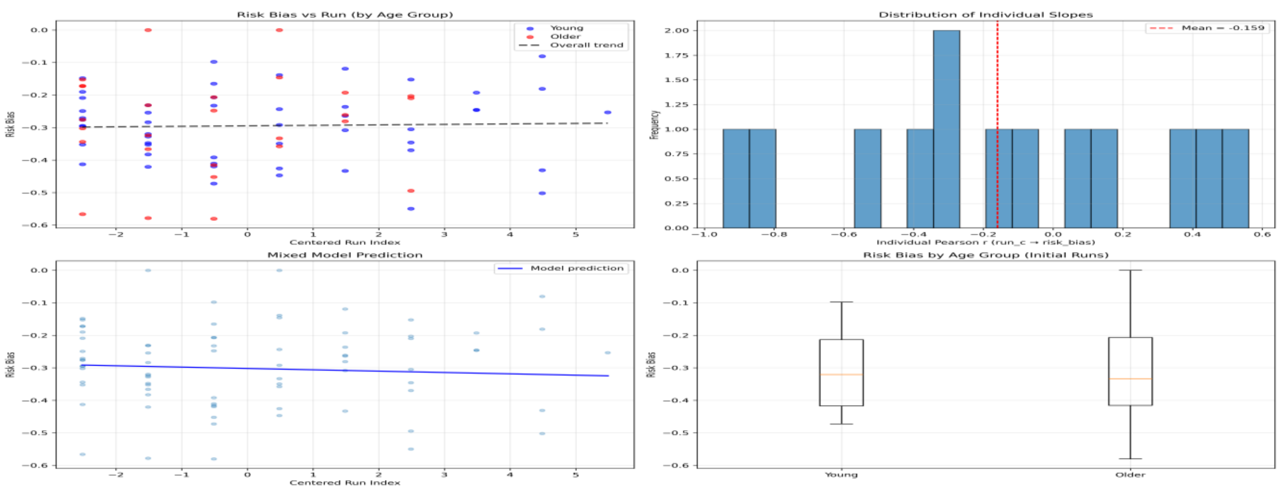

Figure 47. Learning analyses: (A) risk bias vs. centred run index coloured by age group; (B) histogram

To evaluate how environmental conditions influence operator performance, a sensitivity analysis was conducted on the full dataset of port operators using a Central Composite Design (CCD). Two experimental factors were considered, each at two levels:

- **F1: Visibility** optimal visibility [+1] vs. poor visibility $[-1]$
- **F2: Time of Day** daytime [+1] vs. nighttime [-1]

The model included the main effects $F1$and $F2$, as well as the interaction term $F1 \cdot F2$. The resulting sensitivity coefficients for the three Measures of Merit (MoMs) are summarised below.

Table 4. Rsisk Exposure

Effect	Sensitivity
(F2)	(-0.45)
(F1)	(-1.68)
(F1 F2)	(-4.72)

These results indicate that both improved visibility ($F1$) and daylight conditions ($F2$) reduce risk exposure. The interaction term is strongly negative, confirming that the combined presence of good visibility and daylight produces the largest reduction in risk, consistent with operational evidence from port-yard activities.

Table 5. Speed (Operational Efficiency)

Effect	Sensitivity
(F2)	(1.59)
(F1)	(0.04)
(F1 F2)	(1.43)

Both factors contribute positively to increased speed. Daylight has the strongest independent effect, while the interaction indicates that visibility and time of day jointly enhance efficiency, improving operators' ability to complete tasks quickly without sacrificing control or awareness.

Table 6. Correctness of Mission Completion

Effect	Sensitivity
(F2)	(0.10)
(F1)	(-0.10)
(F1 F2)	(0.17)

Correctness shows weaker effects. Visibility alone exhibits a slightly negative contribution, although the interaction term compensates for this, suggesting that poor visibility becomes less detrimental when combined with daylight, likely due to better contextual cues in the environment.

The analysis of mission completion speed between the first and last trial shows a clear improvement over time. This confirms the trend identified earlier: as the experiment progressed, the group became increasingly familiar with the simulated port environment, resulting in faster mission execution.

The next step focuses on the performance of group Port P1 under two conditions:

- No AR mode,
- AR mode, where visual hazard cues are displayed to the user.

In this case, the performance indicator considered is the Success Rate, that is, the percentage of missions completed correctly.

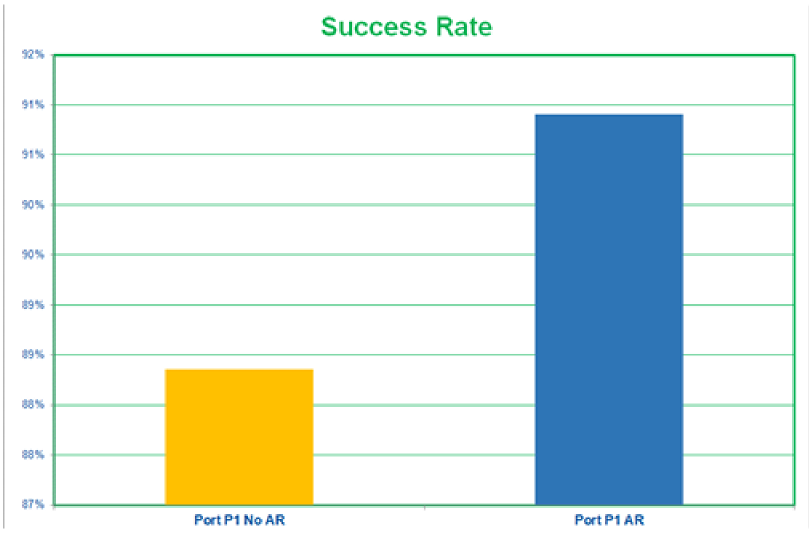

Figure 48. Success Rate of Port operators with AR and without AR

The results show that Port P1 initially achieved a relatively high success rate. However, the introduction of AR mode further increased the success percentage, demonstrating that augmented visual cues help operators detect hazards more effectively and complete tasks with higher accuracy.

A factor that instead appears to increase task difficulty is the presence of fog. Under low-visibility conditions, the results highlight two consistent trends:

- the average speed decreases,
- the average risk exposure increases.

This behaviour reflects the impact of sensory degradation: reduced visibility forces operators to move more cautiously, yet at the same time it compromises their ability to perceive approaching vehicles or nearby obstacles, leading to higher exposure to hazardous situations.

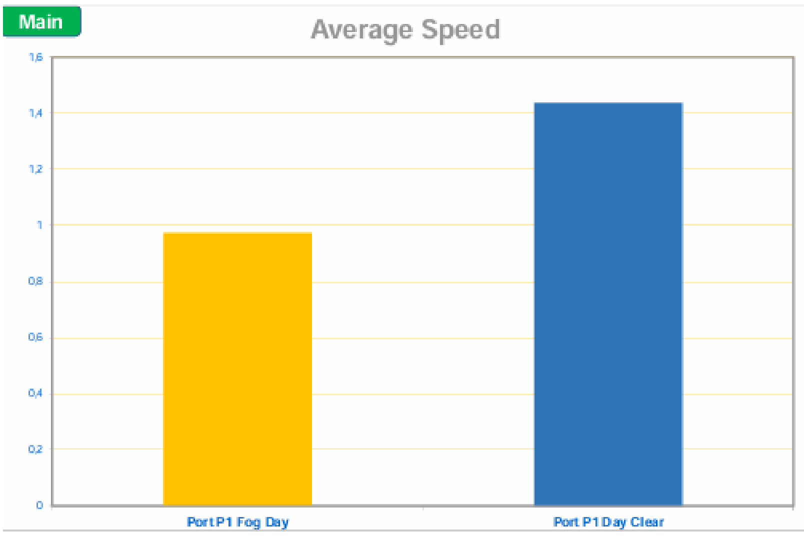

Figure 49. Avg speed with fog and witthout fog

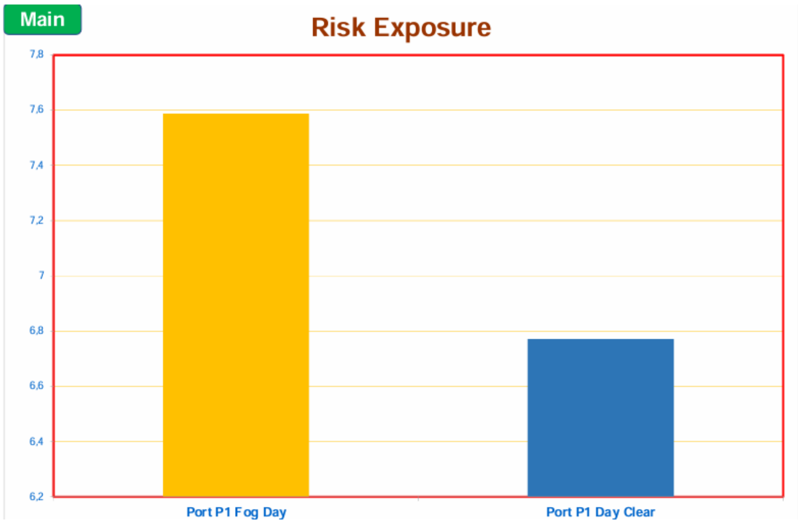

Figure 50. Risk Exposure with and without fog

The experimental results show that:

1. **Operators significantly improved performance** over successive runs, even in complex, dynamic environments.
2. **Objective risk exposure decreased substantially**, demonstrating genuine behavioral adaptation.
3. **Perceived-risk calibration improved but remained biased**, with a systematic underestimation of ~29%.
4. **Learning occurred only when scenario conditions varied**, not when users repeated identical tasks.
5. **Individual differences outweighed demographic predictors**, indicating that cognitive styles, experience, and attention patterns likely determine performance variance.

These findings provide an initial evidence for the training value of COYOTE. The simulator not only supports procedural learning and spatial optimization but also reveals cognitive vulnerabilities that would remain undetected in traditional training. Underestimation of risk is a critical safety issue in port environments; identifying and correcting such biases is essential for reducing accidents.

While Figures 43–45 demonstrate a clear quantitative improvement in terms of Success Rate and Risk Exposure, these results should be interpreted together with qualitative observations related to human factors and operator experience. During the initial phases of the experimentation, a subset of port operators exhibited a cautious attitude toward the use of XR headsets. This hesitation was primarily linked to unfamiliarity with immersive technologies, concerns about physical comfort during prolonged use, and an initial perception that the headset could reduce situational awareness in a safety-critical environment. Such reactions are consistent with well-known dynamics observed when introducing novel technologies in operational contexts characterized by high responsibility and low tolerance for error.

However, this initial resistance was found to be temporary. After a short familiarization period, operators progressively adapted to the XR interface and reported a growing sense of control and engagement within the simulated environment. In particular, the immersive representation of hazards and dynamic elements in the terminal yard contributed to a clearer understanding of spatial relationships, moving risks, and unsafe proximities. This qualitative feedback aligns with the observed reduction in Risk Exposure and suggests that the XR-based experience supports cognitive alignment between perceived risk and actual operational risk, rather than increasing cognitive load.

From a human factors perspective, the results indicate that XR does not act as a distracting layer, but rather as a mediating tool that enhances situational awareness when properly designed and introduced. Operators demonstrated an increased ability to anticipate dangerous situations, maintain safer distances from hazards, and adapt their behavior accordingly. The progressive improvement in Success Rate further suggests that experiential learning within the immersive environment accelerates procedural understanding and reinforces correct operational behavior more effectively than passive training methods.

The acceptance of XR technology was strongly influenced by how the system was presented and integrated into the operational framework. When framed as a training and decision-support tool aimed at improving safety and competence, acceptance was significantly higher than in scenarios where the technology could be perceived as a form of supervision or performance monitoring. This highlights the importance of change management strategies in the deployment of XR solutions within port operations. Early involvement of operators, transparent communication of training objectives, and incremental exposure to immersive tools emerged as key factors in fostering trust and adoption.

Overall, these observations suggest that the feasibility of XR-based training in port environments depends less on technological maturity and more on organizational readiness. The combination of measurable performance improvements and positive qualitative feedback indicates that XR can be effectively integrated into port training programs, provided that adequate attention is given to human factors, onboarding processes, and cultural acceptance. In this sense, XR acts not only as a technological enabler, but as a catalyst for safer operational practices and more resilient human-system interaction.

6.1. Summary of the Results

The experimental results demonstrate that the integrated modelling framework implemented in COYOTE, combining objective risk computation, perceptual risk estimation, behavioral metrics, and task-based performance indicators, provides a robust and informative representation of operator behaviour in safety-critical port operations. Across the experimental campaign, participants consistently improved their operational performance, navigation efficiency, and safety behaviour, even in the absence of explicit training instructions. Duration normalized by distance decreased significantly for most operators, reflecting a more efficient internalization of spatial layout and improved trajectory planning in a dynamic yard populated by moving vehicles and occlusions. Complementarily, total exposure risk (TotRE) showed substantial reductions, confirming that users not only moved faster but did so with greater caution and more refined avoidance strategies. These improvements occurred despite increasing environmental complexity, suggesting that COYOTE effectively induces naturalistic learning loops where familiarization, feedback, and embodied experience converge to reshape behaviour in realistic ways.

The perceptual components of the model revealed equally important insights. Total perceived risk exposure (pTotRE) improved for the majority of participants, indicating that repeated immersion within the virtual environment enhanced their ability to detect cues of hazard and interpret the dynamics of the simulated terminal. However, the systematic underestimation of risk (quantified through the risk-bias index and supported by strong inferential evidence) highlights a persistent gap between real and perceived threat levels. On average, operators underestimated their risk by approximately 29%, a bias of both statistical and operational relevance. The calibration line ($pRtot = 0.15 + 0.92 \cdot Rtot$) shows that although perceived risk scales linearly with objective risk, it remains consistently shifted downward, revealing a cognitive pattern where operators intuitively feel safer than they objectively are. This is a known behavioural vulnerability in industrial environments, often exacerbated by familiarity, routine, and misinterpretation of sensory cues such as occlusions or vehicle trajectories. COYOTE's modelling framework was able not only to quantify this phenomenon but also to monitor its evolution over time, demonstrating that some users improved their calibration while others worsened, a level of granularity impossible to capture with traditional training or observational methods.

Another key finding concerns the learning dynamics revealed by mixed-effects models and correlation analyses. The absence of significant improvements under true-replication conditions (identical scenario repeats) suggests that users do not meaningfully refine their behaviour when exposed to static, repetitive conditions. In contrast, varied scenario exposure (obtained through different fog levels, AR cues, and subtle environmental randomizations) produced robust transfer learning, with participants becoming faster, safer, and perceptually more aligned across sessions. This indicates that COYOTE's ability to modify environmental parameters and embed stochasticity in object placement, visibility, and traffic flows is essential for triggering cognitive adaptation. Such findings provide strong evidence that effective training in dynamic industrial settings cannot rely solely on repetition of scripted procedures, but must involve controlled variability to stimulate generalizable skills, flexible risk assessment, and resilience to uncertainty.

From an operational standpoint, these outcomes confirm the validity and practical usefulness of COYOTE's integrated risk and behavioural models. The objective risk formulation, based on kinetic energy, probability of impact, and geometric constraints, reliably distinguished between safer and more hazardous behaviours across runs, while the normalized severity components allowed meaningful comparison between different users and missions. The perceptual risk model, grounded in sensory cues and cognitive weighting mechanisms, provided a powerful tool for diagnosing miscalibration and identifying users who either fail to recognize danger or overreact to benign stimuli. When combined with performance KPIs the simulator constructs a multi-layered behavioural fingerprint for each operator, capturing not only task execution but also the latent cognitive processes that underpin decision-making in complex operational environments.

Overall, the study demonstrates that COYOTE is not merely a training tool but an advanced experimental platform capable of capturing the interplay between human cognition, dynamic risk, and operational behaviour. The model outputs reveal subtle yet crucial aspects of operator performance, such as risk underestimation, adaptation to environmental variability, selective improvement across tasks, and inter-individual differences in learning trajectories, that cannot be detected through traditional field exercises or classroom training. The implications extend to safety management, certification, ergonomics, and workforce development: COYOTE can be used to assess operator readiness, identify latent vulnerabilities, tailor individualized training, and support the design of safer yard procedures. Furthermore, the hybrid analytical framework combining behavioural telemetry, cognitive modelling, and statistical inference paves the way for data-driven safety culture initiatives in ports and industrial terminals. In essence, the results validate the efficacy and necessity of integrating modelling, simulation, AI-driven risk estimation, and perceptual analytics to enhance safety, learning, and decision-making in modern high-risk operations.

6.2. Discussion on Model Limitations

While the COYOTE modelling framework provides an advanced and multifaceted representation of operator behaviour, risk dynamics, and perceptual processes in terminal-yard environments, its implementation is subject to several important limitations that must be acknowledged to contextualize the results and guide future developments. First, the objective risk model, although grounded in physical principles such as kinetic energy, impact probability, and spatial geometry, inevitably simplifies the complexity of real port operations. Factors such as micro-behaviours of drivers, mechanical malfunctions, unpredictable pedestrian deviations, and rare catastrophic events are either approximated or excluded due to the inherent constraints of real-time simulation. The logistic probability of impact used to modulate risk exposure, despite being flexible and interpretable, cannot fully capture the non-linear and discontinuous behaviours observed in real accidents, where small changes in human reaction time or vehicle trajectory can lead to disproportionately large differences in outcome. Similarly, the perceptual risk model relies on a reduced set of sensory variables (visual angle, occlusion, auditory directionality), which, while capturing essential components of situational awareness, does not incorporate higher-order cognitive processes such as fatigue, attentional tunnelling, stress, or prior experience, all of which significantly influence real-world hazard perception. The behavioural models embedded in COYOTE, including trajectory planning and operator responses, are also bounded by the limitations of keyboard-based control and the assumptions of a controlled virtual environment. Real operators in terminal yards contend with noise, weather variability, workplace fatigue, and multi-tasking demands that cannot be fully reproduced with current simulation fidelity.

The experimental results themselves are influenced by methodological limitations. The sample size, though representative of port contexts, remains insufficient to generalize findings to the broader workforce. The number of true replicates per experimental condition was limited for many users, reducing statistical power for within-condition learning analysis and magnifying the influence of individual variability. Moreover, participants knew they were operating in a simulated environment, which may have altered their behaviour: some users may have acted more cautiously in response to novelty, while others may have taken unrealistic risks knowing that consequences were virtual. Although COYOTE reproduces visual occlusions, traffic flows, fog, and congestion patterns, the absence of tactile cues, peripheral distraction, and real ambient noise can attenuate stress responses that are normally triggered in real operations. This may partially explain the systematic underestimation of risk observed across participants. The augmented-reality guidance system provides highly structured directional cues that may not exist in real settings, potentially producing optimistic estimates of performance gains under guidance modes.

Interoperability constraints also impose indirect limitations. Although COYOTE's architecture supports MS2G principles and maintains compatibility with external simulators and data sources, real-world interoperability introduces timing uncertainties, state-synchronization errors, and semantic mismatches across models. These issues can impact the precision of distributed simulations and complicate the integration of cognitive or autonomous-agent modules based on heterogeneous formalisms. The decision to normalize both real and perceived risk measures on a [0, 1] scale ensures comparability but risks obscuring extreme values and context-specific thresholds that might hold operational meaning in real terminal yards.

Finally, the learning and adaptation trends identified in the study must be interpreted with caution. Transfer learning effects were significant, but the underlying mechanisms remain unclear: improvements may have resulted from enhanced motor control, better scenario comprehension, reduced cognitive load due to environment familiarity, or shifts in risk-taking style. The observed heterogeneity in learning trajectories, some users improving, others deteriorating in risk calibration, suggests that individual cognitive factors, experience levels, or attentional profiles influence behaviour in ways not captured by the current models. Future versions of COYOTE should therefore incorporate richer cognitive representations, physiological monitoring, eye-tracking, and adaptive agents capable of modelling fatigue, under-stress decision-making, and risk compensation behaviours.

In summary, although COYOTE demonstrates strong potential as a training, assessment, and research platform for port-yard safety, its models remain abstractions of a complex reality. Understanding these limitations is essential for interpreting results, refining the simulation framework, and designing future studies that more accurately capture the interplay between human cognition, risk dynamics, and operational variability in real industrial environments.

6.3. Implications for Future Work and Safety Training Programs

The results obtained through the COYOTE experimental campaign provide meaningful insights for the evolution of simulation-based training, safety assessment methodologies, and the design of next-generation decision-support systems in port terminals and other high-risk industrial domains. First, the clear evidence of behavioural and safety improvements across varied simulation runs highlights the crucial role of controlled variability in virtual training environments. Traditional training methodologies, often based on repetitive execution of standard procedures, fail to capture the stochasticity and unpredictability of real operational settings. In contrast, COYOTE demonstrates that dynamic scenario generation, through randomized container placement, variable traffic flows, fog conditions, and optional augmented-reality guidance, induces robust transfer learning, with users becoming consistently faster, more situationally aware, and less exposed to objective risk. These findings suggest that future safety-training programs should incorporate structured variability rather than static repetition, leveraging simulation to expose operators to a broad spectrum of realistic hazards and environmental perturbations.

The persistent underestimation of risk observed across participants has direct implications for safety-culture development and operator certification. The systematic ~29% bias in risk perception indicates that many workers may intuitively underestimate the danger inherent in proximity to moving vehicles, occluded visibility, or suspended loads. Such miscalibration represents a latent risk factor that is often invisible in conventional on-the-job assessments. COYOTE's ability to quantify real versus perceived risk, track their divergence over time, and identify individual calibration trajectories opens new possibilities for personalized training interventions. Operators with strong risk-underestimation profiles could receive targeted modules focused on hazard recognition, auditory-visual cue interpretation, and decision-making under uncertainty. Conversely, operators with excessive risk overestimation could benefit from confidence-building modules to reduce hesitation and increase operational fluency. In this sense, COYOTE functions not only as a training tool but also as a diagnostic instrument capable of detecting cognitive vulnerabilities that directly influence safety outcomes.

From an organizational perspective, the integration of objective and subjective metrics, completion time, accuracy, collision events, TotRE, pTotRE, and risk bias, provides a multidimensional framework for evaluating workforce readiness. These indicators could support new certification protocols, enabling port authorities and terminal operators to assess operator proficiency through quantifiable metrics rather than solely observational or checklist-based evaluations. The data-rich telemetry produced by COYOTE also facilitates longitudinal monitoring, allowing safety managers to track performance trends over weeks or months and identify individuals or teams who may benefit from refresher training. Moreover, the platform could serve as a testbed for evaluating new procedures, technological innovations, or layout modifications before they are deployed in real terminals, thereby reducing the cost and risk associated with real-world trial-and-error.

The implications extend beyond human training to the design of autonomous or semi-autonomous yard systems. The risk models embedded in COYOTE provide a structured way to evaluate how human operators behave around automated vehicles, how they respond to mixed-traffic conditions, and how their risk perception changes in environments where autonomy alters traffic dynamics. These insights may guide the design of human-robot interaction protocols, collaborative safety zones, and decision-support dashboards tailored to hybrid human-autonomy operations.

For future research, several avenues emerge. The observed heterogeneity in learning and calibration trajectories underscores the importance of incorporating richer cognitive models, including attention, fatigue, and stress. Integrating physiological sensing (e.g., heart rate variability, eye-tracking, galvanic skin response) would allow COYOTE to capture latent cognitive states and adapt scenario difficulty in real time. Similarly, expanding the interoperability layer could enable multi-agent distributed simulations that connect COYOTE with port traffic models, crane simulators, cybersecurity simulators, or emergency-response environments under a unified MS2G framework. This would transform COYOTE into a node within a broader ecosystem of interconnected digital twins, supporting strategic planning, training, and operational decision-making across departmental silos.

In summary, the findings from this study strongly support the adoption of simulation-based, AI-enabled training programs as a cornerstone of modern industrial safety strategy. COYOTE demonstrates that high-fidelity virtual environments, coupled with quantitative behavioural analytics, can significantly enhance situational awareness, reduce risk exposure, and reveal cognitive blind spots that traditional approaches cannot detect. As ports continue their transition toward automation, digital transformation, and increasingly complex operational pipelines, simulation platforms like COYOTE will become essential for ensuring safe, efficient, and human-centered operations in the next generation of industrial systems.

Conclusions

The work presented in this dissertation has explored the complexity, risks, and operational challenges that characterize modern port terminals and industrial plants, proposing a comprehensive technological and methodological solution capable of addressing them. These environments, increasingly shaped by dense traffic flows, hazardous materials, sophisticated machinery, and fragmented streams of information, demand new paradigms for safety, training, and decision support. The research has demonstrated that an integrated approach based on Modeling and Simulation, Extended Reality, Artificial Intelligence, and Data Analytics is not only feasible but necessary to sustain the evolving needs of these critical infrastructures. Through this approach, the dissertation has developed a coherent framework capable of enhancing situational awareness, improving operator performance, and supporting strategic decision-making.

At the core of this transformation lies the COYOTE simulator, a real-time, physics-driven digital environment that faithfully reproduces the operational reality of container yards. The simulator represents a synthesis of diverse modeling paradigms: discrete-event logic for process flow, agent-based mechanisms for autonomous entity behavior, stochastic models for uncertainty, physics-based interactions for realism, and cognitive-behavioral constructs for capturing human decision patterns. By combining Unity's 3D engine with advanced navigation algorithms, dynamic obstacle-avoidance logic, and AI-driven behavioral rules, COYOTE provides a multi-layered, immersive representation of port operations that goes far beyond traditional training tools. Vehicles move autonomously through the yard, cranes position and re-handle containers with realistic motion constraints, and operators interact with a world that responds to their decisions in real time. Dangerous goods leaks, reduced visibility, unpredictable machine trajectories, and environmental variations are not merely visual elements but active components of the simulation, affecting risk exposure and mission outcomes.

A particularly innovative contribution of this work is the integration of Extended Reality. The use of immersive headsets, spatial audio, and multisensory cues places the operator at the center of a realistic yet controlled environment, enabling training sessions that would be too hazardous, logistically complex, or economically impractical to perform in reality. Within this space, users must adapt to unpredictable events, manage their movements with respect to autonomous vehicles, detect hazards, and evaluate container conditions while maintaining situational awareness. The XR layer does not serve only as a visual enhancement but becomes a cognitive amplifier that helps induce human responses comparable to those observed on the real yard, creating a fertile ground for the study of human factors and decision-making under stress.

The experiments carried out with professional operators from PSA Genova Pra' represent one of the most significant outcomes of this research. Through more than three hundred and fifty simulation trials conducted under varying levels of traffic intensity, environmental difficulty, and scenario complexity, the study observed substantial improvements in performance indicators. Operators progressively enhanced their task accuracy, reduced their exposure to hazards, and completed missions in shorter times without compromising safety. The simultaneous decrease in both real and perceived risk exposure indicates a maturation of their operational behavior: more deliberate navigation, better assessment of vehicle trajectories, and better planning of their own movements inside the yard. These empirical findings confirm the fundamental hypothesis of the dissertation: a well-designed simulation framework supported by XR technologies can meaningfully improve both competence and situational awareness.

Beyond training and behavioral analysis, the dissertation expands its scope to include a Digital Twin architecture that integrates simulation models, data streams, and AI-driven reasoning. This architecture reflects a broader vision of port terminals as cyber-physical ecosystems in which safety, efficiency, and resilience depend on dynamic interactions among physical assets, digital systems, human operators, and external threats. By linking real-time operational data with predictive simulation, the Digital Twin becomes a strategic tool capable of supporting decision-makers in assessing accident evolution, testing emergency responses, optimizing workflows, and anticipating the effects of hybrid threats. Intelligent agents embedded within this architecture simulate cyber intrusions, equipment failures, antagonistic actions, and cascading effects, offering a realistic yet controllable environment in which to explore multi-domain vulnerabilities.

Taken as a whole, the contributions of this dissertation advance the state of the art in several domains of research and practice. First, they demonstrate the scientific and operational maturity of the MS2G paradigm when extended with XR and AI components. Second, they confirm that simulation-supported training can significantly enhance operator behavior and risk awareness even in highly complex and safety-critical environments. Third, they show that Digital Twin frameworks grounded in interoperable simulation can provide a powerful foundation for strategic decision-making in the face of uncertainty. Finally, they illustrate how human-factor modeling, data analytics, and immersive interaction can be combined to create a new generation of tools for safety engineering and operational resilience.

Despite these achievements, the work also reveals directions in which current technologies and methodologies can evolve further. The cognitive and emotional dimensions of human behavior, although indirectly modeled through measures of risk perception and task performance, could benefit from the integration of physiological sensors or neurocognitive inputs, enabling more accurate and adaptive training strategies. The Digital Twin, while validated through simulation experiments, offers the potential for full operational deployment through real-time IoT integration, sensor networks, and automatic data ingestion from yard management systems. Hybrid threat modeling, currently implemented through intelligent agents and scenario-based analysis, could be expanded into a broader multi-domain simulation framework encompassing cyber, physical, informational, and cognitive layers. Moreover, the integration of reinforcement learning would allow virtual vehicles and agents to autonomously improve their strategies, further enriching the realism and scientific value of the simulation environment.

While the experimental activities presented in this thesis are centered on maritime ports and container terminals, the proposed framework was intentionally designed as a domain-agnostic solution for safety-oriented training and decision support in complex logistics systems. Ports represent a particularly demanding case study due to the coexistence of dense traffic, heterogeneous assets, hazardous materials, and strong coupling between human actions and system-level outcomes. However, these characteristics are not unique to the maritime domain and can be observed, with different manifestations, in other logistics hubs such as airports, rail terminals, and large intermodal facilities.

From a structural perspective, these environments share common features that make the transfer of the framework viable. All involve multi-actor operational spaces where human operators interact with vehicles, automated systems, and infrastructure under time pressure and safety constraints. Airports, for instance, present complex ground operations involving aircraft, service vehicles, and personnel in confined areas, while rail terminals are characterized by tightly synchronized movements, signaling dependencies, and high consequences of procedural deviations. In these contexts, risk does not arise from isolated components, but from the interaction between physical assets, operational procedures, environmental conditions, and human behavior, mirroring the dynamics observed in port terminals.

At the modeling and simulation level, the framework's reliance on interoperable simulation and agent-based representations enables straightforward adaptation across domains. Core elements such as entity behavior modeling, event-driven interactions, and performance and risk indicators can be preserved, while domain-specific assets, rules, and constraints are reconfigured. For example, yard vehicles and container handling operations can be replaced by aircraft ground support equipment or rail shunting units, while maintaining the same logic for measuring success rates, exposure to risk, and procedural correctness. This modularity ensures that the framework remains extensible without requiring a complete redesign for each new application domain.

The XR layer further enhances cross-domain scalability by acting as a unifying interface between simulation models and human operators. Immersive environments allow complex and hazardous scenarios to be represented with high fidelity, regardless of the underlying logistics domain. By adapting virtual assets, interaction metaphors, and hazard representations, the same XR-based approach can support training and rehearsal for airport ground crews, rail maintenance personnel, or operators in intermodal hubs. Importantly, the XR component facilitates experiential learning and situational awareness development in a controlled setting, reducing reliance on real-world exposure for safety-critical training.

Human factors and organizational aspects also support the transferability of the framework. Across logistics hubs, safety performance is strongly influenced by operator awareness, workload management, communication, and adherence to procedures. The results obtained in the port context suggest that simulation- and XR-based training can positively influence these dimensions by aligning perceived risk with actual operational risk. This mechanism is largely independent of the specific domain and can therefore be leveraged in other logistics environments, provided that appropriate calibration and contextualization are performed.

Finally, the broader applicability of the framework opens several avenues for future research and industrial deployment. Cross-domain comparative studies could investigate how risk exposure, technology acceptance, and training effectiveness differ between ports, airports, and rail terminals. Additionally, integration with digital twin initiatives and real-time operational data could further enhance the framework's capability to support not only training, but also strategic planning and resilience analysis across interconnected logistics systems. In this sense, the research contributes a scalable methodological foundation rather than a domain-specific solution, positioning modeling, simulation, and XR as key enablers for safer and more resilient logistics infrastructures.

In conclusion, the dissertation demonstrates that combining high-fidelity simulation, data-driven intelligence, and immersive technologies can fundamentally reshape the way safety, training, and decision support are approached in ports and industrial plants. The COYOTE simulator and the Digital Twin architecture represent not only technological achievements but also conceptual advances that redefine the relationship between human operators and complex operational environments. By providing a scientifically rigorous, empirically validated, and technologically innovative framework, this work lays the foundation for a new generation of intelligent and adaptive safety systems. Beyond their contribution to academic research, these systems have the potential to influence real-world practices, fostering a more resilient, efficient, and safety-oriented culture in critical infrastructures. The path traced by this dissertation points toward a future in which simulation is not a complementary tool but a central component of the operational and strategic management of ports, enabling continuous learning, proactive planning, and enhanced resilience in an increasingly uncertain and interconnected world.

Bibliography

[1] G. Vaggelas, "Port labour challenges and opportunities in the era of port automation and digitalization," in *The International Maritime and Logistics Conference “Marlog 9”*, 2020.

[2] N. G. Leveson, Engineering a safer world: Systems thinking applied to safety, MIT Press, 2016.

[3] UNCTAD, "UNCTAD.org," April 2025. [Online]. Available: https://unctad.org/news/shipping-data-unctad-releases-new-seaborne-trade-statistics.

[4] J. Verschuur, E. E. Koks and J. W. Hall, "Ports’ criticality in international trade and global supply-chains," *Nature Communications 13,* p. 4351, 2022.

[5] C. Arango, A. Bruzzone, L. Nicoletti and P. Cortes, "A simulation-based approach for resource management and control: A real case study of a spanish container terminal," in *International Conference on Harbour, Maritime and Multimodal Logistics Model*, 2012.

[6] O. F. Knudsen and B. Hassler, "IMO legislation and its implementation: Accident risk, vessel deficiencies and national administrative practices," *Marine Policy 35(2),* pp. 201-207, 2011.

[7] Souaiby, N., & El-Hussein, M., "Beirut explosion, the full story.," *Med Emergency, MJEM, 28,* 2020.

[8] B. Fabiano, F. Currò, A. P. Reverberi and R. Pastorino, "Port safety and the container revolution: A statistical study on human factor and occupational accidents over the long period.," *Safety science,* pp. 980-990, 2010.

[9] Z. A. Kadir, R. Mohammad, N. Othman, S. Chelliapan and A. Amrin, "Risk assessment of human risk factors in port accidents.," *International Journal of Mechanical Engineering and Technology, 8(11),* pp. 535-551, 2017.

[10] J. Reason, Managing the Risks of Organizational Accidents, Routledge, 2016.

[11] E. Hollnagel, "Resilience engineering and the built environment," *Building Research & Information, 42(2),* pp. 221-228, 2014.

[12] O. M. Belcore, M. Di Gangi and A. Polimeni, "Connected vehicles and digital infrastructures: A framework for assessing the port efficiency," *Sustainability, 15(10),* p. 8168, 2023.

[13] N. Tsolakis, D. Zissis, S. Papaefthimiou and N. Korfiatis, "Towards AI driven environmental sustainability: an application of automated logistics in container port terminals," *International Journal of Production Research, 60(14),* pp. 4508-4528, 2022.

[14] A. Bruzzone, M. Massei, K. Sinelshchikov, A. Giovannetti, R. Ferrari, A. de Paoli, B. Gadupuri, A. Reverberi, G. Fancello, M. Frosolini, T. Vairo, G. Piroddi, F. Gaborit and G. Paoli, "Innovative virtual lab for improving safety and port operations.," in *proceedings of the 24th International Conference on Harbor, Maritime and Multimodal Logistic Modeling & Simulation, HMS*, Rome, 2022.

[15] A. G. Bruzzone, M. Massei, K. Sinelshchikov, M. Gotelli, A. Giovannertti, A. De Paoli, R. Ferrari and M. Cardelli, "Operations in Ports supported by Simulation, XR and AI," in *25th International Conference on Harbor, Maritime and Multimodal Logistic Modeling & Simulation.*, Athens, 2023.

[16] A. Bruzzone, M. Massei, K. Sinelshichikov, A. De Paoli, A. Giovannetti, F. Longo, G. Fancello, T. Vairo, C. Giliberti and R. Mariconte, "Promoting safety, security, awareness and productivity in port plants," in *5th International Conference on Industry 4.0 and Smart Manufacturing (ISM 2023)*, Prague, 2024.

[17] W. Karwowski, A. Szopa and M. M. Soares, Handbook of standards and guidelines in human factors and ergonomics, CrC Press, 2021.

[18] D. E. Mathew, D. U. Ebem, A. C. Ikegwu, P. E. Ukeoma and N. F. Dibiaezue, "Recent emerging techniques in explainable artificial intelligence to enhance the interpretable and understanding of AI models for human," *Neural Processing Letters, 57(1),* p. 16, 2025.

[19] A. Bruzzone, "MS2G as pillar for developing strategic engineering as a new discipline for complex problem solving," in *Proceedings of the European Modeling and Simulation Symposium*, Budapest, 2018.

[20] F. Goerlandt and J. Montewka, "Maritime transportation risk analysis: Review and analysis in light of some foundational issues," *Reliability Engineering & System Safety,138,* pp. 115-134, 2015.

[21] P. L. Pallis, "Port risk management in container terminals.," *Transportation research procedia, 25,* pp. 4411-4421, 2017.

[22] K. Wawak and T. Małkus, "Managing risks in the transport of dangerous goods: the role of international regulations," *Acta Logistica, 12 (3) ,* 2025.

[23] S. Nguyen, P. S. L. Chen and Y. Du, "Container shipping operational risks: an overview of assessment and analysis," *Maritime Policy & Management, 49 (2),* pp. 279-299, 2022.

[24] T. Xie, X. Lu, G. Wang and F. Lin, "Research on safety risk, prevention and control in port dangerous goods container yard," *Journal of Physics: Conference Series,1910,* 2021.

[25] E. D. Kusumawati, K. Karjono and K. Karmanis, "Review of port management integrated digitization system: A pathway to efficient and sustainable port operations," *Maritime Park: Journal of Maritime Technology and Society,* pp. 55-60, 2023.

[26] N. Yaakub, N. F. Abdul Razak, M. R. Baharuddin and M. A. Mohd Noor, "Stress at the Helm: the Prevalence and Key Risk Factors of Occupational Stress Among Port Workers in the Port Terminal Industry," *Malaysian Journal of Medicine & Health Sciences, 20(3),* 2024.

[27] Y. Wang, X. Shi and D. Xu, "Relationship between overconfidence and risky behavior among ship crew," *Transportation research record, 2674(9,* pp. 500-510, 2020.

[28] M. A. &. F. H. Budiyanto, "Risk assessment of work accident in container terminals using the fault tree analysis method," *Journal of Marine Science and Engineering, 8(6),* p. 466, 2020.

[29] R. Parasuraman and D. H. Manzey, "Complacency and bias in human use of automation: An attentional integration," *Human factors, 52(3),* pp. 381-410, 2010.

[30] F. Tao, Q. Qi, A. Liu and A. Kusiak, "Data-driven smart manufacturing," *Journal of manufacturing systems, 48,* pp. 157-169, 2018.

[31] S. El Mekkaoui, L. Benabbou and A. Berrado, "Machine learning models for efficient port terminal operations: Case of vessels' arrival times prediction," *IFAC-PapersOnLine, 55(10),* pp. 3172-3177, 2022.

[32] Y. Ding, Z. Zhang, K. Chen, H. Ding, S. Voss, L. Heilig, Y. Chen and X. Chen, "Real-Time Monitoring and Optimal Resource Allocation for Automated Container Terminals: A Digital Twin Application at the Yangshan Port," *Journal of Advanced Transportation, 2023(1),* 2023.

[33] I. Progoulakis, N. Nikitakos, D. Dalaklis, A. Christodoulou, A. Dalaklis and R. Yaacob, "Digitalization and cyber physical security aspects in maritime transportation and port infrastructure," *Smart ports and robotic systems: Navigating the waves of techno-regulation and governance,* pp. 227-248, 2023.

[34] S. Karnouskos, P. Leitao, L. Ribeiro and A. W. Colombo, "Industrial agents as a key enabler for realizing industrial cyber-physical systems: Multiagent systems entering industry 4.0.," *IEEE Industrial Electronics Magazine, 14(3),* pp. 18-32, 2020.

[35] M. Indriani, "International Maritime Organization (IMO) guidelines on maritime cyber risk management: Isu dan Tantangan Bagi Indonesia," *Jurnal Majelis, 1,* 2019.

[36] L. Bainbridge, "Ironies of automation," *Analysis, design and evaluation of man–machine systems,* pp. 129-135, 1983.

[37] E. Hollnagel, Safety-I and safety-II: the past and future of safety management, CRC press, 2018.

[38] J. Banks, Discrete event system simulation, Pearson Education India, 2005.

[39] A. Borshchev and A. Filippov, "From system dynamics and discrete event to practical agent based modeling: reasons, techniques, tools," in *22nd International Conference of the System Dynamics Society, July 25 - 29*, Oxford, England , 2004.

[40] S. De Marchi and S. Page, "Agent-based models," *Annual Review of political science,* vol. 17(1), pp. 1-20, 2014.

[41] D. Helbing, "Agent-based modeling," *Social self-organziation,* pp. 25-70, 2012.

[42] E. Babulak and M. Wang, Discrete Event Simulation, 2010.

[43] B. K. Bala, F. Arshad and K. Noh, System Dynamics, 2017.

[44] E. Salas, S. I. Tannenbaum, K. Kraiger and K. A. Smith-Jentsch, "The science of training and development in organizations: What matters in practice," *Psychological science in the public interest, 13(2),* pp. 74-101, 2012.

[45] O. Balci, "A life cycle for modeling and simulation. Simulation," *Simulation: Transactions of the Society for Modeling and Simulation International , 88(7),* pp. 870-883, 2012.

[46] UNCTAD, "Review of Maritime Transport," United Nations Conference on Trade and Development (UNCTAD), Geneva, Switzerland, 2021.

[47] T. Nemer, "The Beirut port explosion: A geoscience perspective," *Seismol. Res. Lett., 21,* p. 2093–2098., 2021.

[48] P. Prugh, "Historical record of ammonium nitrate disaster," *Process Saf. Prog., 39,* 2010.

[49] O. Knudsen and B. Hassler, " IMO legistlation and its implementation: Accident risk, vessel deficiencies and national administrative," *Maritime Policy, 35,* pp. 201-207, 2011.

[50] J. Chen, H. Zheng, L. Wei, Z. Wan, R. Ren, J. Li and Y. Bay, "Factor diagnosis and future governance of dangerous goods ac-cidents," *Environ. Pollut., 257,* 2020.

[51] Z. Pawel and P. Katarzyna, "Analysis of ship accidents based on European statistical surveys.," *Sci. J. Marit. Univ. Szczec, 68,* pp. 17-25, 2021.

[52] P. Antão, M. Caledèron, M. Puig, A. Michail, C. Wooldridge and R. Darbra, "Identification of occupational health, safety, security (OHSS) and environmental performance indicators in port areas," *Safety Sciences, 85,* pp. 266-275, 2016.

[53] R. Khan, J. Yin, F. Mustafa and F. A. O. A., "A data centered human factor analysis approach for hazardous cargo accidents in a port environment," *J. Loss Prev. Process Ind., 75,* 2022.

[54] M. Lecue and R. Darbra, "Accidents in European ports involving chemical substances> Characteristics and trends," *Safety Sciences, 115,* p. 278–284, 2019.

[55] C. Dominguez-Péry, L. Vuddaraju, I. Corbett-Etchevers and R. Tassabehji, "Reducing maritime accidents in ships by tackling humanerror: A bibliometric review and research agenda, 6," *Shipping Trade,* 2021.

[56] J. Kim, J. Kim, G. Lee and J. Park, "MachineLearning/basedmodelsforaccidentpredictionataKoreancontainerport," *Sustainability, 13,* 2021.

[57] L. Kretschmann, "Leading indicators and maritime safety: Predicting future risk with a machine learning approach," *Jorunal of Shipping Trade, 5,* 2020.

[58] A. Rawson, M. Brito, Z. Sabeur and L. Tran-Thanh, "A machine learning approach for monitoring ship safety in extreme weather," *Safety Sciences, 141,* 2021.

[59] Ü. Atak and Y. Arslano˘glu, "Machine learning methods for predicting marine port accidents: A case study in container terminal," *Ships Offhore Structure, 17,* pp. 2480-2487, 2022.

[60] L. Yujian and L. Bo, "A normalized Levenshtein distance metric," *IEEE transactions on pattern analysis and machine intelligence,* vol. 29(6), pp. 1091-1095, 2007.

[61] S. Cahyono, "Comparison of document similarity measurements in scientific writing using Jaro-Winkler Distance method and Paragrafh Vector method," in *IOP Conference Series: Materials Science and Engineering* , 2019.

[62] J. Lauronen, L. Sakari and T. Lehtonen, "How simulation training can benefit from virtual reality extensions? Case: A virtual reality Extension to a simulated ship bridge for emergency steering training," in *International conference on applied human factors and ergonomics*, 2021.

[63] D. Mourtzis and J. Angelopoulos, "Development of an extended reality-based collaborative platform for engineering education: Operator 5.0," *Electronics,* 2023.

[64] M. Gawlik-Kobylinska, M. Urban, G. Gudzbeler and A. Misiuk, "Simulation-based training in the use of the EU-SENSE CBRN reconnaissance device: a case study.," in *Proceedings of the 11th International Defence and Homeland Security Simulation Workshop*, 2021.

[65] L. Birta and G. Arbez, Modelling and Simulation, London: Springer, 2013.

[66] J. Kleijnen, "Design and analysis of simulation experiments," in *International workshop on simulation*, 2015.

[67] D. Hodson and R. Hill, "The Art and science of live, virtual and constructive simulatoin for test and analysis," *The journal of Defense Modeling and Simulation,* vol. 11(2), pp. 77-89, 2014.

Table of Figures

Tables