Proc. of the European Safety and Reliability Conference (ESREL 2026)
Edited by José C. Matos, Paulo B. Lourenço, Michael Beer, Daniel Oliveira,
Edoardo Patelli, Hélder S. Sousa, Rui A. Silva and Monica Santamaria

doi: 10.3850/ESREL2026061419_esrel26-p27803-cd

Bridging the risk awareness gaps between the safety science and AI domains: A perspective with engineering applications

Paul Lee

Department of Marine Technology, Norwegian University of Science and Technology, Norway.
E-mail: paul.lee@ntnu.no

Ingrid Bouwer Utne

Department of Marine Technology, Norwegian University of Science and Technology, Norway.
E-mail: ingrid.b.utne@ntnu.no

Ekaterina Kim

Department of Marine Technology, Norwegian University of Science and Technology, Norway.
E-mail: ekaterina.kim@ntnu.no

In recent years, significant advancements have been observed in the artificial intelligence (AI) domain with a growing adoption in various safety-critical engineering applications. However, numerous AI-related incidents indicate that its intelligence cannot guarantee its safety. This has raised the need for safe and risk-aware AI, enabling deeper understanding of system and operational risk. Efforts have been put in the AI domain to tackle the challenge of "risk awareness", but with a considerable difference from the safety science domain. Therefore, the aim of this article is to identify and assess differences in how risk awareness is understood and used in the safety science and AI domains, focusing on engineering applications. A systematic literature review has been conducted to a) analyze the use of the risk awareness concept in the safety science domain and synthesize a generic description, b) analyze the risk awareness concept in the AI domain, and c) assess the differences between these concepts to identify the gaps. The result highlights the need for proper integration of safety science knowledge for safe AI solutions to be developed and achieved. Furthermore, this study provides a way forward towards future research directions for risk-aware, safe, and trustworthy AI agents.

Keywords: Artificial intelligence, autonomy, safety, risk, risk awareness, risk aware AI.

1. Introduction

1.1. *Background*

Industries have been shifting towards increased autonomy in various engineering applications powered by artificial intelligence (AI), aiming for sustainability and financial gain (Ullah et al. (2020)). Specifically, AI has been employed in relation to various cognitive tasks, where in certain applications it surpassed benchmarks set by previous algorithms or even humans (Shakya et al. (2023)). At higher levels of autonomy, AI is deemed to permeate more into decision making related to safety-critical operations of autonomous systems (Dokas (2026)).

Despite the rapid advances in AI, addressing safety-related issues remains a challenge (Floridi et al. (2018)). Recent incidents indicate the potential harm due to untrustworthy decisions provided by AI during safety-critical scenarios (Liu et al. (2022)), such as in the 216 crashes observed during the testing of advanced driver-assistance systems on highways (Opletal (2025)). This has further raised the need for safer AI and for risk awareness in systems, enabling a deeper understanding of what risk is during operation. Despite efforts to address risk awareness in the AI domain, a considerable difference seems to exist compared to how a risk-aware system is understood in the safety science domain.

1.2. *Aim & objectives*

The aim of this article is to identify and assess differences in how risk awareness is understood and used in the safety science and AI domains, focusing on engineering applications. Furthermore, the article provides recommendations for future research towards bridging the risk awareness gaps between the two domains.

2. Literature review approach

To address the aim, a systematic review of the literature has been performed, consisting of the following phases, conducted on 31st December 2025.

2.1. *Bibliographic database selection*

The first phase entails the selection of bibliographic databases that contain the most relevant corpus of literature. In this article, Scopus has been selected.

2.2. *Documents search*

The second phase entails the search string "*risk aware**" *OR* "*aware* of risk*" employed in the *article title*, *abstract*, *author keywords* search field to capture documents related to risk awareness. Specifically, the double quotes "" have been used to search for a loose phrase, the asterisk wildcard * has been used to replace zero or more characters (i.e., *aware** finds *aware*, *awareness*), while the Boolean operator *OR* has been used to identify documents from either of the loose phrases.

2.3. *Documents filtering*

The third phase entails the refinement of the identified documents with employed filters: *subject area*: *engineering*; *document type*: *article*; *source type*: *journal*; and *language*: *English* to filter in journal articles of peer-reviewed quality with engineering applications.

2.4. *Documents analysis*

The fourth phase entails the analysis of the filtered documents in the safety science and AI domains identified based on the *author keywords* and *source title*. The list of AI-related terms derived from the *author keywords* for the identification of the AI domain is presented in Table A.1. The list of *source titles* used for the identification of the safety science domain that do not contain the AI-related terms is presented in Table A.2, where the top ten according to the number of publications is provided.

3. Results & discussion

A total of 753 journal articles were identified in all domains from 1975 to 2025 with an annual growth rate of 11.5%, as presented in Figure 1. Specifically, 2025 alone constituted 30.3% of all journal articles, which indicates the recent research trend towards risk awareness with engineering applications.

In the AI domain, 166 journal articles were identified with an annual growth rate of 49.6%, Specifically, 2025 alone constituted 50.6% of all journal articles, which indicates the recent research trend towards risk awareness in the AI domain. In the safety science domain, 107 journal articles were identified with an annual growth rate of 6.5%.

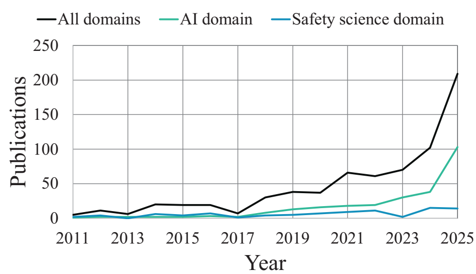

Fig. 1. Publications per year.

3.1. *Risk awareness concept in the safety science domain*

Risk awareness has been addressed in the safety science domain from multiple perspectives, including its relation to risk perception, mitigation, management, and communication, preparedness and proactiveness, resilience, self-efficacy, self-reliance, and self-awareness, as also presented in Figure 2. Though not identified in the literature search, in the resilience-based early warning indicator method, risk awareness has been investigated as part of resilience, including risk understanding, anticipation, and attention (Øien et al. (2012)).

Fig. 2. Word cloud of the *author keywords* in the safety science domain using bibliometrix.

When it comes to the use of the term risk awareness, it largely falls into two "camps". The first camp derives it as an extension of *awareness*, which refers to '*the process of becoming informed and knowledgeable*' (Gibson (2003)). For instance, Latvakoski et al. (2022) referred to risk awareness as the acknowledgement about risk, risk prevention, and mitigation actions. Cisternas et al. (2024) considered it as the information and knowledge about hazards and the ability to recognize potential risk that influences the preparedness measures to be taken. Castañeda et al. (2024) referred to it as the beliefs and knowledge about the hazard.

The second camp derives it as an extension of *situation awareness*, which refers to '*the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future*' (Endsley (2017)). For instance, Bellet and Banet (2012) referred to risk awareness as the dual cognitive abilities of hazard detection through hazard perception and risk management through risk assessment. Chatzimichailidou and Dokas (2016) considered it as the ability to recognize imminent threats or vulnerabilities that can lead to accidents. Lyu et al. (2024) referred to it as hazard identification, assessment of the level of danger, and decision-making.

These two camps exhibit some intrinsic differences, while serving some common overarching objectives. Specifically, the former one can be decomposed into the following three components of 1) information, 2) knowledge, and 3) behavior about risk, whereas the latter one into 1) perception, 2) comprehension, and 3) projection of risk. In that sense, the former camp refers mostly to the properties of risk awareness that answer to the *what*, whereas the latter camp refers to the functions of it that answer to the *how*. Furthermore, the commonality across them is that each component serves the objectives for hazard identification, risk assessment, and decision making, respectively, in a complementary manner.

Based on the above, a generic description of risk awareness may be synthesized as the '*process comprising hazard identification through the perception of risk information, risk assessment through the comprehension of risk knowledge, and decision making through the projection of risk behavior*', as also presented in Figure 3.

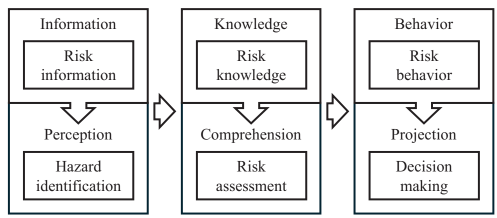

Fig. 3. Synthesized generic process of risk awareness.

3.2. *Risk awareness concept in the AI domain*

Risk awareness has been addressed in the AI domain by employing techniques from various fields of AI, as presented in Figure 4. Specifically, analysis of the *author keywords* suggests that the most frequently used ones related to the fields of AI are in descending order "*reinforcement learning*" (RL) 13.3%, "*deep learning*" (DL) 11.4%, "*machine learning*" (ML) 9.0%, and "*artificial intelligence*" 4.8%. The used risk awareness terms are reviewed based on the above fields of AI.

3.2.1. *Reinforcement learning*

Typical terms related to risk awareness using RL techniques included, but are not limited to:

Measures, e.g., Liu et al. (2025) optimized conditional value-at-risk (CVaR) for supply chain management by maintaining probability distributions over returns and integrating it with constraints into the critic's loss function.

Fig. 4. Word cloud of the *author keywords* in the AI domain using bibliometrix.

Risk functions, e.g., Sun et al. (2023) introduced a collision risk function for robot crowd navigation by evaluating the relative positions and velocities of pedestrians to assess the probability of collision and fusing it with a learning based feature extraction for proximal policy optimization (PPO).

Potential fields, e.g., Wu et al. (2024) employed artificial potential fields for line changing of ground vehicles by incorporating them in task decomposition of branching networks using twin delayed deep deterministic policy gradient (TD3).

Constraints, e.g., Xu and Zhang (2026) introduced dynamic constraints adjustment for supply chain management by reshaping the parameter of a constrained policy optimization based on real-time confidence estimations.

Safety filters, e.g., Lederer et al. (2025) used safety filters for inhibitory control by defining probabilistic safety conditions for a value function to realize probabilistic safety guarantees from its policy.

Barrier functions, e.g., Hu et al. (2025) incorporated barrier functions for ground vehicles by combining risk indicators with Bayesian networks and incorporating them into the loss function of TD3.

Probabilities, e.g., Zhang et al. (2025) predicted the probabilities of ground vehicles' conflict severity with jaywalkers by employing an encoder-decoder module to estimate the severity levels and inform the state space and reward functions.

Risk levels, e.g., Wijegunawardana et al. (2025) integrated risk levels of neighboring cells for cleaning robots by employing failure mode and effect analysis and incorporating them into the reward function to minimize hazardous interactions during complete coverage tasks.

3.2.2. *Deep learning*

Typical terms related to risk awareness using DL techniques, apart from RL techniques, include, but are not limited to:

Risk metrics, e.g., Sun et al. (2024) utilized risk metrics for ground vehicles trajectory prediction by incorporating time-to-collision and moving direction into an edge-enhanced local graph attention encoder.

Risk indices, e.g., Su et al. (2025) developed a collision risk index for lane-changing ground vehicles trajectory predictions by calculating collision probability and intensity based on the predictions estimated by a spatio-temporal graph attention transformer with gated recurrent unit.

Risk levels, e.g., Lee et al. (2025) classified distinct risk levels for scooter riders by labeling time-series data with risk scores and extracting the most important driving behavior through an autoencoder.

Probabilities, e.g., Shree et al. (2021) estimated a probabilistic distribution of scene danger levels for multi-robot search and rescue missions by assessing the semantic similarity between expert-provided natural language descriptions of hazards and robot-captured images and using these danger estimates during path planning.

Uncertainties, e.g., Jha et al. (2025) quantified the epistemic uncertainty for e-commerce advertising by utilizing Bayesian deep learning with self-attention mechanisms to estimate the return on advertisement spend with full probability distribution.

3.2.3. *Machine learning*

Typical terms related to risk awareness using ML techniques, apart from RL and DL techniques, include, but are not limited to:

Measures, e.g., Sabbatella et al. (2024) adopted CVaR for information source optimization by incorporating it into the objective function and utilizing Gaussian process on different risk profiles.

Constraints, e.g., Wu et al. (2021) used a constraint for lithofacies classification under data drift by imposing constraint preservation in the objective function and weighting mechanism to improve the accuracy of pseudo labels.

Likelihood, e.g., Rawson et al. (2021) quantified the likelihood for vessels' incidents in adverse weather by utilizing vessel traffic, weather and casualty data.

Uncertainties, e.g., Yin and Li (2025) quantified the uncertainty for forecasting price strategies by utilizing variance bounds through Gaussian process regression and kernel-based modeling.

3.2.4. *Artificial intelligence*

Typical terms related to risk awareness using AI techniques, apart from RL, DL, and ML techniques, include, but are not limited to:

Risk levels, e.g., Naderpour et al. (2014) estimated risk levels for process systems by employing dynamic Bayesian networks to model the temporal dependencies of abnormal situations and a fuzzy logic system to estimate the likelihood and severity.

Risk maps, e.g., Bremnes et al. (2024) developed risk maps for marine robotics path planning by constructing a holistic spatio-temporal risk model through the integration of systems theoretic process analysis (STPA) into a hierarchical Bayesian belief network to plan path by weighing mission objectives against the probability and consequences of hazardous events.

Supervisory risk control, e.g., Rothmund et al. (2025) developed a supervisory risk controller for industrial drone inspection by employing dynamic decision network derived from STPA to autonomously adjust operational parameters in real-time and maintain acceptable risk levels.

3.3. *Discussion*

3.3.1. *Assessing the risk awareness gaps*

The aforementioned risk awareness concepts in the AI domain are assessed against the generic description of risk awareness in the safety science domain. Specifically, relevant performance criteria (C) are proposed from the synthesized generic process of risk awareness based on the following qualitative mapping process: $C_{1.1\text{-}1.2}$ refer to the definition of risk followed by the performance of hazard analysis; $C_{2.1\text{-}2.5}$ refer to the performance of causal, frequency, and consequence analysis for the risk presentation followed by the performance of risk evaluation and reduction; and $C_{3.1\text{-}3.2}$ refer to the decision making that incorporates and controls risk. The mapping results are presented in Table 1.

The mapping of the risk awareness concepts in the AI domain identifies some limitations, most notably:

Risk definition. The definition of risk is often omitted, considering that approximately only 18.9% of them define it (e.g., see $C_{1.1}$ in Table 1). Without a proper definition, risk is often misunderstood, such as being an equivalent of an expected value, probability, or uncertainty, as critically discussed by Aven (2011). This limitation is consequently reflected in the quality of the investigated risk awareness concepts. An exception is some works, such as Bremnes et al. (2024), where the definition of risk is given, and risk is clearly understood as the combined answer to what can go wrong, how likely it is, and what the consequences are (Kaplan and Garrick (1981)).

Systematic risk assessment methods. The employment of systematic risk assessment methods are often omitted, considering that approximately only 9.4% of them conduct such methods (e.g., see $C_{1.2}$, $C_{2.1-2}$, and $C_{2.4-5}$ in Table 1). Hence, despite the commonly used measures related to risk (e.g., $C_{2.3}$ in Table 1), they lack proper justification on how risk is analyzed, measured, assessed, evaluated, and controlled, leading to a considerable difference compared to its comprehensive understanding in the safety science domain. An exception is some works, such as Rothmund et al. (2025), where hazard analysis, causal and frequency analysis methods are employed to systematically quantify risk based on consequences and probabilities and provide input to develop safe AI and autonomy controllers.

3.3.2. *Recommendations for bridging the gaps in future research*

To enhance risk awareness and safety in the AI domain, some recommendations for future research are:

Journal article	Method	$C_{1.1}$	$C_{1.2}$	$C_{2.1}$	$C_{2.2}$	$C_{2.3}$	$C_{2.4}$	$C_{2.5}$	$C_{3.1}$	$C_{3.2}$
Liu et al. (2025)	RL	✗	✗	✗	✗	✗	✗	✗	✗	✗
Sun et al. (2023)	RL	✗	✗	✗	✗	✗	✗	✗	✗	✗
Wu et al. (2024)	RL	✓	✗	✗	✗	✗	✗	✗	✗	✗
Xu and Zhang (2026)	RL	✗	✗	✗	✗	✗	✗	✗	✗	✗
Lederer et al. (2025)	RL	✗	✗	✗	✗	✗	✗	✗	✗	✗
Hu et al. (2025)	RL	✓	✗	✗	✗	✓	✗	✗	✓	✓
Zhang et al. (2025)	RL	✗	✗	✗	✗	✓	✗	✗	✓	✓
Wijegunawardana et al. (2025)	RL	✗	✓	✓	✓	✓	✗	✗	✓	✗
Sun et al. (2024)	DL	✗	✗	✗	✗	✗	✗	✗	✗	✗
Su et al. (2025)	DL	✗	✗	✗	✗	✓	✗	✗	✓	✗
Lee et al. (2025)	DL	✗	✗	✗	✗	✗	✗	✗	✗	✗
Shree et al. (2021)	DL	✗	✗	✗	✗	✗	✗	✗	✗	✗
Jha et al. (2025)	DL	✗	✗	✗	✗	✗	✗	✗	✗	✗
Sabbatella et al. (2024)	ML	✗	✗	✗	✗	✗	✗	✗	✗	✗
Wu et al. (2021)	ML	✗	✗	✗	✗	✗	✗	✗	✗	✗
Rawson et al. (2021)	ML	✗	✗	✗	✗	✗	✗	✗	✗	✗
Yin and Li (2025)	ML	✗	✗	✗	✗	✗	✗	✗	✗	✗
Naderpour et al. (2014)	AI	✗	✓	✓	✓	✓	✓	✗	✓	✓
Bremnes et al. (2024)	AI	✓	✓	✓	✓	✓	✗	✗	✓	✓
Rothmund et al. (2025)	AI	✓	✓	✓	✓	✓	✗	✗	✓	✓

Risk awareness performance criteria: risk is defined and explained ($C_{1.1}$); hazard analysis is performed ($C_{1.2}$); causal and frequency analysis is performed ($C_{2.1}$); consequence analysis is performed ($C_{2.2}$); risk presentation is performed ($C_{2.3}$); risk evaluation is performed ($C_{2.4}$); risk reduction is performed ($C_{2.5}$); decision making incorporates risk metrics ($C_{3.1}$); decision making controls risk ($C_{3.2}$). ✗: addressed; ✓: not addressed.

Holistic & hybrid approaches. Considering the synthesized process of risk awareness, approaches that address risk awareness holistically should be employed. This includes from the proper definition of risk to decision making that controls and maintains risk to a desirable level. Current approaches lack transparency regarding the methodological connection from hazard identification to decision making. In addition, hybrid approaches that combine systematic risk assessment methods with data-driven methods should be developed, considering their respective strengths and limitations. This includes utilizing the advantages of risk assessment methods to systematically analyze risk, while complementing missing data, processing big data, and identifying hidden patterns by utilizing data-driven and learning approaches.

Risk awareness of AI agents. Novel approaches addressing the risk awareness of AI agents are needed. Specifically, approaches that integrate AI techniques in the risk assessment process are distinct from approaches that integrate risk assessment methods into the life-cycle of AI agents, similar to how physics-informed neural networks are developed to follow the laws of physics. In addition, a clear distinction should be given between approaches that utilize multiple narrow and domain-specific AI agents for risk awareness and approaches that develop a general-purpose AI agent with risk awareness. Both latter approaches are deemed pivotal as a way forward from addressing risk awareness in the AI domain to addressing the intrinsic risk awareness of AI agents.

4. Conclusion

The aim of this article is to identify and assess differences in how risk awareness is understood and used in the safety science and AI domains with engineering applications by conducting a systematic literature review. A generic risk awareness description in the safety domain is synthesized, and various risk awareness concepts in the AI domain are reviewed and assessed. The results in this article show that the risk awareness concept is more comprehensive in the safety science domain than in the AI domain, and this gap is outlined through a performance evaluation.

Recommendations have been given towards addressing the risk awareness gaps by defining risk to avoid misconceptions, employing risk assessment methods to systematically analyze risk, employing holistic and hybrid approaches to address the synthesized process of risk awareness, and developing novel approaches to address the intrinsic risk awareness of AI agents. The results can be used as a basis for improving safety in the AI domain, as they illustrate what needs to be implemented to achieve "truly" risk-aware and safe AI agents. A potential limitation of this review is the omission of further dimensions of risk awareness addressed in other subject areas, such as in medicine, social sciences, and psychology.

Acknowledgement

The research presented in this article has been supported by the Research Council of Norway through the SFI Autoship project (No. 309230).

Appendix A.

Table A.1. List of all AI-related terms.

"*reinforcement learning*" *OR* "*deep learning*" *OR* "*machine learning*" *OR* "*artificial intelligence*" *OR* "*neural network*" *OR* "*bayesian network*" *OR* "*federated learning*" *OR* "*supervised learning*" *OR* "*semisupervised learning*" *OR* "*large language model*" *OR* "*transformer*" *OR* "*long short-term memory*" *OR* "*generative adversarial network*" *OR* "*q-learning*" *OR* "*deep q-network*"

Table A.2. List of the top ten *source titles* in the safety science domain.

Risk Analysis[1]; *Safety Science*[1,2]; *Journal of Risk Research*[1]; *Accident Analysis and Prevention*[1]; *Journal of Disaster Research*[1]; *Applied Ergonomics*[1]; *Reliability Engineering and System Safety*[1]; *International Journal of Occupational Safety and Ergonomics*[1,2]; *International Journal of Disaster Risk Reduction*[2]; *Journal of Flood Risk Management*[1]

All science journal classification based on SCImago journal ranking 2024: *safety, risk, reliability and quality* (1); *safety research* (2).

References

Aven, T. (2011). *Misconceptions of risk*. John Wiley & Sons.

Bellet, T. and A. Banet (2012). Towards a conceptual model of motorcyclists' risk awareness: A comparative study of riding experience effect on hazard detection and situational criticality assessment. *Accident Analysis & Prevention 49*, 154–164.

Bremnes, J. E., I. B. Utne, T. R. Krogstad, and A. J. Sørensen (2024). Holistic risk modeling and path planning for marine robotics. *IEEE Journal of Oceanic Engineering*.

Castañeda, J. V., P. B. Repetto, N. C. Bronfman, and P. C. Cisternas (2024). Community preparedness for natural hazards in two chilean coastal cities. *Journal of Risk Research 27*(7), 806–826.

Chatzimichailidou, M. M. and I. M. Dokas (2016). Risksoap: Introducing and applying a methodology of risk self-awareness in road tunnel safety. *Accident Analysis & Prevention 90*, 118–127.

Cisternas, P. C., L. A. Cifuentes, N. C. Bronfman, and P. B. Repetto (2024). The influence of risk awareness and government trust on risk perception and preparedness for natural hazards. *Risk analysis 44*(2), 333–348.

Dokas, I. M. (2026). From hallucinations to hazards: benchmarking llms for hazard analysis in safety-critical systems. *Safety Science 194*.

Endsley, M. R. (2017). Toward a theory of situation awareness in dynamic systems. In *Situational awareness*, pp. 9–42. Routledge.

Floridi, L., J. Cowls, M. Beltrametti, R. Chatila, P. Chazerand, V. Dignum, C. Luetge, R. Madelin, U. Pagallo, F. Rossi, et al. (2018). Ai4people—an ethical framework for a good ai society: Opportunities, risks, principles, and recommendations. *Minds and machines 28*(4), 689–707.

Gibson, S. D. (2003). The case for 'risk awareness'. *Security Journal 16*(3), 55–64.

Hu, C., Y. Niu, H. Jiang, X. Zhang, and X. Cheng (2025). Decision-making for autonomous driving via a coupled reinforcement learning network combined with risk assessment. *IEEE Robotics and Automation Letters 10*(9), 8994–9001.

Jha, A., A. Bhatia, and K. Tiwari (2025). Bayesian deep learning meets self-attention: A risk-aware approach to advertisement optimization. *IEEE Access*.

Kaplan, S. and B. J. Garrick (1981). On the quantitative definition of risk. *Risk analysis 1*(1), 11–27.

Latvakoski, J., R. Öörni, T. Lusikka, and J. Keränen (2022). Evaluation of emerging technological opportunities for improving risk awareness and resilience of vulnerable people in disasters. *International Journal of Disaster Risk Reduction 80*, 103173.

Lederer, A., E. Noorani, J. S. Baras, and S. Hirche (2025). Risk-sensitive safety filters for reinforcement learning with probabilistic guarantees. *IEEE Transactions on Automatic Control*.

Lee, W.-H., J.-N. Lin, and K.-L. Wen (2025). Risky scooter riders pay more: a usage-based insurance study of a scooter risk-aware fingerprint model without privacy issues. *Journal of Intelligent Transportation Systems*, 1–16.

Liu, H., Y. Wang, W. Fan, X. Liu, Y. Li, S. Jain, Y. Liu, A. Jain, and J. Tang (2022). Trustworthy ai: A computational perspective. *ACM Transactions on Intelligent Systems and Technology 14*(1), 1–59.

Liu, J., J. Wang, and H. Lin (2025). Coordinated physics-informed multi-agent reinforcement learning for risk-aware supply chain optimization. *IEEE Access 13*, 190980–190993.

Lyu, Q., G. Fu, Y. Wang, D. Ban, Z. Wu, and X. Xie (2024). Insufficient safety capability among general aviation practitioners: Evidence from chinese accident data. *Safety Science 180*, 106655.

Naderpour, M., J. Lu, and G. Zhang (2014). A situation risk awareness approach for process systems safety. *Safety science 64*, 173–189.

Øien, K., S. Massaiu, and R. K. Tinmannsvik (2012). Guideline for implementing the rewi method; resilience based early warning indicators. *SINTEF report A 22026*.

Opletal, J. (2025). China's massive adas test: 36 cars, 15 hazard scenarios, 216 crashes.

Rawson, A., M. Brito, Z. Sabeur, and L. Tran-Thanh (2021). A machine learning approach for monitoring ship safety in extreme weather events. *Safety science 141*, 105336.

Rothmund, S. V., C. A. Thieme, I. B. Utne, and T. A. Johansen (2025). Supervisory risk control with application to industrial drone inspection. *Engineering Applications of Artificial Intelligence 157*, 110970.

Sabbatella, A., A. Ponti, A. Candelieri, and F. Archetti (2024). Bayesian optimization using simulation-based multiple information sources over combinatorial structures. *Machine Learning and Knowledge Extraction 6*(4), 2232–2247.

Shakya, A. K., G. Pillai, and S. Chakrabarty (2023). Reinforcement learning algorithms: A brief survey. *Expert Systems with Applications 231*, 120495.

Shree, V., B. Asfora, R. Zheng, S. Hong, J. Banfi, and M. Campbell (2021). Exploiting natural language for efficient risk-aware multi-robot sar planning. *IEEE Robotics and Automation Letters 6*(2), 3152–3159.

Su, H., N. Wang, and X. Wang (2025). Collision risk assessment of lane-changing vehicles based on spatio-temporal feature fusion trajectory prediction. *Electronics 14*(17), 3388.

Sun, S., C. Wang, B. Xiao, X. Liu, C. Shi, R. Sun, and R. Han (2024). Heterogeneous multi-agent risk-aware graph encoder with continuous parameterized decoder for autonomous driving trajectory prediction. *Electronics 14*(1), 105.

Sun, X., Q. Zhang, Y. Wei, and M. Liu (2023). Risk-aware deep reinforcement learning for robot crowd navigation. *Electronics 12*(23), 4744.

Ullah, Z., F. Al-Turjman, L. Mostarda, and R. Gagliardi (2020). Applications of artificial intelligence and machine learning in smart cities. *Computer communications 154*, 313–323.

Wijegunawardana, I., S. B. P. Samarakoon, M. V. J. Muthugala, and M. R. Elara (2025). Risk-aware complete coverage path planning using reinforcement learning. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*.

Wu, S., D. Tian, X. Duan, J. Zhou, D. Zhao, and D. Cao (2024). Continuous decision-making in lane changing and overtaking maneuvers for unmanned vehicles: A risk-aware reinforcement learning approach with task decomposition. *IEEE Transactions on Intelligent Vehicles 9*(4), 4657–4674.

Wu, Y., Y. Yang, W. Lv, J. Chang, Z. Li, D. Feng, T. Xu, and J. Li (2021). Robust unilateral alignment for subsurface lithofacies classification. *IEEE Transactions on Geoscience and Remote Sensing 60*, 1–13.

Xu, Z. and Q. Zhang (2026). Predictive maintenance optimization for industrial equipment via reliable prognosis and risk-aware reinforcement learning. *Complex & Intelligent Systems 12*(1), 11.

Yin, H.-D. and Y.-Y. Li (2025). Bayesian-optimized multi-task gaussian process regression with composite kernels for soybean oil futures forecasting. *IEEE Access*.

Zhang, Z., H. Li, T. Chen, N. Sze, W. Yang, Y. Zhang, and G. Ren (2025). Decision-making of autonomous vehicles in interactions with jaywalkers: A risk-aware deep reinforcement learning approach. *Accident Analysis & Prevention 210*, 107843.