energies

MDPI

Review

Ontology-Based Safety Management for Next-Generation Gaseous Fuel Infrastructure: A Review

Cheolhee Yoon [1], Junghoon Lee [1], Mimi Min [2,*] and Seungho Jung [2,*]

[1] Department of Environmental Engineering, Ajou University, Suwon 16499, Republic of Korea; cjfgml0509@ajou.ac.kr (C.Y.); joel0708@ajou.ac.kr (J.L.)
[2] Department of Environmental and Safety Engineering, Ajou University, Suwon 16499, Republic of Korea
* Correspondence: lovesme@ajou.ac.kr (M.M.); processsafety@ajou.ac.kr (S.J.)

Abstract

Next-generation gaseous fuel infrastructure for low- and zero-carbon gaseous fuels (hydrogen, ammonia, and hydrogen-blended natural gas) is planned to expand, but most announced hydrogen pipeline and storage projects have not reached a final investment decision. Hydrogen infrastructure safety research has centered on material integrity, risk assessment, and standards, while systematic syntheses of ontologies and knowledge graphs for this infrastructure remain relatively limited. This review classifies approaches to safety knowledge by their basis for risk judgment into traditional risk assessment, data-driven, and knowledge-structure-based approaches. Across these approaches, the basis for judgment tends to be fixed in advance, largely during design, training, or knowledge encoding. Computer science, by contrast, has a methodologically mature line of research reasoning only over the parts a query needs. Comparing the two lines reveals a direction that appears promising, in which the ontology remains the knowledge structure and only the necessary analyses run at query time. Prerequisite challenges include confirmation of response times, a shared safety vocabulary, scarce accident data, and verification methods. Close applications were found only outside gaseous fuel infrastructure. These observations rest on abstract-level screening and hold only within the search scope of this review, and the direction has not been demonstrated.

Keywords: chemical safety management; hydrogen safety; ammonia; gaseous fuel infrastructure; ontology; knowledge graph; query-time reasoning; chemical accident prevention

Check for updates

Academic Editor: Song Li

Received: 16 September 2026
Revised: 1 October 2026
Accepted: 10 October 2026
Published: 13 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Renewables made up 46% of the world's installed power capacity by the end of 2024, and the 585 GW of renewable capacity added that year represented 92.5% of all new power capacity [1]. Establishing renewables as the main source of electricity, however, still involves planning questions, including flexibility and adapting to the variability of renewable power [1]. Hydrogen is regarded as a strategic energy carrier for global decarbonization, above all in sectors in which emissions are hard to reduce, such as steelmaking, the chemical industry, long-haul transport, and seasonal storage. Its climate benefits, however, depend heavily on how it is produced, how its infrastructure is built, and how it is integrated at the system level [2]. Worldwide, demand for hydrogen stood at about 100 Mt in 2024. Production of low-emissions hydrogen expanded 10% in 2024, with 1 Mt expected in 2025. Even so, its share of total production stays below 1% [3].

The next-generation gaseous fuel infrastructure addressed in this review comprises the facilities for producing, storing, transporting, and using low- and zero-carbon gaseous fuels. The fuels covered are hydrogen, ammonia, and hydrogen-blended natural gas. Many components of this infrastructure are still at the announcement or review stage. Hydrogen pipeline announcements to 2035 total about 37,000 km, but less than 6% of them have reached a final investment decision (FID). For underground hydrogen storage, about 11 TWh (325 kt of H_2) of capacity would be available if every announced project went ahead, yet projects that have reached FID or started construction account for only 5% [3]. Underground hydrogen storage offers large capacity, geological containment, and low cycling costs [4]. Hydrogen storage in salt caverns is considered a key part of storing hydrogen underground at large scale [5]. Ammonia is being examined as an alternative means of hydrogen storage that would allow CO_2-free energy systems in the future [6], and routes that carry hydrogen blended into existing natural gas pipelines are also under study [7].

Planned hydrogen transmission networks differ considerably from existing hydrogen pipelines in scale and operating conditions. About 5000 km of hydrogen pipelines currently operate worldwide, chiefly linking refineries and chemical plants. They are generally privately owned and located onshore. Existing lines have small diameters and were designed to move hydrogen over relatively short distances, operating at low pressure with steady flow. Planned networks, by contrast, could fall under open-access regulation, use larger diameters and higher pressures, and run along both onshore and offshore routes. Planned networks are also expected to be more flexible in operation, which includes handling pressure swings that may be caused by changing loads and linepack storage [3].

Hydrogen safety research emphasizes a system-level perspective and real-time awareness of system state. It has been argued that hydrogen safety management should evolve from the analysis of individual components toward a system-level view that spans the value chain, from production through end use. A shift from static evaluation to dynamic safety frameworks built on prognostics and health management (PHM) and real-time monitoring can allow risk profiles to be updated continuously and predictive maintenance to be informed by risk [8]. A system dynamics analysis reports that an 80% decline in pipeline monitoring efficiency reduces the total hydrogen delivered to end users by 52%, and it stresses that real-time detection through hydrogen monitoring is especially important [9]. According to a review of risk mitigation measures for hydrogen infrastructure, measures such as leak detection, quick-release valves, emergency ventilation, and barriers of both material and physical kinds offer substantial protection potential, yet how well they work over the long term hinges on how quickly they respond in real time and on how rules are enforced [10].

Accident records have accumulated for hydrogen facilities and existing natural gas pipelines. A preliminary hazard analysis of the Hydrogen Incident and Accident Database (HIAD 2.1) identified 34 pipeline accidents and 28 storage tank accidents over roughly the past 50 years. Most pipeline accidents led to fires, while failures of storage tanks tended more often to escalate into explosions. For pipelines, corrosion and welding problems were the principal hazards [11]. For existing natural gas pipeline networks that could be used for hydrogen-blended transport, long-term statistics on scale and incident frequency are available. Each year, 19 gas transmission system operators in Europe gather incident data covering 150,000 km of pipelines, and a total of 1463 pipeline incidents were recorded from 1970 to 2022. Across the whole period, the failure frequency comes to 0.277 incidents per 1000 km per year, while the most recent 5-year moving average stands at 0.101 incidents per 1000 km per year [12].

Research on hydrogen infrastructure safety has developed mainly along three lines: material integrity, risk assessment methods, and standards and regulation [13]. Material integrity studies have clarified how hydrogen permeates natural gas pipelines carrying hydrogen blends and how it embrittles them [7]. For risk assessment methods, the need for comprehensive risk assessment that combines empirical modeling, computational fluid dynamics (CFD) simulation of hydrogen dispersion, and quantitative risk assessment has been raised [14]. For salt cavern hydrogen storage, a method has been proposed that establishes an evaluation index system, assigns weights, and then assesses leakage risk with a two-dimensional cloud model [5]. On standards and regulation, a review has synthesized hydrogen safety standards and the recent literature to provide a theoretical basis for further research [15]. Another study notes that hydrogen's flammability and explosiveness create safety concerns that hold back its deployment, and it points to the need for tailored regulations, codes, and standards (RCS) [16].

Data-driven approaches have been applied to existing oil and gas pipelines and to the process industries [17,18]. For hydrogen infrastructure, however, emerging technologies such as safety systems based on artificial intelligence (AI) and digital twins are judged to remain underexplored [10]. For existing oil and gas pipelines, it has been reported that artificial neural networks and support vector machines can predict failures more accurately than conventional methods, and combining several algorithms has been recommended to further improve accuracy and prediction time [19]. A review of AI applications in the process industries summarizes achievements in hazard identification, dynamic risk assessment, and early incident detection, while identifying sparse and biased incident data, limited interpretability, and the difficulty of validating algorithms under real plant conditions as persistent challenges [20]. Another assessment holds that comparatively little attention has gone to whether AI could itself create new safety weaknesses [21].

In the safety field, approaches that represent knowledge in explicit structures, such as ontologies and the knowledge graphs built on them, are increasing, but work that systematically synthesizes and reviews them from the perspective of safety in next-generation gaseous fuel infrastructure remains relatively limited. Ontology and knowledge graph research has accumulated attempts in various safety domains to gather knowledge scattered across different sources, such as accident reports, equipment information, regulations, and material properties, into a single structure and to address risk on top of that structure.

This review classifies and examines the approaches to handling safety knowledge from the perspective of safety in next-generation gaseous fuel infrastructure and explores the direction that emerges from their synthesis. The three-way classification is used as an organizing device and is not presented as a result. Its closest precedent in the reviewed literature is the classification of process safety and risk management methods by degree of quantification [22]. The contribution of this review lies in synthesizing the three approaches through three senses of judgment, in observing that safety studies on knowledge structures rarely report in their abstracts when their reasoning is executed, and in narrowing the remaining gap to the use of selective on-demand reasoning over explicit knowledge structures for safety judgment in next-generation gaseous fuel infrastructure. The review is limited to the safety of production, storage, transport, and use infrastructure for hydrogen, ammonia, and hydrogen-blended natural gas, and it does not cover power grids, battery energy storage, or nuclear power. Cases from adjacent domains are cited only for comparison, and each time they are cited, the text states that they come from an adjacent domain. Section 2 divides the approaches to handling safety knowledge into three groups by the basis for risk judgment and reviews, in turn, approaches that rely on expert and team knowledge, approaches that learn from observed data, and approaches that represent knowledge in explicit structures. Section 3 synthesizes this classification and sets out what

tends to become evident in comparison with approaches in adjacent fields. Section 4 discusses the areas where future research is expected, the open challenges, and the limitations of this review, and Section 5 concludes.

Review Method

This review is a structured narrative review rather than a systematic review. Literature was searched in ScienceDirect, Scopus, Web of Science, and Google Scholar, accessed through a university library, between August and September 2026. Search terms combined the safety of hydrogen, ammonia, and hydrogen-blended natural gas infrastructure with the use of ontologies and knowledge graphs, the limitations of expert-based and data-driven assessment, and query-time reasoning, including ontology-based data access.

The search centered on studies published from 2020 to 2026, and earlier studies were included where they define the methods discussed or remain key sources. Peer-reviewed articles, standards, and reports from authoritative institutions were included when their abstract or equivalent text could be obtained, and they supported a point in this review. Power grids, battery energy storage, and nuclear power were excluded, and cases from adjacent domains were cited only for comparison. Repeated searches, together with studies suggested by the authors and assessed against the same inclusion criteria, led to 117 included studies. During revision, three further studies were identified and assessed against the same criteria. Screening and the classifications in the tables rest on abstract-level statements. The numbers of records identified and screened at each stage were not recorded, so no Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram is presented.

In this review, "within the search scope of this review" refers to the studies collected with these databases, search terms, and period and included in this review, and the three studies added during revision. The evidence supporting each cell of Tables 1–5 is summarized in Supplementary Table S1.

2. Classification of Safety Research Approaches

Safety research approaches fall into three groups according to the basis for risk judgment (Figure 1). Traditional risk assessment relies on the knowledge of experts and teams, data-driven approaches learn from observed data, and knowledge-structure-based approaches represent the basis for judgment in explicit knowledge structures. This classification serves to organize the literature and to locate knowledge-structure-based approaches, and traditional risk assessment and data-driven approaches are reviewed as reference points for that comparison. This review distinguishes three senses of judgment. One sense is when the basis for judgment is fixed, for example at design time, at training time, or when knowledge is encoded in a structure. A second is when an inference is executed. A third is whether conclusions are computed in advance and stored before they are needed. This section reviews them in that order and treats knowledge-structure-based approaches in the most detail, because research on them has been growing quickly.

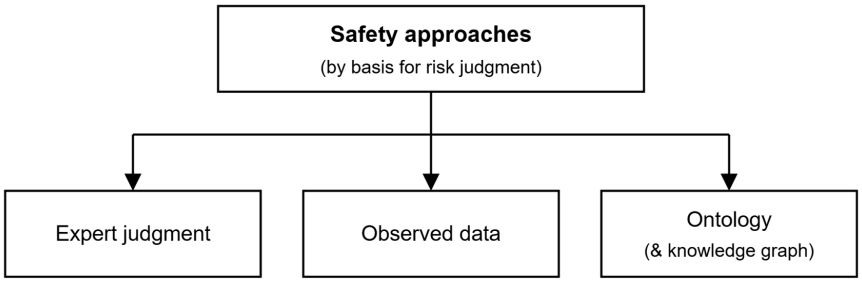


Figure 1. Classification of safety research approaches by the basis for risk judgment.

2.1. Traditional Risk Assessment

A review of process safety and risk management methods and models developed over past decades classifies them as qualitative, semi-quantitative, quantitative, or hybrid [22]. This classification reflects how far an assessment is quantified. It does not consider the form in which the underlying knowledge is represented.

Representative techniques include hazard and operability (HAZOP) analysis, quantitative risk assessment (QRA), and layer of protection analysis (LOPA). HAZOP, a process hazard analysis technique, examines how departures from design conditions affect a system and thus covers operability problems in addition to hazards [23,24]. QRA is used in design, licensing, and operation [25], and in LOPA, facility risk tolerance criteria are sometimes allocated to individual scenarios or events [26] (Table 1). These techniques have also been applied to facilities for low- and zero-carbon gaseous fuels. Quantitative risk assessment of hydrogen refueling stations has compared the calculated risk with as low as reasonably practicable criteria [27,28], and the same technique has been applied to two designs of ammonia-fueled ships [29] (Tables 1 and 2). Consequence analysis has been reported for a steam reforming process used to produce hydrogen [30] and for a large-scale hydrogen jet fire in a process plant [31] (Table 1). Layer of protection analysis has been carried out for an off-site hydrogen refueling station [32] (Table 1).

In expert-driven techniques such as HAZOP, the behavior of analysts and the way teams work are important determinants of result quality, yet empirical research on how these factors affect hazard analysis is almost nonexistent [33] (Table 1). Tools based on artificial intelligence have also been developed for HAZOP, but the existing tools do not offer both complete automation of report generation and a knowledge base comprehensive enough to detect hazards and operational malfunctions [34].

Studies have also addressed when assessments are made and updated, and they report a shift from static to dynamic assessment. The inability of QRA to update its risk picture has led to several dynamic risk assessment approaches, and the proposed way forward is a more precise QRA that can be dynamically updated to aid real-time decision-making [35]. A review of chemical process industry risk assessment methods conducted with the PRISMA methodology likewise identifies a shift toward dynamic and systemic assessments [36] (Table 1).

The validation of QRA and the scenarios it captures have also received attention. Compared with its wide application, QRA has received far less academic research on its validity and validation [25], and many major accidents follow scenarios that the analysis did not foresee [37] (Table 1).

Accident models grounded in systems theory also belong to the approaches that rely on expert and team knowledge. Most accident causation models were developed long ago, and it has been argued that traditional models may have limited ability to identify accident causes in modern industry, where socio-technical systems are complex. Systems-theoretic accident models were introduced in response, yet these models are used mainly by academic researchers [38]. A study of systems thinking approaches holds that these approaches may lower the possibility of surprises by deepening the understanding of the system and its risks, but that they need to be complemented by other approaches to adequately support decisions on risk issues [39] (Table 1).

The strength of approaches that rely on expert and team knowledge is their systematic character, as their procedures are standardized and tied into regulatory frameworks. Judgment in these approaches rests on the knowledge of experts and teams, and the reference point of an assessment is largely set at the design stage [40,41].

2.2. Data-Driven Approaches

Approaches that learn from observed data use machine learning (ML) and AI to obtain the basis for risk judgment from the data themselves.

In chemical process monitoring, a field close to the scope of this review, ML and data analytics have made great progress in recent decades, but models tested on benchmark cases like the Tennessee Eastman simulation still face substantial barriers when deployed in real industrial settings. As models grow more complex, they generalize poorly, become less transparent, and struggle to detect faults and diagnose their root causes at the same time. Current research is dominated by an algorithm-centric bias, marked by heavy reliance on simulated benchmarks, inadequate interpretability of models, and a mismatch with industry's need for lightweight, operable solutions [42]. A review of real-time fault detection and diagnosis in the Industry 4.0 context likewise reports difficulties with data quality, model interpretability, and the complexity of integration [43] (Table 1).

The nature of process operation and safety data sets the conditions under which data-driven models operate. A study reporting industrial experience with soft sensors in the chemical industry notes that changes in equipment, feedstock, sensors, and operating strategy keep most processes away from a true steady state, so soft sensor models have to be updated at regular intervals. The prevailing maintenance practice of rebuilding models from recent data, however, is not sufficiently robust when process data are corrupted by outliers, shutdowns, and non-steady-state transients. As an alternative, the study proposes an indicator based on prediction residuals that detects long-term declines in prediction performance [44]. Safety applications also contend with scarce data, and a benchmark framework for predicting rare abnormal events in chemical processes obtains its training data from forward flux sampling (FFS) simulations [45] (Table 1).

Reviews of applications and prospects point to challenges and new risks alongside reported gains. A review of safety-critical AI in the process industries finds that earlier studies have mostly stressed the safety benefits of ML and deep learning while paying less attention to the possibility that AI itself introduces new safety-critical vulnerabilities, and it sets out seven categories of AI-induced risk [21] (Table 1).

A methodological literature has also developed around how data-driven models behave, with particular attention to separating correlation from causation and to distribution shift. Regarding correlation and causation, one perspective argues that many deep learning failures may be read as different symptoms of one underlying problem, known as shortcut learning. In this view, a shortcut is a decision rule that scores well on standard benchmarks yet does not carry over to harder test conditions such as real-world settings [46]. When a model gains performance by exploiting spurious correlations and artifacts in a dataset instead of causal relationships or task-relevant features, the phenomenon is described as the Clever Hans effect [47] (Table 1).

For distribution shift, one synthesis observes that most practical applications must deal with dataset shift in some form [48]. A mismatch between training and test distributions can markedly reduce the accuracy of ML systems once they are deployed in real environments. With standard training, performance on out-of-distribution data falls well below in-distribution performance, and the gap persists even for models trained with existing methods designed to handle distribution shift [49]. ML systems that depend heavily on the properties of their inputs tend to fail silently, producing wrong outputs without signaling an error [50]. In the data streams analyzed by industrial AI, a change in distribution may leave the knowledge of models trained on past data out of step with the current state, and detecting such a change early is especially difficult when the corresponding ground truth is not immediately available [44,51] (Table 1).

Problems with the generalization of data-driven models, their use in real settings, and the reliability of reported performance have also been documented in fields outside the scope of this review. A power grid study argues that giving too little weight to domain knowledge poses a serious risk to the practicality of data-driven methods and that ignoring grid-specific spatiotemporal patterns can yield infeasible or meaningless predictions for new inputs [52]. In a medical example unrelated to safety engineering, models for predicting treatment outcomes in schizophrenia were highly accurate within the trial used to develop them, yet their out-of-sample performance was no better than chance [53]. Work on battery state-of-health estimation and on reproducibility in machine-learning-based science has likewise reported limited applicability when conditions change and overly optimistic conclusions caused by data leakage [54,55] (Table 1).

On explainability, part of the literature is critical of post hoc explanation. There is a hope that methods for explaining black box models after the fact might ease the problems these models cause. It has been argued, however, that explaining black boxes rather than building interpretable models from the outset risks entrenching bad practice and can cause serious harm, and that models should be designed to be interpretable by nature [56]. Even Shapley values, a widely used explanation technique, run into mathematical problems when used to measure feature importance, and remedies for these problems inevitably add complexity, for example by requiring causal reasoning [57] (Table 1).

In approaches that learn from observed data, judgment rests on the observed data used for training, and the trained models still require updating after deployment as data distributions and operating conditions change.

Table 1. Literature on traditional risk assessment and data-driven approaches.

Study	Year	Method and Target	Characteristics	Limitations	Statements on Assessment Timing	Study Type
Traditional risk assessment						
[22] Khan et al.	2015	Process safety and risk management methods and models	Classified as qualitative, semi-quantitative, quantitative, or hybrid	-	-	Review
[13] Min et al.	2025	Trends in methodologies for hydrogen risk assessment (theoretical, experimental, and simulation-based)	Focuses on risk in stationary and transportation applications, covering storage systems and transport infrastructure	-	-	Review
[23] Dunjó et al.	2010	Hazard and operability (HAZOP) analysis	Process hazard analysis technique covering both the hazards and the operability problems of a system	-	Explores how departures from design conditions affect the system	Review
[24] Jung et al.	2025	Combined application of HAZOP and job safety analysis for preventing fires and explosions in chemical plants	Describes HAZOP as the most common safety evaluation method, since it works through the causes and outcomes of every abnormal deviation in plant facilities	Additional risk factors were found with job safety analysis that the HAZOP process had not identified	-	Journal article
[25] Goerlandt et al.	2017	Quantitative risk assessment (QRA)	Widely applied across industries as a tool for improving safety	Far less academic study of its validity and validation than of its application	Used in design, licensing, and operation	Review

Table 1. *Cont.*

Study	Year	Method and Target	Characteristics	Limitations	Statements on Assessment Timing	Study Type
[35] Villa et al.	2016	QRA and dynamic risk assessment approaches	General approach unchanged since its emergence in the early 1980s, while continuing to evolve in various forms	Cannot update the risk picture	Proposed way forward is a more precise QRA that can be dynamically updated to aid real-time decision-making	Review
[26] Baybutt	2017	Layer of protection analysis (LOPA)	Facility risk tolerance criteria allocated to individual scenarios or events	Allocation is fraught with difficulty and can make results inaccurate within one facility and inconsistent between facilities	-	Journal article
[30] Lee et al.	2024	Quantitative risk assessment of steam reforming in hydrogen production plants (PHAST consequence analysis)	Reports impact distances for jet fire, vapor cloud explosion, and flash fire	-	Offered as baseline data for designing and operating such facilities	Journal article
[32] Kim et al.	2025	Off-site hydrogen refueling station examined through HAZOP followed by layer of protection analysis	Compares an approach counting only automated protection layers with one that also counts mechanical and procedural layers	-	Proposes a hybrid in which all safeguards are weighed at the early design stage and functional safety governs the final design stage	Journal article
[27] Jeon et al.	2023	Quantitative risk assessment of hydrogen refueling stations in urban areas (PHAST and HyRAM)	Societal risk evaluated with a probit model, and the calculated risk placed below as low as reasonably practicable criteria	-	-	Journal article
[28] Kwak et al.	2024	Quantitative risk assessment of an urban hydrogen refueling station (SAFETI)	Quantifies the effect of the shutoff valve and proposes minimum installation requirements	-	-	Journal article
[31] Jang and Jung	2016	Computational fluid dynamics analysis of a hydrogen jet fire at high pressure in a process plant	Evaluates jet fire consequences under high temperature and radiant heat, with results checked against another commercial code	-	-	Journal article
[40] Yoon et al.	2025	Inherent safety design of a liquefied natural gas terminal (transition context) through section isolation, with accident scenarios assessed by computational fluid dynamics	Finds that explosion impacts vary across sections and that spatial configuration and operating conditions matter	-	States that accident impacts should be calculated in advance so that safety measures enter at the design phase	Journal article
[41] You et al.	2019	Explosion risk analysis of dual mixed refrigerant liquefaction configurations for floating liquefied natural gas (transition context)	Similar efficiency between the two configurations, but explosion risk differs with the optimal refrigerant composition and the number of devices	-	Offered as guidance for managing risk while the process concept is still being chosen	Journal article

Table 1. *Cont.*

Study	Year	Method and Target	Characteristics	Limitations	Statements on Assessment Timing	Study Type
[33] Braarud and Simensen	2025	Expert-driven techniques such as HAZOP and STPA (System-Theoretic Process Analysis)	Analyst behavior and team processes shape the quality of results	Almost no empirical research on how these factors affect hazard analysis	-	Development of an observation protocol with initial results
[34] Elhosary and Moselhi	2024	HAZOP automation (AI-based HAZOP tools)	Structured, systematic method for identifying and mitigating hazards and operational issues	Current tools do not combine full automation of reporting with a comprehensive knowledge base	-	Review
[36] Ab Rahim et al.	2024	Chemical process industry risk assessment methods (process safety, security, and resilience)	12 major methods identified in the literature from 2000 to 2022	-	Shift toward dynamic, systemic assessments	Systematic literature review (PRISMA)
[37] Pasman et al.	2017	Quantitative risk assessment (QRA)	Used for prediction, for designing risk reduction measures, and as a basis for safety decisions. Large-scale computerization expected to become essential for keeping QRA practical	Distrusted partly because uncertainty estimates are usually left out. Many major accidents arise from scenarios the analysis did not foresee	-	Journal article
[38] Yousefi et al.	2019	Accident causation models AcciMap, FRAM (Functional Resonance Analysis Method), and STAMP (Systems-Theoretic Accident Model and Processes)	Systems-theoretic models introduced in response to the argument that traditional models may be limited in identifying accident causes in complex socio-technical environments	Used largely in academic research	-	Comparison study
[58] Patriarca et al.	2022	STAMP and the related techniques STPA and CAST (Causal Analysis based on System Theory)	Overview of system types, domains, and the coverage and completeness of analytical and methodological steps in 321 documents since 2003	-	-	Scoping review
[39] Bjerga et al.	2016	Systems thinking approaches (STAMP, FRAM)	May reduce the possibility of surprises by improving understanding of the system and its risks	Need to be complemented by other approaches to support decisions on risk issues adequately	-	Journal article
Data-driven approaches						
[19] Al-Sabaeei et al.	2023	ML for failure prediction in existing oil and gas pipelines (artificial neural networks, support vector machines)	More accurate failure prediction than conventional methods	Combining multiple algorithms recommended to improve accuracy and prediction time	-	Systematic review

Table 1. *Cont.*

Study	Year	Method and Target	Characteristics	Limitations	Statements on Assessment Timing	Study Type
[20] Selvam et al.	2026	AI for safety in the chemical process industries (early incident detection, hazard identification, dynamic risk assessment)	Reported gains in detecting anomalies and classifying near misses	Sparse and biased incident data, limited interpretability, algorithms hard to validate in real plant conditions	-	Review
[21] Devarajan et al.	2026	Safety-critical AI in the process industries	Applied to leak detection, hazard identification, predictive maintenance, and emergency shutdown support	Little attention to safety vulnerabilities that AI itself may introduce. Seven categories of AI-induced risk proposed	-	Review
[17] Jiao et al.	2023	Machine learning for predicting toxic dispersion casualties (quantitative property-consequence relationship model)	Builds a leak scenario database and estimates the affected range	The model needs expansion to more toxic chemicals and leak scenarios	-	Journal article
[18] Phark et al.	2022	Machine learning for emergency evacuation order decisions in chemical accidents (convolutional neural network)	Supervised learning on United States incident data, with weighted training applied because the learning data are asymmetric	-	-	Journal article
[42] Ji et al.	2025	ML and data analytics for chemical process monitoring	Tested on benchmark processes, including Tennessee Eastman	Barriers to real industrial deployment, heavy reliance on simulated benchmarks	-	Journal article
[43] Leite et al.	2025	ML for real-time fault detection and diagnosis in Industry 4.0	Fault prediction accuracy	Data quality, model interpretability, integration complexity	-	Review
[44] Lu and Chiang	2018	Soft sensor maintenance in the chemical industry	Most processes do not operate near a true steady state	Rebuilding models on recent data is not robust when data are contaminated by outliers, downtime, and transients	Periodic model updating needed. An indicator based on prediction residuals detects long-term performance degradation	Industrial experience report
[45] Sudarshan and Seider	2025	Prediction of rare abnormal events in chemical processes	Comparison of ML algorithms of differing complexity, with training data generated by FFS simulations	-	-	Benchmark framework
[59] Özdemir and Yıldırım	2025	ML across chemical engineering	Expanding beyond fault detection, process modeling, and signal processing into other fields. Transfer learning, automated ML, and physics-informed ML expected to gain attention	Discusses major challenges and limitations	-	Review

Table 1. *Cont.*

Study	Year	Method and Target	Characteristics	Limitations	Statements on Assessment Timing	Study Type
[46] Geirhos et al.	2020	Shortcut learning in deep learning	Treats many deep learning failures as symptoms of one underlying problem	Decision rules that do well on standard benchmarks but do not carry over to harder test conditions such as real-world settings	-	Perspective
[47] Pathak et al.	2026	Clever Hans effect (exploiting spurious correlations and artifacts in datasets)	Roadmap for robust AI includes causal integration	Current evaluation methods mask such behavior	-	Review
[48] Moreno-Torres et al.	2012	Dataset shift in classification	Most real-world applications face some form of shift	Relevant literature is scattered	-	Unifying framework based on reviewing and comparing key studies
[49] Koh et al.	2021	Benchmark of distribution shifts in real-world settings	Distribution shift can substantially reduce the accuracy of deployed ML systems	Out-of-distribution performance well below in-distribution performance, with the gap persisting under existing mitigation methods	-	Benchmark
[50] Rabanser et al.	2019	Methods for detecting dataset shift	-	ML systems that depend strongly on input properties tend to fail silently	-	Empirical study
[51] Lobo et al.	2023	Concept drift in industrial AI	Non-stationary data streams can shift the distribution and make model knowledge obsolete	Early detection of change is difficult when ground truth is not readily available	Early detection of change is key to updating model knowledge	Short conference paper
[52] Li et al.	2023	Data-driven approaches for power grids (out of scope)	Analysis of grid behavioral patterns using real operational data	Insufficient attention to domain knowledge poses a risk to practicality. Ignoring grid-specific patterns can yield infeasible or meaningless predictions	-	Conference paper
[54] Mbagaya et al.	2025	ML for battery state-of-health prediction (out of scope)	Cross-battery generalization tested on a cell not used in training	Conventional approaches have limited applicability when conditions differ. Transfer learning across battery chemistries identified as a future direction	-	Comparative review

Table 1. *Cont.*

Study	Year	Method and Target	Characteristics	Limitations	Statements on Assessment Timing	Study Type
[60] Chesterman et al.	2023	Wind turbine condition monitoring based on normal behavior modeling (out of scope)	Techniques implemented and compared on data from five operating wind farms	Choice of techniques driven by industrial partners' requirements, such as small training sets and inexpensive model training and upkeep	-	Review with implementation comparison
[53] Chekroud et al.	2024	Schizophrenia treatment outcome prediction models (medicine, out of scope)	High accuracy within the trial used for development	No better than chance out of sample. Pooling data from multiple trials did not improve predictions	-	Journal article
[55] Kapoor and Narayanan	2023	Data leakage in machine-learning-based science (out of scope)	Affects 294 papers in 17 fields, with overly optimistic conclusions in some cases	After the errors are fixed, complex ML models offer no substantive improvement over older logistic regression models	-	Literature survey
[56] Rudin	2019	Explaining black box models for high-stakes decisions	Expectation that post hoc explanation methods will mitigate the problems	Explaining black boxes may perpetuate bad practice and cause great harm. Inherently interpretable models should be designed instead	-	Journal article
[57] Kumar et al.	2020	Shapley value explanations as feature importance measures	-	Mathematical problems arise, and remedies inevitably add complexity, including a need for causal reasoning	-	Conference paper

Table 2. Status of the three types of next-generation gaseous fuel infrastructure.

Type	Scale and Operating Conditions	Main Hazards	Current Risk Assessment	Real-Time Judgment and Monitoring by	Current Position of Ontologies	Status Within the Search Scope	Supporting Studies	Citation Qualification (Classification by This Review)
Hydrogen pipelines and hydrogen-blended gas networks	About 37,000 km of hydrogen pipelines announced by 2035. Expected to be able to handle pressure swings that may be caused by changing loads and linepack storage	Overpressure and leakage during transport of blended gas, with long-term hydrogen embrittlement still under study. Conversion of natural gas networks brings some degree of safety risk. Most pipeline incidents led to fires, and corrosion and welding issues are the main hazards	-	Data-driven models (68 ms prediction with decisions every 1 s from a convolutional neural network (CNN) in a refueling station digital twin, and ML-based leakage monitoring of buried hydrogen-blended pipelines from vibration signals). Real-time detection through hydrogen monitoring is reported as particularly important. Pipeline digital twins have emerged for real-time monitoring, but a lack of standard architectures and interoperability frameworks limits their adoption	Causal networks extracted from accident reports and converted to Bayesian networks, and extracted knowledge stored and visualized, for existing natural gas and gas polyethylene pipelines (transition context)	Studies in which knowledge structures take part in judgment at the time of operation are still rare	[3,7,9,11,61–67]	The real-time judgment agents are CNN and ML models, not ontologies. The pipeline digital twin review concerns subsea, oil and gas, and water pipelines, not hydrogen pipelines. The ontology cases concern existing gas pipelines in a transition context

Table 2. *Cont.*

Type	Scale and Operating Conditions	Main Hazards	Current Risk Assessment	Real-Time Judgment and Monitoring by	Current Position of Ontologies	Status Within the Search Scope	Supporting Studies	Citation Qualification (Classification by This Review)
Large-scale hydrogen storage	About 11 TWh of capacity if all announced underground hydrogen storage projects go ahead, with 5% at final investment decision or under construction. Underground gas storage is subject to high-frequency cyclic loading	Potential leakage risk from the particular structure of salt caverns and the inherent risk of hydrogen storage. Fatigue failure from high-frequency cyclic loading. Rapid dispersion and cryogenic hazards from liquid hydrogen releases. Open research questions on cyclic geomechanics, interactions among hydrogen, rock, and microbes, and liner performance	Leakage risk assessment for salt cavern hydrogen storage based on an evaluation index system and weights. Static safety grade assigned to a horizontal salt cavern gas storage facility over the long term, based on numerical simulation, the analytic hierarchy process, and fuzzy comprehensive evaluation	Existing sensing technologies are deficient at low temperatures. A review of underground hydrogen storage offers insights into AI-based monitoring	Topological analysis of a hydrogen value chain knowledge graph built from an incident database	Real-time monitoring cases comparable to those for pipelines or refueling stations were not sufficiently identified within the search scope of this review	[3–5,68–71]	The static safety grade case and the fatigue failure statement concern gas storage and are not transferred to hydrogen storage (only the hydrogen part is cited)
Ammonia transfer and bunkering	Under the given inputs, altering wind speed by 50% can alter the maximum footprint of the 1100 ppm cloud by as much as 100% in small releases and by 663% in medium releases	Toxicity, flammability, and corrosiveness make safety harder to achieve than for other alternative fuels. In confined spaces, zones of fatal impact form within 1.675 s after leakage and ignition	Quantitative risk assessment applied to two ammonia-fueled ship designs, a containership and a tanker, identifying a risk reduction measure that current design requirements do not mandate	Recirculating flow delays sensor response, prolonging exposure	The closest case is a hydrogen accident knowledge graph whose stored knowledge completes a Bayesian network structural model	Ontology or knowledge graph studies that directly address ammonia were not found within the search scope of this review	[6,29,72–74]	The bunkering review and the fuel preparation room study have the same first author. The quantitative risk assessment covers two ship designs

Note. Cases marked as out of scope (power grids, batteries, wind power, medicine, and machine-learning-based science) are cited only as evidence for methodological points. A dash (-) means that the cited abstract does not report the item. The label (transition context) marks natural gas and liquefied natural gas facilities cited as existing gas infrastructure. The Study type column gives the document type stated in the abstract or the bibliographic record.

2.3. Knowledge-Structure-Based Approaches

The typical forms of knowledge-structure-based approaches are ontologies and the knowledge graphs built on them. An ontology can serve as a tool for storage and retrieval or as a tool for reasoning (Figure 2).

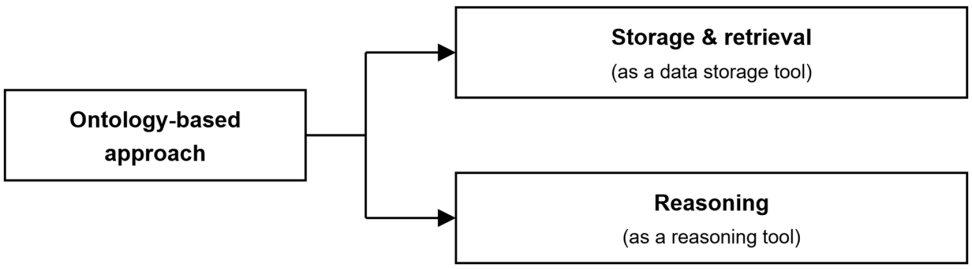


Figure 2. Two uses of ontologies.

This review groups the studies by how the ontology is used, separating cases in which it stays at the level of representation and retrieval from cases in which it is combined with a reasoning engine to judge risk directly. The former are discussed below as ontologies for storage and retrieval and the latter as ontologies for reasoning. Use rather than domain serves as the criterion because the same kind of use appears both inside and outside the domains covered by this review. Table 2 lists the target domain, use, reasoning agent, and reasoning timing of each study, while the text describes representative cases and shared characteristics.

2.3.1. Ontologies for Storage and Retrieval

Studies that use ontologies for storage and retrieval start from the problem of scattered information. In large, complex chemical processes, the information needed for process monitoring is spread across heterogeneous sources, and the OntoSafe ontology was proposed in response to provide semantics for abnormal situation management. OntoSafe makes it possible to integrate all the information required to judge the state of the process and records the underlying links so that changes in process descriptors spread consistently [75]. Work on hazardous chemical management and petrochemical alarm management has likewise used ontologies and knowledge graphs to bring together isolated or non-interoperable information [76,77] (Table 3).

Studies that structure accident experience with ontologies follow the same pattern of use. Accident databases are cumbersome because many people have written their entries over long periods in different styles, whereas enriching such databases semantically can make retrieving information more efficient. For accident databases in the chemical process industry, a method has been proposed that fills a predefined ontology structure automatically with information extracted by natural language processing. The resulting ontology-based database offers extra ways to explore accidents that human experts or computer systems can use. Combining several such databases means that knowledge can be reused across companies and countries [78]. Later studies structured hazardous chemical accident reports and HAZOP analysis reports as knowledge graphs [79,80] (Table 3).

The same use appears within the domains covered by this review, although the cases are unevenly distributed. A study of hydrogen accidents takes text fragments with clear causal relations from accident reports and applies a deep learning algorithm to extract knowledge. The extracted knowledge is stored in a Neo4j database as a knowledge map, and the hydrogen accident risk factors obtained from it complete the structural model of a Bayesian network [72]. A study of the entire hydrogen value chain holds that the heterogeneous structure of a knowledge graph enables broad information mapping by

identifying many kinds of entities and relationships, and that topological analysis can uncover hidden patterns. Its data come from a public record of hydrogen incidents between 2010 and 2023 [68]. The analysis in the hydrogen value chain study is a post hoc topological analysis of accumulated incident data rather than real-time reasoning during operation. Unlike the cases involving hydrogen and chemical processes, ontology or knowledge graph studies that directly address ammonia were not found within the search scope of this review.

The same pattern appears in existing natural gas pipelines that could be used to carry hydrogen blends. Studies of existing natural gas pipelines and gas polyethylene pipelines extracted knowledge from accident reports and pipeline operation data and organized it as knowledge graphs, and one of them converted the extracted causal network into a Bayesian network structure [61,62] (Table 3).

The most recent studies also use ontologies for storage and retrieval. In research on emergency response decisions for chemical process accidents, the entities and relations of a knowledge graph were used to strengthen the reasoning of large language models, and in safety management research on methanol production, an accident tree reasoning model designed separately from the safety information knowledge base identifies accident causes [81,82] (Table 3).

Studies on ontologies for storage and retrieval share a common character. Ontologies and knowledge graphs act as a representation layer that gathers scattered knowledge, gives it structure, and makes it searchable. The reasoning that judges risk is largely left to other engines outside the ontology. The reasoning engine differs from study to study and includes Bayesian networks, accident trees, large language models, statistical and topological analysis, and an external distributed intelligence system (see the reasoning agent column in Table 2). This division of roles between the representation layer and an external reasoning engine is not tied to particular substances or periods, and it also holds in the hydrogen domain covered by this review [68,72]. Even in the petrochemical alarm management study, which is closest to the time of operation, the output of the ontology ends up as modules for storing and retrieving information [77]. In studies from 2025 to 2026 that draw on deep learning and large language models as well, the ontology remains confined to representation and structuring [80,81].

2.3.2. Ontologies for Reasoning

In ontologies for reasoning, the ontology is combined with a reasoning engine to judge and diagnose risk directly, and this use is found both inside and outside the domains covered by this review.

Ontology-based dynamic reasoning has also been proposed within chemical processes. An ontology framework for process monitoring, which conceptualizes equipment, control systems, and hazards and formalizes them in description logic, seeks to show that description logic reasoning alone, without assistance from external agents, can support process monitoring as well as fault detection and diagnosis. Through forward chaining, its description logic reasoner inserts implicit facts into the ontology, working from current measurements toward a characterization of hazards. The process monitoring system can also check the consistency of its knowledge and explain the results formally [83]. Another study in process safety presents a knowledge-based framework that automatically generates HAZOP worksheets. It embeds expert knowledge of process and plant safety in an ontological model and then applies a reasoning algorithm built on semantic reasoners, which identifies hazards and operability problems in the manner of a HAZOP study [84]. The output of the latter study is a worksheet, and its abstract does not report when the reasoning is executed.

Adjacent fields outside the domains covered by this review have accumulated more cases of dynamic reasoning. Studies of smart work environments, metro tunnel construction, and construction safety inspection combine ontologies with logic-based rules, rule inference, or safety rule libraries in an effort to identify and manage hazards during operation or on construction sites [85–87] (Table 3). All three cases belong to adjacent domains outside the scope of this review and are not precedents for next-generation gaseous fuel infrastructure.

The four cases in cybersecurity and the Internet of Things, which lie outside the scope of this review, are not independent but successive work by the same research group, and they approach dynamic risk management through ontology-based reasoning [88] (Table 3). In a study that models risk in administrative domains in real time, all information is kept as knowledge in ontologies so that reasoning engines can be applied, and the new knowledge inferred in this way is used again in later reasoning [89].

Other safety fields outside the domains covered by this review that apply ontologies to safety management are covered by dedicated reviews of ontology use, including reasoning. A railway safety review identifies heterogeneous data, scalability, and missing semantic standards as remaining challenges, and a review in disaster response points to open issues in integrating and reasoning over semantic data in real time [90,91] (Table 3).

The rules used in studies of ontologies for reasoning vary and include the Semantic Web Rule Language (SWRL), rules in SPARQL (the query language for the Resource Description Framework, RDF), constraints defined with logic-based rules, and safety rule libraries, as listed in the rule approach column of Table 2. In one case, the abstract describes description logic reasoning alone, without mentioning rules [83]. In ontologies for reasoning, judgment rests on knowledge represented explicitly in the ontology, and reasoning is performed by an engine coupled to the ontology.

Table 3. Literature on knowledge-structure-based approaches and adjacent computer science research.

Study	Year	Target	Use or Line of Research (Classification by This Review)	Role of the Ontology	Reasoning Agent	Rule Approach	Reasoning Timing (Classification by This Review, Based on Abstracts)
Knowledge-structure-based approaches							
[75] Natarajan et al.	2012	Chemical process monitoring (case study of offshore oil and gas production)	Storage and retrieval	Provides semantics for abnormal situation management and integrates information for judging process state	Distributed intelligence system (ENCORE)	N/A	N/A (storage and retrieval)
[76] Zheng et al.	2021	Hazardous chemical management	Storage and retrieval	Ontology structure for hazardous chemical management	N/A	N/A	N/A (storage and retrieval)
[77] Song et al.	2025	Petrochemical alarm management	Storage and retrieval	Information fusion through local and global ontologies, with storage and retrieval modules	N/A	N/A	N/A (storage and retrieval)
[78] Single et al.	2020	Chemical process industry accident databases	Storage and retrieval	Accident database structure filled with extracted information	N/A	N/A	N/A (storage and retrieval)

Table 3. *Cont.*

Study	Year	Target	Use or Line of Research (Classification by This Review)	Role of the Ontology	Reasoning Agent	Rule Approach	Reasoning Timing (Classification by This Review, Based on Abstracts)
[92] Du et al.	2024	Management of hazardous chemical incidents	Storage and retrieval	Ontology construction techniques used to build the knowledge graph	N/A	N/A	N/A (storage and retrieval)
[79] Zhao et al.	2022	Petrochemical HAZOP analysis reports	Storage and retrieval	HAZOP ontology built with a top-down seven-step rule	N/A	N/A	N/A (storage and retrieval)
[80] Zhao et al.	2026	Hazardous chemical accident investigation reports	Storage and retrieval	Ontology representing key accident factors	Statistical analysis and exploration of topological relationships	N/A	Not reported in the abstract (statistical and topological analysis of accident investigation reports)
[72] Luan et al.	2024	Hydrogen accidents	Storage and retrieval	Knowledge map storing accident knowledge	Bayesian network	N/A	N/A (storage and retrieval)
[68] Karimi Dehkordi et al.	2026	Hydrogen value chain incidents	Storage and retrieval	Information mapping that identifies entities and relationships	Topological analysis	N/A	Not reported in the abstract (topological analysis of incident data)
[61] Bai et al.	2023	Natural gas pipelines (transition context)	Storage and retrieval	Extraction of causal networks from accident reports	Bayesian network (structure converted after quantifying correlations with the Decision-Making Trial and Evaluation Laboratory, DEMATEL)	N/A	N/A (storage and retrieval)
[62] Yu et al.	2025	Gas polyethylene pipelines (transition context), failure modes of material aging and third-party damage	Storage and retrieval	Storage and visualization of extracted knowledge	N/A	N/A	N/A (storage and retrieval)
[93] Tang et al.	2025	Commissioning new equipment in power grids (out of scope)	Storage and retrieval	Support for generating commissioning plans and identifying risks	N/A	N/A	N/A (storage and retrieval)
[81] Zheng et al.	2025	Emergency response to chemical process accidents	Storage and retrieval (retrieval-augmented variant)	Scenario and emergency response ontologies designed in seven steps	Large language model	N/A	N/A (storage and retrieval)
[82] Liu et al.	2023	Methanol production	Storage and retrieval	Safety information knowledge base (storage, processing, and reuse)	Accident tree reasoning model	N/A	N/A (storage and retrieval)

Table 3. *Cont.*

Study	Year	Target	Use or Line of Research (Classification by This Review)	Role of the Ontology	Reasoning Agent	Rule Approach	Reasoning Timing (Classification by This Review, Based on Abstracts)
[83] Musulin et al.	2013	Chemical process monitoring	Dynamic reasoning	Equipment, control, and hazard concepts formalized in description logic	Description logic reasoner	Not mentioned in the abstract (description logic reasoning)	Implicit facts added through forward chaining (timing not stated in the abstract)
[84] Single et al.	2020	Process and plant safety (HAZOP worksheets)	Dynamic reasoning	Embeds expert knowledge of process and plant safety	Algorithm that reasons through semantic reasoners	-	Not reported in the abstract (worksheets generated automatically)
[85] Teimourikia and Fugini	2017	Smart work environments (out of scope)	Dynamic reasoning	Ontological knowledge base of work environment safety	Run-time safety management methodology (monitor, analyze, plan, and execute loop)	Constraints defined with logic-based rules	Cannot be determined from the abstract
[87] Liu et al.	2025	Metro tunnel construction (out of scope)	Dynamic reasoning	Domain ontology formalizing safety risk knowledge	Rule inference	Rules (type not specified in the abstract)	No query-time configuration in the abstract
[86] Shen et al.	2022	Construction safety inspection (building information modeling, BIM) (out of scope)	Dynamic reasoning	Ontology-based safety rule library	Automatic inspection mechanism combining BIM and safety rules	Safety rule library	No query-time configuration in the abstract
[88] Riesco and Villagrá	2019	Cybersecurity (out of scope)	Dynamic reasoning	Basis of an integrated framework for dynamic risk assessment and management	Semantic reasoner	SWRL	No query-time configuration in the abstract
[89] Vega-Barbas et al.	2019	Cybersecurity and administrative domains (out of scope)	Dynamic reasoning	Stores information as knowledge so that reasoning engines can be applied	Reasoning engine	Rule language not specified in the abstract	Inferred knowledge reused in later reasoning (advanced execution stated in the abstract)
[94] Sánchez-Zas et al.	2023	Cybersecurity (out of scope)	Dynamic reasoning	Ontology describing types of anomalies	Knowledge inference and risk level calculation through SPARQL rules	SPARQL rules	No query-time configuration in the abstract
[95] Sánchez-Zas et al.	2024	Internet of Things cyber risk management (out of scope)	Dynamic reasoning	Basis of an interoperability framework for risk management	Logical reasoning	Rule language not specified in the abstract	Cannot be determined from the abstract

Table 3. *Cont.*

Study	Year	Target	Use or Line of Research (Classification by This Review)	Role of the Ontology	Reasoning Agent	Rule Approach	Reasoning Timing (Classification by This Review, Based on Abstracts)
[90] De Bartolomeo and De Nicola	2025	Railway safety (out of scope, review of 53 papers)	Dynamic reasoning (review)	Supports formalizing safety knowledge, identifying hazards, responding to emergencies, and modeling accident causation	N/A	N/A	N/A
[91] Soularidis et al.	2024	Search and rescue missions (out of scope)	Dynamic reasoning (review)	Review of technologies for integrating and reasoning over semantic data in real time	N/A	N/A	N/A
Adjacent computer science research (not safety domain)							
[96] Motik et al.	2019	Computer science (datalog)	Maintenance algorithms for materialization	-	Datalog implementations	Datalog programs	Precomputed and stored, updated whenever facts change
[97] Belhajjame and Mejri	2023	Computer science (enterprise knowledge graphs)	Incremental maintenance of materialization	-	-	-	Incremental maintenance of changes during construction (not sensor changes during operation)
[98] Nenov et al.	2015	Computer science (main-memory store)	Materialization (high-performance implementation)	-	Parallel datalog reasoning based on materialization	Datalog	Materialization computed and incrementally updated
[99] Poggi et al.	2008	Computer science (access to existing data sources)	Query time (ontology-based data access)	Ontology language based on description logic and suited to reasoning over many instances	Query answering that combines ontology-level reasoning with mapping processing	-	At query answering
[100] Xiao et al.	2018	Computer science (relational data sources, survey)	Query time (survey of ontology-based data access)	Semantic paradigm for accessing data repositories	N/A	-	-
[101] Calvanese et al.	2017	Computer science (querying relational data sources)	Query time (query rewriting)	Represents the domain concepts to which data sources are mapped	Query rewriting	-	Virtual approach that does not materialize triples
[102] Xiao et al.	2019	Computer science (data integration and access)	Query time (virtual knowledge graph)	Virtually maintained graph that embeds domain knowledge	-	-	-

Table 3. *Cont.*

Study	Year	Target	Use or Line of Research (Classification by This Review)	Role of the Ontology	Reasoning Agent	Rule Approach	Reasoning Timing (Classification by This Review, Based on Abstracts)
[103] Ahmetaj	2020	Computer science (query rewriting for ontology-mediated query answering)	Query time (query rewriting)	-	Query rewriting	-	-
[104] Botoeva et al.	2019	Computer science (non-relational data sources, MongoDB experiment)	Query time (virtual approach)	-	-	-	Virtual approach
[105] Calvanese et al.	2021	Computer science (biomedical resources, tutorial)	Query time (virtual knowledge graph)	High-level conceptual view of data	-	-	-
[106] Ranatunga et al.	2025	Computer science (environmental geospatial data integration)	Query time (ontology-based data access)	Extension of the GeoSPARQL ontology	SPARQL queries	-	-
[107] Urbani et al.	2011	Computer science (large datasets)	Query time (backward chaining)	-	Backward chaining reasoner	Web Ontology Language (OWL) Horst fragment	Computed at query time without pre-computation
[108] Qin et al.	2021	Computer science (ontology-mediated query answering)	Partial materialization	-	-	-	Partial consequences computed before online queries
[109] Sequeda et al.	2014	Computer science (data access systems)	Hybrid of query rewriting and materialization	-	Mappings implemented as Structured Query Language (SQL) views	-	Cost model decides which views to materialize
[110] Kharlamov et al.	2017	Computer science (data access in a petroleum company)	Query time (industrial deployment)	Ontology exposing data conceptually	-	-	-
[111] Dell'Aglio et al.	2017	Computer science (highly dynamic data)	Stream reasoning	-	N/A	-	Continuous queries (not the same as query-time configuration)
[112] Bonte et al.	2025	Computer science (Internet of Things and sensor data streams)	Stream reasoning (survey of stream processing)	-	N/A	-	Processing, querying, and reasoning inside a set time window
[113] Tataroğlu Özbulak et al.	2025	Computer science (heterogeneous streaming data)	Stream reasoning combined with knowledge graphs and ML	-	Combination of knowledge graphs and ML	-	Real-time reasoning (not query-time reasoning)
[114] Stavropoulou et al.	2024	Intelligent manufacturing processes (out of scope)	Query-driven digital twin	Semantic knowledge used by the digital twin	Digital twin	-	Queries orchestrate operations (not risk reasoning)

Table 3. *Cont.*

Study	Year	Target	Use or Line of Research (Classification by This Review)	Role of the Ontology	Reasoning Agent	Rule Approach	Reasoning Timing (Classification by This Review, Based on Abstracts)
Close cases in adjacent safety domains							
[115] Bilidas et al.	2023	Wildfire risk management (out of scope)	Query time (ontology-based data access)	Semantic query layer over a spatio-temporal data cube		-	-
[116] Noor	2022	Drug–drug interactions, medication safety (out of scope)	Dynamic reasoning	Knowledge graph integrating mechanistic information	Backward chaining inference algorithm	Complex rules (type not specified in the abstract)	-

Note: Reasoning timing is based on abstract-level statements in each study and does not presume the internal implementation of individual studies. The adjacent computer science research consists of methodological literature from computer science rather than applications in the safety domain, and its terms (materialization, query rewriting, backward chaining, stream reasoning) are explained in Sections 3.2 and 3.3. Columns marked (classification by this review) give this review's classification, not statements of the cited studies. A dash (-) means that the cited abstract does not report the item, and N/A (not applicable) means that the item does not apply, such as storage and retrieval uses without a reasoning engine, review articles, and surveys. The labels (out of scope) and (transition context) are used as in Table 1. In the Target column, the label Computer science marks methodological literature from computer science.

The preceding sections have classified safety research approaches by the basis for risk judgment. Traditional risk assessment rests on the knowledge of experts and teams, data-driven approaches on observed data, and knowledge-structure-based approaches on explicitly represented knowledge. Knowledge-structure-based approaches divide further into cases in which the ontology remains a representation layer and cases in which it is combined with a reasoning engine. Section 3 examines what emerges when these approaches are compared.

3. Synthesis

3.1. Comparison of Judgment Timing

The approaches reviewed in Section 2 can be compared using the three senses of judgment defined there. In the three approaches, the basis for judgment tends to be fixed before the point of need (Figure 3). For each approach, Table 4 gathers the basis for judgment, the point at which that basis is fixed, the point at which inference is executed, the update mode, and the reported limitations.

In traditional risk assessment, judgment rests on the knowledge of experts and teams, and its reference point lies largely at the design stage. Limitations have also been reported regarding validation and scenario coverage, decision support, and knowledge representation, and existing AI-based HAZOP tools do not offer both complete automation of report generation and a comprehensive knowledge base [25,34] (Tables 1 and 4).

In data-driven approaches, the basis for judgment is the data used for training. Most processes require models to be updated periodically, and performance gaps caused by distribution shift persist even in models trained with methods designed to address it [44,49] (Table 4). The future directions proposed within data-driven approaches include physics-

informed ML [59] and causal integration [47], and both point toward bringing in knowledge from outside the data.

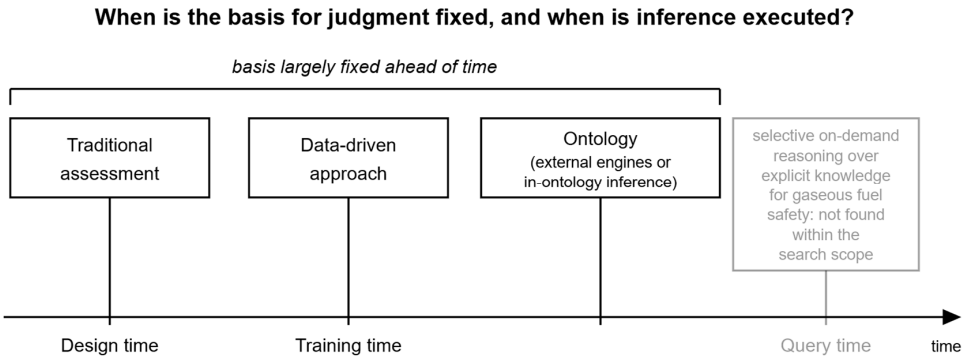


Figure 3. Timing of judgment in each approach: when the basis for judgment is fixed and when inference is executed (classification by this review). The query-time cell refers to next-generation gaseous fuel infrastructure, and close cases in adjacent domains are discussed in Section 3.3.

In knowledge-structure-based approaches, the basis for judgment is fixed when knowledge is encoded in the ontology, and inference is either handed over to engines outside the ontology or carried out within it. In the studies discussed in Section 2.3.1, the ontology serves as the representation layer, and judgment is performed by a separate engine coupled to it [72] (Table 3), whereas in Section 2.3.2 the ontology itself becomes the agent of judgment. Among these, one study in chemical processes reports that a description logic reasoner inserts implicit facts into the ontology through forward chaining [83], and one study in an adjacent domain reports that inferred knowledge is used again in later reasoning [89] (Table 4). The other studies do not report in their abstracts when their reasoning is executed.

Table 4. Timing of judgment in the three approaches and adjacent computer science research.

Category	Basis for Judgment (Classification by This Review)	Timing of Judgment (Classification by This Review)	Update Mode	Reported Limitations	Representative Studies	Qualification
Traditional risk assessment	Knowledge of experts and teams	Basis fixed largely at the design stage	Cannot update the risk picture (which prompted dynamic risk assessment approaches)	Academic research on validity and validation is scarce relative to application	[23,25,35] (see Table 1)	QRA is also used in operation
Data-driven approaches	Data used for training	Basis fixed by the data at training time	Periodic model updating needed in most processes	Barriers to real industrial deployment, and performance gaps from distribution shift that persist under mitigation methods	[42,44,49] (see Table 1)	Cases outside the domains of this review serve only as evidence for methodological points
Knowledge-structure-based approaches	Explicitly represented knowledge	Basis fixed when knowledge is encoded. Inference executed by external engines or within the ontology, with timing rarely reported in abstracts	Implicit facts added to the ontology through forward chaining in one study, and inferred knowledge reused in later reasoning in one adjacent-domain study	-	[72,83,89] (see Table 2)	Reasoning timing is rarely reported in abstracts and is not asserted for individual studies

Table 4. *Cont.*

Category	Basis for Judgment (Classification by This Review)	Timing of Judgment (Classification by This Review)	Update Mode	Reported Limitations	Representative Studies	Qualification
Adjacent computer science research (summary, computer science, not safety domain)	Ontology-level reasoning and mappings to source data	Inference executed in advance (materialization), at query time (backward chaining), or partly before online queries	Materialization is recomputed whenever the knowledge base changes, and backward chaining needs no precomputation	Query rewriting incurs a runtime cost	[99,107,108] (see adjacent computer science research in Table 2)	Update frequency is not the same as the time window of real-time safety judgment. Not applications in the safety domain

Note: Columns marked (classification by this review) give this review's classification using the three senses of judgment defined in Section 2, not statements of the cited studies. Their entries in the top three rows are synthesis statements from Section 3.1, and the individual studies are listed in Tables 1 and 3. A dash (-) means that the cited abstracts do not report the item. In the knowledge-structure-based row, the reported limitations cell is a dash, and the observation that the ontology stays a representation layer in storage and retrieval uses, with judgment performed by an external engine, is a synthesis statement from Section 3.1. The adjacent computer science research consists of methods from computer science, not applications in the safety domain. The entries in the adjacent computer science row summarize the cited computer science studies.

Approaches that rely on expert and team knowledge, approaches that learn from observed data, and approaches that represent knowledge in explicit structures share a common feature in that the basis for judgment is fixed before the state changes. In traditional assessment, the basis for judgment is largely fixed at the design stage. Data-driven approaches fix it through the data available at training time. In knowledge-structure-based approaches, the basis for judgment is fixed when knowledge is encoded, and the abstracts rarely report when inference is executed. Of the three senses defined in Section 2, the gap therefore concerns when inference is executed and whether conclusions are computed in advance, not when the basis for judgment is fixed. A case that applies selective on-demand reasoning over an explicit knowledge structure to safety judgment in next-generation gaseous fuel infrastructure did not appear within the search scope of this review, and close cases in adjacent domains are discussed in Section 3.3.

This observation, however, needs a qualification. Because the judgments of reasoning timing for knowledge-structure-based approaches rest on abstract-level statements in each study, they can say nothing definite about the internal implementation of individual studies. They can describe only what tends to hold for knowledge-structure-based approaches as a whole.

3.2. Reasoning Timing in Computer Science and the Semantic Web

In computer science, a body of research has accumulated around materialization. Materialization means computing in advance, and storing, the facts that can be obtained by reasoning. Datalog implementations often precompute and store every fact entailed by a program and its explicit facts, so that queries are answered directly over facts that already exist [96]. Materialized facts, however, have to be refreshed each time the explicit facts change, and recomputing every fact from the beginning can be highly inefficient, so many maintenance algorithms have been developed to cut the cost of updates [96]. Incremental maintenance of knowledge graphs is also treated as a research topic in its own right [97]. The changes addressed in that work, however, arise while a knowledge graph is being built, not in sensor data during operation. The RDF store RDFox supports main-memory parallel datalog reasoning and reports storing as many as 9.2 billion triples and a reasoning throughput of as much as 6.1 million triples each second [98]. RDFox

reached this performance on the premise of main-memory parallel reasoning, which shows that materialization works well when sufficient resources are available.

Computer science and the Semantic Web also offer methodological research on reasoning only over the parts needed at the time a query is posed, and these studies are not applications in the safety domain. The foundation of this query-time line of research is ontology-based data access (OBDA). OBDA started from the need of organizations for powerful yet efficient access to their existing data sources, and it introduced an ontology language based on description logic that is suited to reasoning over large numbers of instances. Its query answering combines ontology-level reasoning with mapping processing and efficient access to the source data [99]. A survey describes OBDA as a paradigm that has seen active development and study over the last ten years [100]. Ontop, an open-source OBDA system, presents as its key feature a virtual approach that avoids materializing triples by relying on query rewriting [101]. OBDA is also known as the virtual knowledge graph paradigm, in which graphs that remain virtual and carry domain knowledge take the place of the rigid relational tables that formed the integration layer [102]. Query rewriting is regarded as one of the most promising approaches to answering queries in OBDA [103].

The scope of OBDA has extended beyond relational databases, and the framework has been generalized so that arbitrary data sources can be queried with SPARQL [104]. The experiments in this generalization study, however, compared the document database MongoDB with a triple store. A practical tutorial on biomedical resources explains that an ontology lets users see the data at a high conceptual level and that the interplay between data and domain knowledge yields richer answers [105]. OBDA research continues, and in 2025 an OBDA-based framework was proposed for integrating heterogeneous environmental geospatial data [106].

The trade-offs between materialization and query-time reasoning are set out together in one study. Materialization allows very efficient answers at query time but comes at the price of an expensive closure computation performed up front, which has to be repeated whenever the knowledge base changes. Backward chaining needs no such precomputation and therefore suits knowledge bases that change more often, but it must do more computation at query time [107]. Backward chaining traces back, at the moment a query is posed, only the facts that the query requires. In this comparison, knowledge bases that change more often are those with a high update frequency, which is not the same as the time window of real-time safety judgment. A backward chaining reasoner has also been reported to scale to 1 billion triples [107].

Query-time approaches are not always advantageous. In ontology-mediated query answering, implicit information can be infinite and therefore cannot be stored in full, and query rewriting, the technique widely adopted to handle this, incurs the cost of rewriting queries at runtime [108]. As an alternative, partial materialization has been proposed, which computes only partial consequences before online queries instead of materializing everything [108]. A bidirectional evaluation system that combines query rewriting and materialization has also been presented, and it uses a cost model to decide which views to materialize [109]. Query rewriting and materialization are not mutually exclusive options.

An industrial application of OBDA is the data access project at the oil and gas company Statoil. The modules developed were integrated into a single OBDA system, deployed, and evaluated with engineers and real data [110]. Its purpose, however, was data access, not safety or risk reasoning. The authors of the Statoil study wrote that OBDA systems remain rare in real industrial settings, even though the research community has paid them much attention [110].

These reports from adjacent fields all describe general findings in computer science and are not results measured in next-generation gaseous fuel infrastructure.

3.3. Synthesis of the Two Lines of Research

When Section 3.1 is compared with Section 3.2, the use of knowledge structures in the safety field has moved from representation to reasoning, while the studies rarely report in their abstracts when that reasoning is executed. In computer science, by contrast, a methodological line of research that reasons only over the parts needed when a query is posed has matured, driven by scalability. No case that applies selective on-demand reasoning over an explicit knowledge structure to safety judgment in next-generation gaseous fuel infrastructure was found within the search scope of this review. Close cases appear in adjacent domains, each with a limitation. A wildfire risk management system computes fire risk from a harmonized data cube and uses geospatial ontology-based data access for semantic querying, so ontology-based access serves querying rather than the risk computation [115]. A study in medication safety, outside engineering safety, applies a rule-based backward chaining algorithm over a knowledge graph to identify evidence for drug–drug interactions [116]. None of these cases concerns next-generation gaseous fuel infrastructure.

The shared character of existing dynamic reasoning studies can be seen in cases of chemical process monitoring, HAZOP automation, and safety management for smart work environments. The chemical process monitoring study formalizes its ontology in description logic and, without assistance from external agents, works from current measurements to a characterization of hazards. Its reasoner uses forward chaining to insert implicit facts into the ontology, and the abstract does not report when this reasoning is executed [83]. The HAZOP automation study identifies hazards and operability problems with an ontology and a semantic reasoner, and its output is a worksheet [84]. The smart work environment study, which lies outside the domains covered by this review, claims to support safety management during operation, yet its reasoning structure is a loop made up of constraints defined with logic-based rules on the ontology together with monitoring, analysis, planning, execution, and knowledge [85]. Whether rules are used, and whether a study targets design time or runtime, differs from study to study, but among the studies examined in this section, a configuration that reasons only over the parts needed at the time a query is posed is not evident from the abstracts.

Comparing the use of knowledge structures in the safety field with query-time reasoning in computer science, one direction emerges. It is a configuration that keeps the ontology as the structure of knowledge but does not produce risk analyses in advance, running only the necessary parts at query time, when judgment becomes necessary. It is not a new technique but a direction that becomes visible when the two bodies of literature are compared.

Some research in computer science lies close to query-time configurations, but none of the studies reviewed in this section matches a configuration that reasons about risk at query time in next-generation gaseous fuel infrastructure. Stream reasoning, which applies inference techniques to highly dynamic data, is an independent research area that marked ten years in 2017 [111]. In 2025, a survey in the same line covered languages and systems for RDF stream processing and raised the requirement that data streams arriving continuously from Internet of Things devices and sensors be processed, queried, and reasoned over within a given time window [112]. The paradigm of stream reasoning, however, is continuous querying, which is not the same as a configuration triggered when a query is posed. A survey of challenges under streaming conditions concludes that combining knowledge graphs and ML greatly enhances the ability to handle those challenges [113], but what this conclusion points to is the combination of two methods, not query-time reasoning. As a case closer in form, in a manufacturing process study that combines knowledge graphs and digital twins, the digital twin runs suitable queries against

the knowledge graph to coordinate how virtual objects and their physical counterparts operate and to set their parameters [114]. In that study, however, what the queries drive is the coordination of operations rather than risk reasoning, and the target is a manufacturing process, not safety.

The conditions under which a query-time direction appears promising are limited. Query-time reasoning avoids paying the cost of updating precomputed results but pays a computational cost at query time instead [107,108], and a system that combines query rewriting with materialization and uses a cost model to decide which views to materialize has also been presented [109]. Running analysis at query time is therefore not always advantageous. Under conditions of large scale and frequent state changes, where precomputed results quickly become outdated, giving query-time execution a greater share may be reasonable. Next-generation gaseous fuel infrastructure is close to such conditions. Hydrogen pipeline announcements to 2035 total about 37,000 km, and most of these projects have yet to reach a final investment decision. Planned transmission networks are expected to be able to handle pressure swings that may be caused by changing loads and by linepack storage [3]. A diagnosis that shifting from static evaluation toward dynamic safety frameworks means risk profiles can be updated continuously has also already been put forward [8]. The means proposed in that review, however, are quantitative risk assessment, prognostics and health management, digital twins, and sensor monitoring, not ontologies.

The observations and direction obtained from the synthesis come with three qualifications. First, statements about when existing studies execute inference rest on abstract-level statements, and most abstracts do not report it, so the reasoning timing of individual studies is not asserted. Second, the statement that query-time approaches are still rare in the safety field is limited to the search scope of this review and does not mean that such studies do not exist. Third, the direction above reflects what tends to emerge when the literature is synthesized; it has not been demonstrated in this review, and because this review did not measure performance, it claims no superiority of any kind.

Section 4 examines how the direction that emerged from the synthesis could apply to each type of next-generation gaseous fuel infrastructure.

4. Future Research Directions and Challenges

To examine what research the synthesis in Section 3 could lead to in next-generation gaseous fuel infrastructure, this section addresses in turn the areas where research on ontology use is expected, the challenges that need to be resolved for such research to advance, and the limitations of this review itself. Throughout this section, statements with citations report applications and findings in the cited studies, whereas statements about configurations that run analysis at query time describe the direction set out by this review, which has not been demonstrated.

4.1. Areas Where Ontology Applications Are Expected

The areas where research on ontology use is expected are three types of next-generation gaseous fuel infrastructure, namely hydrogen pipelines and hydrogen-blended gas networks, large-scale hydrogen storage, and ammonia transfer and bunkering. The three types share three conditions, namely scale and dispersion, changing operating states, and heterogeneous data. On scale and dispersion, hydrogen pipeline announcements to 2035 total about 37,000 km, and if every announced underground hydrogen storage project goes ahead, about 11 TWh of capacity will be available [3]. The second condition, changing operating states, takes a different form in each type. Planned hydrogen transmission networks are expected to be able to handle pressure swings that may be caused by changing loads and by linepack storage [3], while fatigue failure from high-frequency cyclic loading

is identified as a critical risk in underground gas storage [69]. In ammonia bunkering, the size of the risk is sensitive to external conditions, and under the given inputs, altering wind speed by 50% can alter the maximum footprint of the 1100 ppm cloud by as much as 100% in small releases and by 663% in medium releases [73]. On the third condition, heterogeneous data, regulations, codes, and standards for hydrogen distribution and use are at a stage where tailored ones are recommended [16], and in the adjacent settings of hazardous chemical management and petrochemical alarm management, safety knowledge has been reported to be split across isolated databases and non-interoperable information [76,77].

Scale and dispersion, changing operating states, and heterogeneous data connect with the points identified in Section 3. If the size of a risk changes greatly with the state, an assessment produced in advance for the state at one moment loses its basis as soon as the state changes. Precomputation incurs update costs whenever changes occur [96,107]. This correspondence, however, is a structural analogy, not a measurement of the same phenomenon. The changes discussed in the literature on materialization and backward chaining are changes in knowledge bases, whereas the changes in the conditions above are changes in operating and weather conditions.

The status of each type is summarized in Table 4. For hydrogen pipelines and hydrogen-blended gas networks, specific safety concerns such as overpressure, leakage, and long-term hydrogen embrittlement have been reported [63,64] (Table 2), and the real-time judgment reported for this type and for refueling facilities is handled by data-driven models. In the digital-twin-based safety model for a hydrogen refueling station and in the leakage monitoring model for buried hydrogen-blended natural gas pipelines, it is neural networks and ML models that judge risk, not ontologies [65,66]. Ontologies have been used to structure accident knowledge for existing gas pipelines in a transition context [61,62] (Table 2).

For large-scale hydrogen storage, research gaps and leakage risks have been reported, but what is responsible for real-time monitoring is less clear than for pipelines or refueling stations. Current assessment for the storage type relies on precomputation, and a study assessing how safe a horizontal salt cavern gas storage facility remains over the long term produces a static safety grade, although its subject is gas storage rather than hydrogen storage [70] (Table 2). A review of underground hydrogen storage also offers insights into AI-based monitoring [4], but real-time monitoring cases comparable to those for pipelines or refueling stations were not sufficiently identified within the search scope of this review.

Ammonia transfer and bunkering combine difficult safety assurance and fast accident development with current assessments that sit at the design stage. Its toxicity, flammability, and corrosiveness make ammonia harder to handle safely than other alternative fuels [6], and in confined spaces such as a ship's fuel preparation room, zones of fatal impact form within 1.675 s after leakage and ignition [74]. The fuel preparation room study and the bunkering review mentioned above have the same first author. A quantitative risk assessment of ammonia-fueled ships was applied to two designs, a containership and a tanker, and identified a risk reduction measure not mandated in current design requirements [29] (Table 2). Ontology or knowledge graph studies that directly address ammonia were not found within the search scope of this review. The closest case is the hydrogen accident knowledge graph whose stored knowledge completes the structural model of a Bayesian network [72].

The same pattern recurs in hydrogen pipelines and hydrogen-blended gas networks, large-scale hydrogen storage, and ammonia transfer and bunkering. Current risk assessments are produced in advance [29,70], and where real-time judgment is required, data-driven models fill the role [65,66]. Ontologies are used in post hoc analysis of accident records, and even in alarm management, which is closest to the time of operation, the

output of the ontology ends up as storage and retrieval modules [77]. Studies in which knowledge structures take part in judgment at the time of operation are still rare in all three types. Future research is expected to address this gap. One form such research could take is a configuration in which an ontology maintains pipeline sections, material properties, regulations, and sensor histories as its structure and runs risk analysis as queries at the moment the state changes. Such a configuration, however, is a prospect of what may be possible and does not imply that it performs better than current data-driven monitoring.

4.2. Open Challenges

Before research that uses an ontology as its structure and runs risk analysis at query time can advance, challenges need to be resolved at the technical, domain, and methodological levels. Table 5 summarizes these challenges and the current status reported in the literature for each level. Viewed from the standpoint of operating real infrastructure, these challenges come down to four questions. One question concerns data acquisition, because recirculating flow in confined spaces can delay sensor response and existing sensing technologies are deficient at low temperatures [71,74]. Another question concerns keeping knowledge and models current, as soft sensor models need periodic updating [44], early detection of drift is important for updating model knowledge [51], and materialized facts need updating each time the explicit facts change [96]. Computational efficiency raises a further question, since a refueling station digital twin reports millisecond-scale predictions with decisions every second [65], whereas query-time reasoning performs more computation when a query is posed [107,108]. The last question is integration with digital twins and AI-based monitoring. The literature presents these as means for dynamic safety frameworks [8], yet they remain underexplored for hydrogen infrastructure [10], and the lack of standard architectures and interoperability frameworks limits their adoption [67]. These are prerequisite questions for the direction set out in Section 3.3, and this review does not answer them.

At the technical level, real-time query answering performance, the scale of the mapping layer, and partial reasoning with modularization remain open challenges. Query rewriting incurs a cost at runtime [108], while the time window available for safety judgments in next-generation gaseous fuel infrastructure is narrow (Table 5). Whether query-time reasoning fits within this time window has not yet been confirmed, and the literature in the related research areas comes entirely from computer science and adjacent domains, so none of it reports results measured in next-generation gaseous fuel infrastructure. The cost of building a mapping layer for large facilities and maintaining it during operation was not identified within the search scope of this review. Reasoning over only the necessary part presupposes that this part can first be identified, yet finding a minimal-size module of an ontology is known to be a computationally hard problem [117]. Query-time approaches therefore do not become lightweight automatically.

Knowledge-structure-based approaches carry limitations of their own, and the reviewed literature reports them mostly outside the safety field. Among the safety-domain sources reviewed, a railway safety review is the only one that states such limitations directly, naming data heterogeneity, scalability, and the absence of semantic standardization as remaining challenges [90]. A commentary on foundation models in chemical process safety also lists ontology drift among the barriers to deploying those models in safety-critical environments [118]. In computer science, materialized facts need updating each time the explicit facts change, and recomputing all of them anew can be very inefficient [96]. Backward chaining avoids this upfront cost but shifts more computation to query time [107], and query rewriting pays its cost at runtime [108]. Knowledge graphs under construction also undergo changes that need to be handled interactively and efficiently [97]. Data-driven

approaches and digital twins face difficulties of their own, such as barriers to real-world industrial deployment [42] and challenges in data integration, model validation, and scalability [67], so these limitations do not place any of the approaches above the others.

At the domain level, the standardization of safety vocabularies, constraints on data acquisition, and the scarcity of accident data remain challenges. Standards for representing sensors and their observations already exist, but they only represent observations and do not reason about risk [119], and what is missing is a shared vocabulary for the safety of next-generation gaseous fuels (Table 5). States that go unobserved or arrive late, such as delayed sensor response in confined spaces and deficient sensing at low temperatures in liquid hydrogen environments [71,74], stay the same whether reasoning happens earlier or later. That accidents are rare also means that there are few training samples (Table 5), and introducing knowledge structures does not change this scarcity.

At the methodological level, the challenges are verification and trust. In a configuration that does not produce results in advance, what is to be verified must first be redefined, but this review did not find literature that answers this question. Explicit knowledge structures can present reasoning paths formally and are therefore aligned with demands for explainability and auditability, yet a method for reconstructing reasoning performed at query time after the fact and presenting it to regulators has not been established within the search scope of this review. Although a survey of challenges under streaming conditions concludes that combining knowledge graphs and ML significantly enhances the capacity to deal with them [113], having such combinations available is not the same as having ways to verify the combined systems. A survey of the verification and evaluation of neurosymbolic AI does not address the verification of query-time reasoning [120] (Table 5).

Table 5. Prerequisite challenges for research on query-time execution.

Layer	Challenge	Current Status Reported in the Literature	Status Within the Search Scope of this Review	Supporting Studies	Field (Classification by This Review)	Citation Qualification (Classification by This Review)
Technical	Real-time query answering performance	Query rewriting incurs a runtime cost. The hydrogen refueling station case reports 68 ms prediction with decisions every 1 s, and the ammonia accident in a confined space develops within 1.675 s. Stream reasoning research raises open challenges and a need for processing, querying, and reasoning inside a set time window. Disaster response research reviews unresolved issues in real-time semantic integration and reasoning	Whether query-time reasoning fits within this time window has not yet been confirmed	[65,74,91,108,111,112]	Computer science (query rewriting, stream reasoning), in scope (hydrogen refueling station, ammonia), adjacent (disaster response)	Computer science and adjacent literature do not report results measured in next-generation gaseous fuel infrastructure. The judgment agent in the refueling station case is a CNN, not an ontology
Technical	Scale of the mapping layer	Query answering in ontology-based data access combines ontology-level reasoning with processing of mappings. The deployment at an oil and gas company went through module integration, integration with company infrastructure, and evaluation with engineers and real data	The cost of building a mapping layer for large facilities and maintaining it during operation was not identified within the search scope of this review	[99,110]	Computer science, data access in the oil and gas industry	The industrial deployment serves data access, not safety or risk reasoning

Table 5. *Cont.*

Layer	Challenge	Current Status Reported in the Literature	Status Within the Search Scope of this Review	Supporting Studies	Field (Classification by This Review)	Citation Qualification (Classification by This Review)
Technical	Partial reasoning and modular-ization	Finding minimal-size modules is a well-established hard computational problem, often infeasible for highly expressive languages, and practical approximations can yield unnecessarily large modules. Materialization maintenance algorithms have been neither compared with one another nor examined for practical applicability	Query-time approaches do not become lightweight automatically	[96,117]	Computer science	The difficulty of module extraction is cited as evidence against the direction emerging from the synthesis. The maintenance algorithm literature is not a critique of materialization
Domain	Standardization of safety domain vocabularies	Standards for representing sensors and observations exist. For hydrogen distribution and use, tailored regulations, codes, and standards are recommended. A review of integrity management and standards for hydrogen-blended pipelines points out shortcomings in current transport. Missing standard architectures and interoperability frameworks restrict the adoption of pipeline digital twins. A railway safety review also names the absence of semantic standardization as a challenge	What is missing is a shared vocabulary for the safety of next-generation gaseous fuels	[16,64,67,90,119]	Domain-independent standard, in scope (hydrogen distribution and use, hydrogen blending), adjacent (pipeline engineering, railways)	The sensor standard represents observations but does not reason about risk. The specific shortcomings are not specified in the abstract. The pipeline digital twin review does not concern hydrogen pipelines
Domain	Constraints on data acquisition	Existing sensing technologies are deficient at low temperatures in liquid hydrogen environments. Recirculating flow delays sensor response in confined spaces	States that go unobserved or arrive late stay the same whether reasoning happens earlier or later	[71,74]	In scope (liquid hydrogen, ammonia)	N/A
Domain	Scarcity of accident data	The moving average failure frequency of European gas transmission pipelines over the most recent 5 years is 0.101 incidents per 1000 km per year. Training data for a benchmark on predicting rare process events come from simulated sampling. A review of AI in the process industries cites sparse and biased incident data as a challenge. Research on real-time fault diagnosis calls for refining techniques for handling imbalanced data	Introducing knowledge structures does not change the scarcity condition	[12,20,43,45]	Natural gas transmission statistics (transition context), chemical and process industries	The failure frequency is a natural gas transmission statistic, not a failure frequency for hydrogen pipelines
Method-ological	Verification	QRA validity and validation have received far less academic study than applications of the method	No literature was found that answers how to redefine what is verified in a configuration that does not produce results in advance	[25]	Safety-related quantitative risk analysis	N/A
Method-ological	Explainability and auditability	Argument that inherently interpretable models should be designed for high-stakes decisions. Remedies in explanation techniques add further complexity, including a need for causal reasoning. Seven categories of AI-induced risk proposed	A method for reconstructing query-time reasoning after the fact and presenting it to regulators has not been established within the search scope of this review	[21,56,57]	General ML, AI in the process industries	N/A

Table 5. *Cont.*

Layer	Challenge	Current Status Reported in the Literature	Status Within the Search Scope of this Review	Supporting Studies	Field (Classification by This Review)	Citation Qualification (Classification by This Review)
Method-ological	Verification of combined knowledge structures and ML	Conclusion that combining knowledge graphs and ML significantly enhances the ability to handle stream challenges. Verification techniques for neurosymbolic AI are partly sufficient for handling symbolic and subsymbolic parts separately, but some applications need adjustments or new approaches	Having such combinations available is not the same as having methods to verify the combined systems	[113,120]	Computer science	The neurosymbolic verification survey does not address the verification of query-time (on-demand) reasoning. The combination conclusion does not concern query-time reasoning

Note: The column Status within the search scope of this review does not quote the cited studies. It gives the statements of absence in Section 4.2 and this review's own judgments, and the computer science and adjacent domain literature does not report results measured in next-generation gaseous fuel infrastructure. The columns Field and Citation qualification (classification by this review) give this review's classification, not statements of the cited studies. N/A in the column Citation qualification means that no qualification is needed. In the column Field, in scope, adjacent, and (transition context) follow the scope defined in Section 1. This review frames the challenges for the direction set out in Section 3.3, a direction that has not been demonstrated.

4.3. Limitations of This Review

This review has four limitations.

First, the judgments of reasoning timing for knowledge-structure-based approaches rest on abstract-level statements in each study, and the full texts of individual papers, particularly their implementation sections, were not examined. The reasoning timing of individual studies was therefore not asserted.

Second, the observation that query-time approaches are still rare in the safety field is limited to the search scope of this review and does not mean that such studies do not exist. The search depended on particular databases, languages, and keywords. Some studies were not included because the relevant text could not be obtained, and the numbers of records at each stage were not recorded (Section Review Method).

Third, the direction outlined in Section 3 is conceptual and has not been demonstrated in this review. Because this review neither implemented any system nor measured performance, it does not claim that this direction is faster or more accurate than current approaches.

Fourth, the literature on query-time approaches comes from computer science, and there is a gap in transferring it to the safety domain. Even the case closest to industry was a deployment for data access, not for safety or risk reasoning [110]. The methodological maturity of the query-time reasoning line of research in computer science does not guarantee that it can be transferred to the safety domain.

5. Conclusions

Next-generation gaseous fuel infrastructure is planned to expand as large, dispersed facilities whose operating states change, and demands on its safety management are growing. A diagnosis that shifting from static evaluation toward dynamic safety frameworks means risk profiles can be updated continuously has already been put forward in the field [8]. To explore future directions from the perspective of safety in next-generation gaseous fuel infrastructure, this review has therefore examined how safety knowledge is handled, grouping the approaches by the basis for risk judgment. Traditional risk assessment rests

on the knowledge of experts and teams, data-driven approaches on observed data, and knowledge-structure-based approaches on explicitly represented knowledge.

Two findings emerged from synthesizing the approaches examined. First, they share a common feature in that the basis for judgment is fixed before the state changes. In traditional assessment, the basis for judgment is largely fixed at the design stage. Data-driven approaches fix it through the data available at training time. Knowledge-structure-based approaches fix it when knowledge is encoded, and the abstracts rarely report when inference is executed. One study in chemical processes reasons from current measurements with description logic alone, without external agents, and forward chaining then adds the inferred facts to the ontology [83]. Second, computer science and the Semantic Web have a methodologically mature line of research on reasoning only over the parts needed when a query is posed, although industrial use remains limited, as reported for ontology-based data access [110]. An approach that rewrites queries instead of materializing triples has been implemented as a working system [101], and backward chaining, which needs no precomputation, has been described as suited to knowledge bases that change more often [107]. No case that applies selective on-demand reasoning over an explicit knowledge structure to safety judgment in next-generation gaseous fuel infrastructure was found within the search scope of this review, although close cases with their own limitations exist in adjacent domains [115,116].

Combining the two observations reveals a direction that appears promising. It is a configuration that keeps the ontology as the structure of knowledge but does not produce risk analyses in advance, running only the necessary parts at query time, when judgment becomes necessary. Next-generation gaseous fuel infrastructure has conditions under which this direction could be meaningful. Its scale is large and planned to grow further [3], the size of risk changes with operating states and external conditions, and the data to be integrated are heterogeneous. The prerequisite challenges, however, are many. On the technical side, whether query answering fits within time windows of seconds has not yet been confirmed, and the cost of building a mapping layer was not identified within the search scope of this review. Identifying the necessary parts is itself known to be computationally hard. On the domain side, a shared vocabulary for the safety of next-generation gaseous fuels has not yet been standardized. In addition, data acquisition itself is constrained under extreme conditions, and accident data are scarce. On the methodological side, a method for verifying a configuration that does not produce results in advance has not been established within the search scope of this review.

The contribution of this review lies in synthesizing the three approaches through three senses of judgment and in comparing the use of knowledge structures in the safety field with query-time reasoning in computer science to set out the direction that emerges between them. The empirical validation of that direction remains a task for future work. Because the direction set out here is conceptual, this review claims no superiority of any kind. The statement that query-time approaches are rare is also limited to the search scope of this review and does not mean that such studies do not exist. Judgments about when existing studies execute inference rest on abstract-level statements, and most abstracts do not report this timing, so the reasoning timing of individual studies has not been asserted. Whether knowledge structures can move beyond the layer of representation and take part at the moment of judgment remains open. That question is left for future examination under these qualifications.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/en19204842/s1. Table S1: Evidence for each cell of Tables 1–5.

Author Contributions: Conceptualization, C.Y.; methodology, C.Y., J.L. and M.M.; validation, C.Y., M.M. and S.J.; data curation, C.Y., J.L. and M.M.; writing—original draft preparation, C.Y., J.L. and M.M.; writing—review and editing, C.Y., J.L., M.M. and S.J.; supervision, M.M. and S.J.; All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Specialized Graduate School Program for Chemical Safety Management through the Korea Environmental Conservation Institute (KECI), funded by the Ministry of Climate, Energy and Environment of the Korean government (Grant No. S2026A072500001).

Data Availability Statement: No new data were created in this study. The evidence supporting each table cell is summarized in Supplementary Table S1.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. International Renewable Energy Agency. *Renewable Capacity Statistics 2025*; International Renewable Energy Agency: Abu Dhabi, United Arab Emirates, 2025.
2. Kovač, A.; Nooshmand, S. Hydrogen in the energy transition: A system-level review of production pathways, infrastructure, and deployment challenges. *Energies* **2026**, *19*, 3342. [CrossRef]
3. International Energy Agency. *Global Hydrogen Review 2025*; IEA: Paris, France, 2025.
4. Hawez, H.K.; Omar, S.R.; Alwan, L.L. Underground hydrogen storage: A comprehensive review of technologies, geological formations, and future prospects. *Energies* **2026**, *19*, 2760. [CrossRef]
5. Bai, W.; Shi, X.; Zhu, S.; Wei, X.; Huang, Y. Leakage risk assessment system for salt cavern hydrogen storage. *Deep Undergr. Sci. Eng.* **2025**, *4*, 548–561. [CrossRef]
6. Duong, P.A.; Ryu, B.R.; Song, M.K.; Nguyen, H.V.; Nam, D.; Kang, H. Safety assessment of the ammonia bunkering process in the maritime sector: A review. *Energies* **2023**, *16*, 4019. [CrossRef]
7. Zheng, Z.; Liu, D.; Sun, X.; Wang, Y.; Zhao, Y.; Xu, J. Addressing the hydrogen embrittlement challenge in future hydrogen pipelines: A multiscale review from mechanisms to material design. *Metals* **2026**, *16*, 433. [CrossRef]
8. Groth, K.M.; Li, R.; Al-Douri, A.; Ustolin, F.; Yang, R. Hydrogen safety for production, infrastructures, vehicles and aviation. *Nat. Rev. Clean Technol.* **2026**, *2*, 476–491, Correction in *Nat. Rev. Clean Technol.* **2026**, *2*, 548.. [CrossRef]
9. Shourideh, M.; Yasseri, S.; Bahai, H. Hydrogen pipelines safety using system dynamics. *Hydrogen* **2025**, *6*, 81. [CrossRef]
10. Mensah, R.A.; Correa, A.; Asante-Okyere, S.; Wallmark, C.; Försth, M. A comprehensive review of risks and mitigation strategies for safe hydrogen infrastructure deployment. *Front. Chem. Eng.* **2026**, *8*, 1721648. [CrossRef]
11. Asante-Okyere, S.; Mensah, R.A.; Sandström, J.; Försth, M. Risk and safety assessment of hydrogen pipelines and storage tanks using preliminary hazard analysis. *Front. Chem. Eng.* **2025**, *7*, 1722173. [CrossRef]
12. European Gas Pipeline Incident Data Group. *Gas Pipeline Incidents: 12th Report of the European Gas Pipeline Incident Data Group (Period 1970–2022)*; Doc. No. VA 23.0304; EGIG: Groningen, The Netherlands, 2023.
13. Min, M.; Yoon, C.; Yoo, N.; Kim, J.; Yoon, Y.; Jung, S. Hydrogen risk assessment studies: A review toward environmental sustainability. *Energies* **2025**, *18*, 229. [CrossRef]
14. Calabrese, M.; Portarapillo, M.; Di Nardo, A.; Venezia, V.; Turco, M.; Luciani, G.; Di Benedetto, A. Hydrogen safety challenges: A comprehensive review on production, storage, transport, utilization, and CFD-based consequence and risk assessment. *Energies* **2024**, *17*, 1350. [CrossRef]
15. Gregorovičová, E.; Pospíšil, J. Hydrogen safety in energy infrastructure: A review. *Energies* **2025**, *18*, 5470. [CrossRef]
16. An, Y.; Loh, T.Y.; Bhaskara, P.S.; Sin, S.M.I. Review of safety regulations, codes, and standards (RCS) for hydrogen distribution and application. *Int. J. Green Energy* **2024**, *21*, 1757–1765. [CrossRef]
17. Jiao, Z.; Zhang, Z.; Jung, S.; Wang, Q. Machine learning based quantitative consequence prediction models for toxic dispersion casualty. *J. Loss Prev. Process Ind.* **2023**, *81*, 104952. [CrossRef]
18. Phark, C.; Kim, S.; Jung, S. Development to emergency evacuation decision making in hazardous materials incidents using machine learning. *Processes* **2022**, *10*, 1046. [CrossRef]
19. Al-Sabaeei, A.M.; Alhussian, H.; Abdulkadir, S.J.; Jagadeesh, A. Prediction of oil and gas pipeline failures through machine learning approaches: A systematic review. *Energy Rep.* **2023**, *10*, 1313–1338. [CrossRef]
20. Selvam, D.C.; Raja, T.; Vaghela, D.R.; Meena, M.; Tripathy, S.; Kumar, B.; Manjunath, H.R.; Mehar, K.; Devarajan, Y. Artificial intelligence in process safety: A review of opportunities, challenges, and future directions for the chemical process industries. *J. Loss Prev. Process Ind.* **2026**, *100*, 105917. [CrossRef]

21. Devarajan, Y.; Acharya, S.K.; Jakhar, S.K.; Nagappan, B.; Saxena, S.A.; Naik, Y.; Jayabal, R.; Mehar, K. A review of model uncertainty, cyber–physical threats and assurance frameworks for safety-critical AI in process industries. *IET Cyber-Phys. Syst. Theory Appl.* **2026**, *11*, e70062. [CrossRef]
22. Khan, F.; Rathnayaka, S.; Ahmed, S. Methods and models in process safety and risk management: Past, present and future. *Process Saf. Environ. Prot.* **2015**, *98*, 116–147. [CrossRef]
23. Dunjó, J.; Fthenakis, V.; Vílchez, J.A.; Arnaldos, J. Hazard and operability (HAZOP) analysis. A literature review. *J. Hazard. Mater.* **2010**, *173*, 19–32. [CrossRef] [PubMed]
24. Jung, S.; Kim, H.; Kang, C. Efficiency analysis of the integrated application of hazard operability (HAZOP) and job safety analysis (JSA) compared to HAZOP alone for preventing fire and explosions in chemical plants. *Processes* **2025**, *13*, 88. [CrossRef]
25. Goerlandt, F.; Khakzad, N.; Reniers, G. Validity and validation of safety-related quantitative risk analysis: A review. *Saf. Sci.* **2017**, *99*, 127–139. [CrossRef]
26. Baybutt, P. Overcoming challenges in using layers of protection analysis (LOPA) to determine safety integrity levels (SILs). *J. Loss Prev. Process Ind.* **2017**, *48*, 32–40. [CrossRef]
27. Jeon, B.-i.; Kim, S.; Yoon, Y.S.; Jung, S. Quantitative risk assessment of hydrogen refueling station in Cheonan City of South Korea. *Energies* **2023**, *16*, 7138. [CrossRef]
28. Kwak, H.; Kim, M.; Min, M.; Park, B.; Jung, S. Assessing the quantitative risk of urban hydrogen refueling station in Seoul, South Korea, using SAFETI model. *Energies* **2024**, *17*, 867. [CrossRef]
29. Franks, A.; Parkar, S.; Hansen, C.R. Use of quantitative risk assessment to enhance the safety of ships using ammonia as fuel. *J. Saf. Sustain.* **2024**, *1*, 127–140. [CrossRef]
30. Lee, J.; Kwak, H.; Jung, S. Quantitative risk assessment of steam reforming process by hydrogen generator, using PHAST model. *Energies* **2024**, *17*, 5704. [CrossRef]
31. Jang, C.B.; Jung, S. Numerical computation of a large-scale jet fire of high-pressure hydrogen in process plant. *Energy Sci. Eng.* **2016**, *4*, 406–417. [CrossRef]
32. Kim, Y.; Han, S.; Song, H.; Jung, S. Risk assessment of an off-site hydrogen refueling station: A hybrid IEC 61511-CCPS LOPA framework. *Energies* **2025**, *18*, 6242. [CrossRef]
33. Braarud, P.Ø.; Simensen, J.E. Analyst behaviour and team processes during hazard analysis: The development of an observation protocol and initial results from evaluating HAZOP sessions. *Saf. Sci.* **2025**, *181*, 106694. [CrossRef]
34. Elhosary, E.; Moselhi, O. Automation for HAZOP study: A state-of-the-art review and future research directions. *J. Inf. Technol. Constr.* **2024**, *29*, 750–777. [CrossRef]
35. Villa, V.; Paltrinieri, N.; Khan, F.; Cozzani, V. Towards dynamic risk analysis: A review of the risk assessment approach and its limitations in the chemical process industry. *Saf. Sci.* **2016**, *89*, 77–93. [CrossRef]
36. Ab Rahim, M.S.; Reniers, G.; Yang, M.; Bajpai, S. Risk assessment methods for process safety, process security and resilience in the chemical process industry: A thorough literature review. *J. Loss Prev. Process Ind.* **2024**, *88*, 105274. [CrossRef]
37. Pasman, H.J.; Rogers, W.J.; Mannan, M.S. Risk assessment: What is it worth? Shall we just do away with it, or can it do a better job? *Saf. Sci.* **2017**, *99*, 140–155. [CrossRef]
38. Yousefi, A.; Rodriguez Hernandez, M.; Lopez Peña, V. Systemic accident analysis models: A comparison study between AcciMap, FRAM, and STAMP. *Process Saf. Prog.* **2019**, *38*, e12002. [CrossRef]
39. Bjerga, T.; Aven, T.; Zio, E. Uncertainty treatment in risk analysis of complex systems: The cases of STAMP and FRAM. *Reliab. Eng. Syst. Saf.* **2016**, *156*, 203–209. [CrossRef]
40. Yoon, C.; Bang, B.-H.; Lee, K.; Jung, S.; Min, M. Inherent safety design for LNG terminals through risk assessment. *Korean J. Chem. Eng.* **2025**, *42*, 2713–2728. [CrossRef]
41. You, W.; Park, J.; Jung, S.; Lim, Y. Risk and efficiency analysis of dual mixed refrigerant liquefaction process configurations for floating liquefied natural gas at conceptual design stage. *Process Saf. Prog.* **2019**, *38*, 87–98. [CrossRef]
42. Ji, C.; Ma, F.; Rao, J.; Wang, J.; Sun, W. Bridging the gap in chemical process monitoring: Beyond algorithm-centric research toward industrial deployment. *Processes* **2025**, *13*, 3809. [CrossRef]
43. Leite, D.; Andrade, E.; Rativa, D.; Maciel, A.M.A. Fault detection and diagnosis in Industry 4.0: A review on challenges and opportunities. *Sensors* **2025**, *25*, 60. [CrossRef] [PubMed]
44. Lu, B.; Chiang, L. Semi-supervised online soft sensor maintenance experiences in the chemical industry. *J. Process Control* **2018**, *67*, 23–34. [CrossRef]
45. Sudarshan, V.; Seider, W.D. Advancing machine learning in Industry 4.0: Benchmark framework for rare-event prediction in chemical processes. *Comput. Chem. Eng.* **2025**, *194*, 108929. [CrossRef]
46. Geirhos, R.; Jacobsen, J.-H.; Michaelis, C.; Zemel, R.; Brendel, W.; Bethge, M.; Wichmann, F.A. Shortcut learning in deep neural networks. *Nat. Mach. Intell.* **2020**, *2*, 665–673. [CrossRef]

47. Pathak, A.K.; Gupta, M.; Jain, G. Unmasking the Clever Hans effect in AI models: Shortcut learning, spurious correlations, and the path toward robust intelligence. *Front. Artif. Intell.* **2026**, *8*, 1692454. [CrossRef] [PubMed]
48. Moreno-Torres, J.G.; Raeder, T.; Alaiz-Rodríguez, R.; Chawla, N.V.; Herrera, F. A unifying view on dataset shift in classification. *Pattern Recognit.* **2012**, *45*, 521–530. [CrossRef]
49. Koh, P.W.; Sagawa, S.; Marklund, H.; Xie, S.M.; Zhang, M.; Balsubramani, A.; Hu, W.; Yasunaga, M.; Phillips, R.L.; Gao, I.; et al. WILDS: A benchmark of in-the-wild distribution shifts. In *Proceedings of the 38th International Conference on Machine Learning*; Meila, M., Zhang, T., Eds.; PMLR: Cambridge, MA, USA, 2021; Volume 139, pp. 5637–5664.
50. Rabanser, S.; Günnemann, S.; Lipton, Z.C. Failing loudly: An empirical study of methods for detecting dataset shift. In *Advances in Neural Information Processing Systems 32 (NeurIPS 2019)*; Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E., Garnett, R., Eds.; Curran Associates: Brooklyn, NY, USA, 2019.
51. Lobo, J.L.; Laña, I.; Osaba, E.; Del Ser, J. On the connection between concept drift and uncertainty in industrial artificial intelligence. In Proceedings of the 2023 IEEE Conference on Artificial Intelligence (CAI), Santa Clara, CA, USA, 5–6 June 2023; pp. 171–172. [CrossRef]
52. Li, S.; Drgona, J.; Abhyankar, S.; Pileggi, L. Power grid behavioral patterns and risks of generalization in applied machine learning. In *Companion Proceedings of the 14th ACM International Conference on Future Energy Systems (e-Energy '23 Companion), Orlando, FL, USA, 20–23 June 2023*; ACM: New York, NY, USA, 2023; pp. 106–114. [CrossRef]
53. Chekroud, A.M.; Hawrilenko, M.; Loho, H.; Bondar, J.; Gueorguieva, R.; Hasan, A.; Kambeitz, J.; Corlett, P.R.; Koutsouleris, N.; Krumholz, H.M.; et al. Illusory generalizability of clinical prediction models. *Science* **2024**, *383*, 164–167. [CrossRef] [PubMed]
54. Mbagaya, L.; Reddy, K.; Botes, A. Machine learning techniques for battery state of health prediction: A comparative review. *World Electr. Veh. J.* **2025**, *16*, 594. [CrossRef]
55. Kapoor, S.; Narayanan, A. Leakage and the reproducibility crisis in machine-learning-based science. *Patterns* **2023**, *4*, 100804. [CrossRef] [PubMed]
56. Rudin, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat. Mach. Intell.* **2019**, *1*, 206–215. [CrossRef] [PubMed]
57. Kumar, I.E.; Venkatasubramanian, S.; Scheidegger, C.; Friedler, S. Problems with Shapley-value-based explanations as feature importance measures. In *Proceedings of the 37th International Conference on Machine Learning (ICML)*; PMLR: Cambridge, MA, USA, 2020; Volume 119, pp. 5491–5500. [CrossRef]
58. Patriarca, R.; Chatzimichailidou, M.; Karanikas, N.; Di Gravio, G. The past and present of System-Theoretic Accident Model And Processes (STAMP) and its associated techniques: A scoping review. *Saf. Sci.* **2022**, *146*, 105566. [CrossRef]
59. Özdemir, P.; Yıldırım, R. ML@ChemE: Past, present, and future of machine learning in chemical engineering. *ChemBioEng Rev.* **2025**, *12*, e70012. [CrossRef]
60. Chesterman, X.; Verstraeten, T.; Daems, P.-J.; Nowé, A.; Helsen, J. Overview of normal behavior modeling approaches for SCADA-based wind turbine condition monitoring demonstrated on data from operational wind farms. *Wind Energ. Sci.* **2023**, *8*, 893–924. [CrossRef]
61. Bai, Y.; Wu, J.; Ren, Q.; Jiang, Y.; Cai, J. A BN-based risk assessment model of natural gas pipelines integrating knowledge graph and DEMATEL. *Process Saf. Environ. Prot.* **2023**, *171*, 640–654. [CrossRef]
62. Yu, Z.; Zhou, J.; Wang, Q.; Zhou, H.; Song, Y.; Zong, C. Construction of knowledge graph for gas polyethylene pipelines based on ALBERT-BiGRU-CRF. *Sci. Rep.* **2025**, *15*, 25002. [CrossRef] [PubMed]
63. Mahajan, D.; Tan, K.; Venkatesh, T.; Kileti, P.; Clayton, C.R. Hydrogen blending in gas pipeline networks—A review. *Energies* **2022**, *15*, 3582. [CrossRef]
64. Tian, X.; Pei, J. Study progress on the pipeline transportation safety of hydrogen-blended natural gas. *Heliyon* **2023**, *9*, e21454. [CrossRef] [PubMed]
65. An, N.Y.; Yang, J.H.; Song, E.; Hwang, S.-H.; Byun, H.-G.; Park, S. Digital twin-based hydrogen refueling station (HRS) safety model: CNN-based decision-making and 3D simulation. *Sustainability* **2024**, *16*, 9482. [CrossRef]
66. Liu, C.; Zhu, S.; Yin, Y.; Xiao, K.; Chen, X.; Liu, W.; Li, Y. A leakage monitoring technology for buried hydrogen-doped natural gas pipelines based on vibration signal with machine learning. *Int. J. Hydrogen Energy* **2025**, *131*, 118–135. [CrossRef]
67. Azimi, H.; Shoghi, R.; Shiri, H. Digital twin technology in pipeline engineering: A study review of applications, challenges, and future directions. *Technologies* **2026**, *14*, 479. [CrossRef]
68. Karimi Dehkordi, M.; Sattari, F.; Lefsrud, L. Exploring incident patterns in the hydrogen value chain using knowledge graphs: A roadmap toward targeted risk control. *Reliab. Eng. Syst. Saf.* **2026**, *267*, 111790. [CrossRef]
69. Wang, X.; Zhang, Z.; Zhang, J.; Zhao, Y.; Wang, Z. Sustainable transition of underground gas storage: A unified engineering framework from methane and carbon dioxide to hydrogen. *Sustainability* **2026**, *18*, 4622. [CrossRef]

70. Wang, J.; Zhao, P.; Liu, X.; Lin, Y.; Song, Z.; Liu, W.; Wang, J. Long-term safety analysis and evaluation on the horizontal salt cavern gas storage. *J. Energy Storage* **2026**, *147*, 120279. [CrossRef]
71. Vogiatzaki, K.; Tretola, G.; Cesmat, L. Digital twins for cryogenic hydrogen safety: Integrating computational fluid dynamics and machine learning. *Hydrogen* **2025**, *6*, 110. [CrossRef]
72. Luan, T.; Li, H.; Wang, K.; Zhang, X.; Li, X. Knowledge graph-based Bayesian network risk assessment model for hydrogen accidents. *Int. J. Hydrogen Energy* **2024**, *81*, 927–941. [CrossRef]
73. Yang, M.; Lam, J.S.L. Risk assessment of ammonia bunkering operations: Perspectives on different release scales. *J. Hazard. Mater.* **2024**, *468*, 133757. [CrossRef] [PubMed]
74. Duong, P.A.; Bae, J.-W.; Lee, C.; Yang, D.H.; Kang, H. Numerical study on ammonia dispersion and explosion characteristics in confined space of marine fuel preparation room. *J. Mar. Sci. Eng.* **2025**, *13*, 1235. [CrossRef]
75. Natarajan, S.; Ghosh, K.; Srinivasan, R. An ontology for distributed process supervision of large-scale chemical plants. *Comput. Chem. Eng.* **2012**, *46*, 124–140. [CrossRef]
76. Zheng, X.; Wang, B.; Zhao, Y.; Mao, S.; Tang, Y. A knowledge graph method for hazardous chemical management: Ontology design and entity identification. *Neurocomputing* **2021**, *430*, 104–111. [CrossRef]
77. Song, X.; Yin, F.; Zhao, D. Knowledge graph-based alarm management in petrochemical enterprises: A study on fusion and analysis of multi-source heterogeneous information. *J. Loss Prev. Process Ind.* **2025**, *97*, 105706. [CrossRef]
78. Single, J.I.; Schmidt, J.; Denecke, J. Knowledge acquisition from chemical accident databases using an ontology-based method and natural language processing. *Saf. Sci.* **2020**, *129*, 104747. [CrossRef]
79. Zhao, Y.; Zhang, B.; Gao, D. Construction of petrochemical knowledge graph based on deep learning. *J. Loss Prev. Process Ind.* **2022**, *76*, 104736. [CrossRef]
80. Zhao, K.; Lu, X.; Wan, L.; Zhang, L.; Jin, Y.; Wen, P.; Gao, J.; He, M.; Wang, Q.; Zhang, L. From data to insight: Building a knowledge graph for risk analysis of hazardous chemical accidents. *Chin. J. Chem. Eng.* **2026**, *91*, 92–110. [CrossRef]
81. Zheng, C.; Chen, G.; Chen, H.; Xu, Q.; Zhao, Y.; Zhao, Y.; Yang, Y. Chemical process safety domain knowledge graph–enhanced LLM for efficient emergency response decision support. *Can. J. Chem. Eng.* **2025**, *103*, 4814–4843. [CrossRef]
82. Liu, M.; Huang, R.; Xu, F. Research on the construction of safety information ontology knowledge base and accident reasoning for complex hazardous production systems—Taking methanol production process as an example. *Sustainability* **2023**, *15*, 2568. [CrossRef]
83. Musulin, E.; Roda, F.; Basualdo, M. A knowledge-driven approach for process supervision in chemical plants. *Comput. Chem. Eng.* **2013**, *59*, 164–177. [CrossRef]
84. Single, J.I.; Schmidt, J.; Denecke, J. Ontology-based computer aid for the automation of HAZOP studies. *J. Loss Prev. Process Ind.* **2020**, *68*, 104321. [CrossRef] [PubMed]
85. Teimourikia, M.; Fugini, M. Ontology development for run-time safety management methodology in Smart Work Environments using ambient knowledge. *Future Gener. Comput. Syst.* **2017**, *68*, 428–441. [CrossRef]
86. Shen, Q.; Wu, S.; Deng, Y.; Deng, H.; Cheng, J.C.P. BIM-based dynamic construction safety rule checking using ontology and natural language processing. *Buildings* **2022**, *12*, 564. [CrossRef]
87. Liu, P.; Jin, X.; Shang, Y. Ontology-based automated knowledge identification of safety risks in Metro shield construction. *J. Asian Archit. Build. Eng.* **2026**, *25*, 1295–1324. [CrossRef]
88. Riesco, R.; Villagrá, V.A. Leveraging cyber threat intelligence for a dynamic risk framework. *Int. J. Inf. Secur.* **2019**, *18*, 715–739. [CrossRef]
89. Vega-Barbas, M.; Villagrá, V.A.; Monje, F.; Riesco, R.; Larriva-Novo, X.; Berrocal, J. Ontology-based system for dynamic risk management in administrative domains. *Appl. Sci.* **2019**, *9*, 4547. [CrossRef]
90. De Bartolomeo, M.; De Nicola, A. Ontologies and knowledge graphs for railway safety. *Safety* **2025**, *11*, 100. [CrossRef]
91. Soularidis, A.; Kotis, K.I.; Vouros, G.A. Real-time semantic data integration and reasoning in life- and time-critical decision support systems. *Electronics* **2024**, *13*, 526. [CrossRef]
92. Du, W.; Wang, X.; Zhu, Q.; Jing, X.; Liu, X. CPBA-CLIM: An entity-relation extraction model for ontology-based knowledge graph construction in hazardous chemical incident management. *Sci. Prog.* **2024**, *107*, 368504241235510. [CrossRef] [PubMed]
93. Tang, W.; Zhang, Y.; Mao, X.; Jia, H.; Lv, K.; Shan, L.; Qiao, Y.; Jiang, T. Construction and application of knowledge graph for power grid new equipment start-up. *Energies* **2025**, *18*, 5471. [CrossRef]
94. Sánchez-Zas, C.; Villagrá, V.A.; Vega-Barbas, M.; Larriva-Novo, X.; Moreno, J.I.; Berrocal, J. Ontology-based approach to real-time risk management and cyber-situational awareness. *Future Gener. Comput. Syst.* **2023**, *141*, 462–472. [CrossRef]
95. Sánchez-Zas, C.; Larriva-Novo, X.; Villagrá, V.A.; Rivera, D.; Marín-Lopez, A. A methodology for ontology-based interoperability of dynamic risk assessment frameworks in IoT environments. *Internet Things* **2024**, *27*, 101267. [CrossRef]

96. Motik, B.; Nenov, Y.; Piro, R.; Horrocks, I. Maintenance of datalog materialisations revisited. *Artif. Intell.* **2019**, *269*, 76–136. [CrossRef]
97. Belhajjame, K.; Mejri, M.-Y. Online maintenance of evolving knowledge graphs with RDFS-based saturation and why-provenance support. *J. Web Semant.* **2023**, *78*, 100796. [CrossRef]
98. Nenov, Y.; Piro, R.; Motik, B.; Horrocks, I.; Wu, Z.; Banerjee, J. RDFox: A highly-scalable RDF store. In *The Semantic Web—ISWC 2015*; Arenas, M., Corcho, O., Simperl, E., Strohmaier, M., d'Aquin, M., Srinivas, K., Groth, P., Dumontier, M., Heflin, J., Thirunarayan, K., et al., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2015; pp. 3–20. [CrossRef]
99. Poggi, A.; Lembo, D.; Calvanese, D.; De Giacomo, G.; Lenzerini, M.; Rosati, R. Linking data to ontologies. In *Journal on Data Semantics X*; Spaccapietra, S., Ed.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2008; pp. 133–173. [CrossRef]
100. Xiao, G.; Calvanese, D.; Kontchakov, R.; Lembo, D.; Poggi, A.; Rosati, R.; Zakharyaschev, M. Ontology-based data access: A survey. In Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI-18), Stockholm, Sweden, 13–19 July 2018; pp. 5511–5519. [CrossRef] [PubMed]
101. Calvanese, D.; Cogrel, B.; Komla-Ebri, S.; Kontchakov, R.; Lanti, D.; Rezk, M.; Rodriguez-Muro, M.; Xiao, G. Ontop: Answering SPARQL queries over relational databases. *Semant. Web* **2016**, *8*, 471–487. [CrossRef]
102. Xiao, G.; Ding, L.; Cogrel, B.; Calvanese, D. Virtual knowledge graphs: An overview of systems and use cases. *Data Intell.* **2019**, *1*, 201–223. [CrossRef]
103. Ahmetaj, S. Rewriting approaches for ontology-mediated query answering. *KI Künstl. Intell.* **2020**, *34*, 523–526. [CrossRef] [PubMed]
104. Botoeva, E.; Calvanese, D.; Cogrel, B.; Corman, J.; Xiao, G. Ontology-based data access—Beyond relational sources. *Intell. Artif.* **2019**, *13*, 21–36. [CrossRef]
105. Calvanese, D.; Lanti, D.; Mendes De Farias, T.; Mosca, A.; Xiao, G. Accessing scientific data through knowledge graphs with Ontop. *Patterns* **2021**, *2*, 100346. [CrossRef] [PubMed]
106. Ranatunga, S.; Ødegård, R.S.; Jetlund, K.; Onstein, E. Use of semantic web technologies to enhance the integration and interoperability of environmental geospatial data: A framework based on ontology-based data access. *ISPRS Int. J. Geo-Inf.* **2025**, *14*, 52. [CrossRef]
107. Urbani, J.; van Harmelen, F.; Schlobach, S.; Bal, H. QueryPIE: Backward reasoning for OWL Horst over very large knowledge bases. In *The Semantic Web—ISWC 2011*; Aroyo, L., Welty, C., Alani, H., Taylor, J., Bernstein, A., Kagal, L., Noy, N., Blomqvist, E., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2011; pp. 730–745. [CrossRef]
108. Qin, X.; Zhang, X.; Yasin, M.Q.; Wang, S.; Feng, Z.; Xiao, G. SUMA: A partial materialization-based scalable query answering in OWL 2 DL. *Data Sci. Eng.* **2021**, *6*, 229–245. [CrossRef]
109. Sequeda, J.F.; Arenas, M.; Miranker, D.P. OBDA: Query rewriting or materialization? In practice, both! In *The Semantic Web—ISWC 2014*; Mika, P., Tudorache, T., Bernstein, A., Welty, C., Knoblock, C., Vrandečić, D., Groth, P., Noy, N., Janowicz, K., Goble, C., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2014; pp. 535–551. [CrossRef]
110. Kharlamov, E.; Hovland, D.; Skjæveland, M.G.; Bilidas, D.; Jiménez-Ruiz, E.; Xiao, G.; Soylu, A.; Lanti, D.; Rezk, M.; Zheleznyakov, D.; et al. Ontology based data access in Statoil. *J. Web Semant.* **2017**, *44*, 3–36. [CrossRef]
111. Dell'Aglio, D.; Della Valle, E.; van Harmelen, F.; Bernstein, A. Stream reasoning: A survey and outlook. *Data Sci.* **2017**, *1*, 59–83. [CrossRef]
112. Bonte, P.; Callé, C.; Curé, O.; Kondylakis, H.; Tommasini, R. Languages and systems for RDF stream processing, a survey. *VLDB J.* **2025**, *34*, 50. [CrossRef]
113. Tataroğlu Özbulak, G.A.; Manzo, G.; Shrestha, Y.R.; Calbimonte, J.-P. A comprehensive survey of stream reasoning and its integration with knowledge graphs. *Knowl. Inf. Syst.* **2025**, *67*, 11225–11257. [CrossRef]
114. Stavropoulou, G.; Tsitseklis, K.; Mavraidi, L.; Chang, K.-I.; Zafeiropoulos, A.; Karyotis, V.; Papavassiliou, S. Digital twin meets knowledge graph for intelligent manufacturing processes. *Sensors* **2024**, *24*, 2618. [CrossRef] [PubMed]
115. Bilidas, D.; Mantas, A.; Yfantis, F.; Stamoulis, G.; Koubarakis, M.; Kondylatos, S.; Prapas, I.; Papoutsis, I. Fire risk management using data cubes, machine learning and OBDA systems. In *Proceedings of the 31st ACM International Conference on Advances in Geographic Information Systems*; ACM: New York, NY, USA, 2023; pp. 1–4. [CrossRef]
116. Noor, A. Integrating mechanistic information to predict drug-drug interactions and associated relevance for decision support. In *2022 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS)*; IEEE: New York, NY, USA, 2022; pp. 1–4. [CrossRef]
117. Armas Romero, A.; Kaminski, M.; Cuenca Grau, B.; Horrocks, I. Module extraction in expressive ontology languages via datalog reasoning. *J. Artif. Intell. Res.* **2016**, *55*, 499–564. [CrossRef]
118. Yang, H.; Kareck, T.L.; Wang, Q. New era of AI in chemical process safety: Foundation models. *ACS Chem. Health Saf.* **2026**, *33*, 171–179. [CrossRef]

119. Haller, A.; Janowicz, K.; Cox, S.; Le Phuoc, D.; Taylor, K.; Lefrançois, M. (Eds.) *Semantic Sensor Network Ontology*; W3C Recommendation, 19 October 2017; W3C: Wakefield, MA, USA; OGC: Arlington, VA, USA, 2017. Available online: https://www.w3.org/TR/vocab-ssn/ (accessed on 30 September 2026).
120. Renkhoff, J.; Feng, K.; Meier-Doernberg, M.; Velasquez, A.; Song, H.H. A survey on verification and validation, testing and evaluations of neurosymbolic artificial intelligence. *IEEE Trans. Artif. Intell.* **2024**, *5*, 3765–3779. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.