Automation in Cyber Risk Management: A Literature Review of Evidence, Gaps, and Emerging Directions

Antonio Belli[1,*], Nicolò Maunero[1,*] and Paolo Prinetto[2]

[1]IMT Scuola Alti Studi Lucca, Lucca, Italy
[2]CINI Cybersecurity National Lab, Torino, Italy

Abstract

Organizations that run large or complex ICT infrastructures face a persistent flow of vulnerabilities, alerts, configuration deviations, and regulatory demands. Although risk management standards and security frameworks are widely adopted and generally mature, daily practice still relies on periodic, interview-based assessments that are only weakly connected to operational telemetry and often reflect subjective biases rather than the actual system or organization behavior. At the same time, automation technologies (continuous monitoring, standardized security data formats, and advances in machine learning and language models) are reshaping how security evidence can be collected, normalized, and interpreted. This paper presents a structured literature review of research and practice on automation in cyber risk management, with a focus on risk and governance frameworks; continuous monitoring and "evidence-as-code"; scenario-based methods; human factors such as information overload; alert and decision fatigue; and emerging work on compound/cascading cyber risk. The review consolidates contributions from standards bodies, empirical studies, and recent academic work, and identifies structural gaps, including the weak alignment between telemetry and governance processes, the scarcity of reusable scenario definitions, and the limited availability of simple, explainable models that connect continuous evidence with risk estimation. The paper concludes with future research directions towards integrated, automation-assisted cyber risk governance, particularly for mid-to-large ICT enterprises.

Keywords

cybersecurity risk, automation, literature review, research directions

1. Introduction

Organizations operating large and interconnected ICT infrastructures increasingly rely on risk management frameworks to guide security investments, compliance, and accountability. However, in many enterprises, risk assessment remains a predominantly periodic [1] and documentation-driven exercise: risks are elicited through interviews, registered in static spreadsheets [2]. This approach struggles to keep pace with infrastructures that are continuously changing (cloud adoption, rapid software delivery, third-party dependencies) and with security evidence that is now abundant, heterogeneous, and generated at high frequency (logs, vulnerabilities, identity events, configuration drift).

Automation is often presented as the obvious answer, yet practice shows a recurrent mismatch: automation can easily increase the volume of signals without necessarily improving the quality of governance decisions [3][4]. Moreover, across major datasets [5][6][7][8] ransomware and credential/identity-driven intrusion remain prevalent, while third-party / supply-chain involvement is increasingly visible. In parallel, NIS2 [9] increases pressure for demonstrable and continuously maintained risk management.

For these reasons, we propose a structured literature review to underline recurring problems that limit the practical value of automation in cyber risk management, identify structural gaps and propose future research directions towards integrated, automation-assisted cyber risk governance.

The motivation for this review is pragmatic: organizations are not lacking tools to collect evidence, but they are lacking robust approaches to turn evidence into risk decisions that are understandable, auditable, and sustainable over time. For this reason, we deliberately focus on the parts of the automation

Joint National Conference on Cybersecurity (ITASEC & SERICS 2026), February 09-13, 2026, Cagliari, IT
*Corresponding author.
antonio.belli@imtlucca.it (A. Belli); nicolo.maunero@imtlucca.it (N. Maunero); paolo.prinetto@cybersecnatlab.it (P. Prinetto)
0000-0002-4331-1066 (N. Maunero)

landscape that are most directly relevant to governance outcomes, rather than on automation as "more security tooling". In particular, we emphasize (i) methods that explicitly account for human-factor constraints (attention, cognitive load, decision fatigue) rather than assuming infinite analyst capacity; (ii) approaches that connect continuous operational evidence to scenario-level risk representation; and (iii) models that acknowledge interdependencies and contagion effects across incidents. Hence, we frame the literature through three main problems:

1. **P1** (*decision overload/alert fatigue*): security and risk stakeholders face an expanding flow of alerts and metrics, while their attention and decision capacity remain bounded.
2. **P2** (*evidence-to-governance gap*): risk registers and interview-based assessments tend to drift from operational reality, because day-to-day telemetry is rarely transformed into governance-ready risk views.
3. **P3** (*compound and cascading risk*): under crisis conditions, events interact (e.g., ransomware diverting resources while credential theft enables further compromise), but many methods still treat risks as isolated and independent.

This paper, therefore, examines how far existing standards, methods, and tools support the identified necessities, and what remains missing for effective adoption in mid-to-large ICT environments. It consolidates contributions from standards bodies and institutional guidance together with academic work on continuous monitoring, evidence-as-code, maturity models, ontologies, and emerging research on compound/cascading cyber risk. The goal is not to propose a new automation architecture, but to clarify what the current body of knowledge enables, where it fails, and which research directions appear most promising when the objective is governance-grade risk management rather than tool-level optimization.

This paper contributes by: (i) structuring the landscape of automation in cyber risk management through the lens of P1–P3; (ii) synthesizing the main solution families and their limitations in terms of evidence-to-decision traceability and human sustainability; and (iii) outlining emerging directions towards scenario-centered, human-aware, and evidence-driven automation that are realistically adoptable in mid-to-large ICT enterprises.

In this paper, mid-to-large ICT enterprises refer to organizations operating heterogeneous, continuously changing ICT estates (multi-service, multi-team environments) where risk governance requires structured registers, recurring management review, and regulatory accountability (e.g., NIS2-like obligations). As a practical proxy, this includes organizations above typical Small and medium-sized enterprise (SME) complexity thresholds (e.g., multiple business units and centralized IAM/SOC/GRC functions), even when headcount-based definitions vary by sector.

This paper is structured as follows. Section 2 summarizes the main standards/frameworks used in cyber risk governance and highlights automation touchpoints. Sections 3-6 describe the review method, synthesize the literature by P1–P3, discuss gaps, and outline research directions.

2. Background. Risk Management Standards and Frameworks

This section provides the conceptual and normative baseline for the review by summarizing the main risk management standards and institutional frameworks used in practice. We focus on what they prescribe (and what they leave open) with respect to risk-based governance, control selection, and evidence collection, highlighting the specific touchpoints that can enable - or constrain - automation. The discussion covers ISO/IEC 27001 and related ISO standards, enterprise-risk principles [10], and complementary guidance from European and NIST sources, to frame the rest of the paper on a shared vocabulary and scope.

2.1. ISO/IEC 27001: Risk-Based Information Security Management

ISO/IEC 27001:2022 [11] defines the requirements for an information security management system (ISMS) that is explicitly risk–based. Organizations are required to understand their context and stakeholders,

define the ISMS scope, and establish processes to assess and treat information security risks as part of overall planning.

Risk assessment and risk treatment appear both in the planning phase and in operation, reinforcing the idea of an iterative cycle rather than a one-off exercise. Monitoring, measurement, analysis, and evaluation, internal audit, and management review close the loop, with continual improvement. Annex A then points to a reference set of controls aligned with ISO/IEC 27002:2022 [12], to be selected based on risk treatment decisions.

From the perspective of this paper, ISO/IEC 27001 provides three important baselines: (i) it formalizes risk as the organizing principle for information security governance; (ii) it embeds risk activities into a PDCA-like cycle (plan, operate, evaluate, improve); (iii) it links governance decisions to a standardized control set, but without specifying how continuous operational telemetry should feed risk estimates.

This creates a first tension with day-to-day practice in complex ICT environments: while the standard is compatible with continuous updating, many organizations implement it through periodic, interview-based reviews and static risk registers.

2.2. ISO/IEC 27002: Control Catalogue and Attribute-Based Views

ISO/IEC 27002:2022 [12] provides the reference catalogue of information security controls used by ISO/IEC 27001. In the 2022 version of the standard, controls are grouped into four themes (organizational, people, physical, and technological) and described with purpose and implementation guidance.

Another key evolution in the 2022 edition is the introduction of control attributes (Annex A). Each control is tagged along several dimensions: control type (preventive, detective, corrective), information security properties (confidentiality, integrity, availability), cybersecurity concepts aligned with Identify–Protect–Detect–Respond–Recover, operational capabilities (e.g., threat and vulnerability management, identity and access management), and security domains (governance & ecosystem, protection, defence, resilience). Organizations are encouraged to extend this model with their own attributes (for example, events, assets involved, implementation state, or even risk scenarios) to generate tailored views and filtering of controls.

Two aspects are directly relevant to the observed problems: (i) scenario and event attributes. The standard explicitly suggests associating controls to "events" or risk scenarios, managed via spreadsheets or databases, to accelerate risk treatment and ensure no necessary control is overlooked; (ii) continuous monitoring and alerting. The guidance on logging and monitoring (e.g., control 8.16 "Monitoring activities") emphasizes automated tools, real-time or periodic monitoring, large-scale data handling, and alerting tuned to organizational baselines. It explicitly mentions the use of machine learning and AI to enhance anomaly detection and the need to minimize false positives.

These provisions acknowledge the availability of rich telemetry and automated detection, but they stop at the control level: they describe what should be monitored and how alerts should be handled, without providing a structured method to translate this stream of evidence into updated risk scenarios or governance-ready risk metrics.

2.3. ISO/IEC 27005: Information Security Risk Management

ISO/IEC 27005:2022 [13] focuses specifically on the information security risk management process to support the implementation of an ISMS. It structures the process into context establishment, risk assessment (identification, analysis, evaluation), risk treatment, risk acceptance, and risk communication and consultation, with monitoring and review across all phases.

The standard promotes the use of risk scenarios that combine assets, threats, vulnerabilities, and consequences. It offers guidance on: (i) identifying assets (including owners and business processes); (ii) identifying threats and vulnerabilities (internal, external, accidental, deliberate); (iii) determining consequences (including loss of effectiveness, operational disruption, financial and reputational damage); (iv) choosing appropriate risk analysis methods (qualitative, quantitative, or hybrid); (v) assessing likelihood and impact using available information and expert judgment.

Another relevant limitation is that, although the standard allows for combinations of risks and acknowledges that consequences can arise from sequences of events, it does not offer a tractable methodology for compound or cascading risk suitable for governance dashboards.

2.4. ISO 31000: Enterprise Risk Management Principles

ISO 31000:2018 [10] provides generic guidelines for risk management across domains. It defines risk as the "effect of uncertainty on objectives" and sets out principles, a management framework, and a process applicable at strategic, program, and operational levels.

The principles stress that risk management should be: (i) integrated into all organizational activities; (ii) structured and comprehensive; (iii) customized to the organizational context; (iv) inclusive, taking into account stakeholders' knowledge and perceptions; (v) dynamic, recognizing that risks can change quickly; (vi) based on the best available information, while acknowledging limitations, uncertainties, and biases.

The process section (clause 6) describes an iterative cycle of communication and consultation, scope/context and criteria definition, risk assessment (identification, analysis, evaluation), risk treatment, and monitoring, review, recording and reporting.

ISO 31000 contributes two key angles: (i) it legitimizes the idea that risk management has to balance rich information with human cognitive limits: "best available information" is required, but the standard explicitly warns about uncertainties, data limitations, and the influence of stakeholders' perceptions and biases; (ii) it frames risk combinations and sequences as part of the analysis, but again leaves the operational modeling (e.g., co-occurrence metrics, attack-graph models) to domain-specific methods.

2.5. European work on interoperable risk management frameworks

ENISA's *Interoperable EU Risk Management Framework* [14] analyses a broad set of risk management methods and frameworks with a specific focus on how easily their components can be combined and reused. It defines interoperability in this context as the ability for risk management components (asset taxonomies, threat and vulnerability catalogues, calculation methods, control catalogues) to reuse information from other frameworks. For this paper, ENISA's work is relevant because it explicitly treats risk frameworks themselves as composable objects. It provides a vocabulary to compare methods such as ISO 27005, NIST 800-30/39 [15][16], and national frameworks, and to reason about where automation can realistically "plug" telemetry and external catalogues into existing processes. At the same time, the report confirms that differences in valuation scales, scenario structure and residual-risk formulas still hinder reuse of scenarios and metrics across tools and organizations—directly connecting to the problem of fragmented, non-standardized inputs to automated risk estimation (Problems 2 and 3).

2.6. Threat-landscape methodologies and cyber threat intelligence as risk inputs

The *ENISA Cybersecurity Threat Landscape (CTL) Methodology* [5] formalizes how threat intelligence can be transformed into a structured "threat landscape" through a cycle of Direction, Collection, Processing, Analysis & Production, Dissemination. The methodology stresses the need to define intelligence requirements and audience ("direction"), to plan and validate heterogeneous sources, and to normalize and enrich data through taxonomies and CTI formats before producing analytical outputs. In practice this means that a large volume of raw events and reports is filtered, tagged and aggregated into a smaller set of threat scenarios, trends and indicators meant to be consumable by decision-makers.

Methodologies like CTL are implicitly designed to mitigate information overload: they introduce explicit steps for source validation, language processing and consolidation, and they rely on standard taxonomies and CTI frameworks to keep heterogeneous data manageable. However, CTL stops short of prescribing how its outputs should feed into specific risk registers or governance frameworks; the mapping from threat landscape products to concrete changes in risk scenarios, control priorities or residual-risk estimates is left to local practice.

In parallel, operational knowledge bases (e.g., ATT&CK [17], KEV [18], OSCAL [19]) provide machine-readable artefacts for techniques, exploited vulnerabilities, and control/assessment structures. The core gap is therefore not data availability, but the lack of stable model alignment and interfaces that connect CTI/telemetry artefacts to scenario-based governance views (P1 and P2).

2.7. NIST guidance on continuous monitoring, risk registers and ERM integration

NIST SP 800-39 defines a multi-tier risk management approach (organization, mission/business process, information system) and emphasizes the need to treat cybersecurity risk as an integral part of enterprise risk, governed through formal processes and risk appetite statements. NIST SP 800-30 focuses on the risk assessment process itself, detailing preparation, execution, communication and maintenance activities, and highlighting the need to consider both likelihood and impact while accounting for existing controls. SP 800-137 [20] on Information Security Continuous Monitoring (ISCM) extends this view to an ongoing cycle where security controls are monitored with frequencies tailored to their volatility and criticality, and where automation is explicitly encouraged for highly volatile and machine-checkable controls (e.g., automatic disabling of inactive accounts, detection of unauthorized components).

NISTIR 8286 [21] makes the connection to enterprise risk explicit by advocating cybersecurity risk registers that can be "rolled up" into an enterprise risk register aligned with strategic, operational, reporting and compliance objectives. It discusses how consistent scales for likelihood, impact and residual risk, together with explicit mappings between control families and enterprise objectives, can support aggregation and prioritization at board level, and stresses the role of governance bodies (CIO/CISO councils, enterprise risk steering committees) in coordinating this integration.

CSF 2.0 (National Institute of Standards and Technology, 2024) generalizes these ideas into a sector-agnostic taxonomy of cybersecurity outcomes structured around the functions Govern, Identify, Protect, Detect, Respond, Recover, and introduces "Profiles" and "Tiers" to describe target and current postures and the maturity of risk governance practices. The framework explicitly encourages organizations to treat cybersecurity alongside financial, privacy and supply-chain risks, and to use the CSF as a communication device between technical teams and senior management.

Taken together, these documents assume the existence – and provide guidance - for structured risk registers, continuous monitoring data and governance processes, and they provide vocabulary and process steps for integrating them. What they do not fully specify is how telemetry and standardized evidence sources (e.g., CTI, KEV, OSCAL assessments) should be mapped to reusable risk scenarios and simple, explainable metrics that can be understood and acted upon by non-specialist decision-makers (Problem 1 and 2). The link between continuous monitoring data to decision-quality, ERM-aligned risk views remains mostly manual.

3. Methodology

This paper adopts a structured literature review rather than a fully protocol-driven systematic review. The goal is not exhaustiveness, but to consolidate evidence around three governance problems introduced in Section 1: (P1) decision and alert overload; (P2) drift between periodic, interview-based risk estimates and operational reality; and (P3) limited practical support for compound and cascading cyber risks in governance processes.

3.1. Search strategy and corpus construction

The review combined three complementary source types: (i) normative and institutional documents that define governance vocabulary and expectations; (ii) academic and practitioner research on automation methods, models and tools; and (iii) practice-oriented, machine-readable artifacts (e.g., control catalogues, CTI knowledge bases) that enable evidence-as-code pipelines.

Corpus construction proceeded in three steps:

- **Baseline corpus**: we collected the current versions of major standards and institutional guidance referenced in Section 2 (ISO/IEC 27001/27002/27005, ISO 31000; ENISA interoperability and threat-landscape methodology; NIST risk assessment, ERM integration and continuous monitoring guidance; and related regulatory and incident-report sources).
- **Targeted database searches**: we queried IEEE Xplore, ACM Digital Library, Scopus and Google Scholar using problem-driven keyword sets (Table 1), iteratively refined to capture both governance-level automation and technical modeling work.
- **Snowballing**: we performed backward and forward snowballing from key contributions identified in steps (i)–(ii), with additional targeted searches on human-factors and systemic-risk terminology to avoid a tool-centric bias.

Corpus snapshot:

- Total references: 45
- Core mapped sources (Table 3): 42
- Time span: 1946-2025
- Since 2020: 36/45
- Source-type breakdown (mapped corpus): 42 mapped sources; of those 24 research/tooling, 5 normative/standard, 6 institutional guidance, 2 report/incident, 4 machine-readable artifacts, 1 foundational.

3.2. Screening and inclusion criteria

We screened sources by title/abstract and then full text. Inclusion required: (i) an explicit link to information/cybersecurity risk management or closely related ICT risk domains; (ii) substantive treatment of automation, formal modeling, or human-factor constraints; and (iii) relevance to at least one of P1–P3. Given the heterogeneity of sources (standards, institutional reports, empirical studies, modeling papers and tool descriptions), we did not apply PRISMA-style flow diagrams or formal quality-score checklists.

3.3. Coding, thematic clustering and synthesis

To support a problem-driven synthesis, each included contribution was coded along two orthogonal axes: (a) the governance problem(s) it primarily addresses (P1–P3), and (b) the thematic cluster(s) it contributes to (T1–T6). Clusters were defined to reflect recurring automation approaches in the corpus: T1 governance frameworks; T2 continuous monitoring and evidence-as-code; T3 process-level automation and toolkits; T4 data-driven risk estimation; T5 human factors; and T6 compound/systemic risk models. This dual coding makes overlaps explicit (e.g., continuous monitoring papers may also raise P1 concerns via alert proliferation) and allows a compact summary matrix of the reviewed literature (Tables 2-3).

Limitations. The search is time-bounded and English-language biased. The corpus also mixes heterogeneous source types (standards, institutional reports, tool descriptions and academic papers), so the synthesis is conceptual rather than statistical. The explicit P1–P3 and T1–T6 coding mitigates this heterogeneity by making the review lens transparent, but it does not eliminate selection bias.

Figure 1 can visually support the description of the 3 Problems.

Figure 1: Evidence-to-governance pipeline and where the three governance problems arise: P1 (decision/alert overload) at the decision surface, P2 (evidence-to-governance drift) at the scenario mapping and register alignment layer, and P3 (compound/cascading risk) when interdependencies must be represented beyond single-scenario views. Where no 'P' is indicated, the evidence is relevant for all the problems.

4. State-of-the-Art and related reviews

Several literature reviews already address adjacent pieces of the "automation in cyber risk assessment" space, but they typically do so through model-centric (e.g., Dynamic Risk Assessment), domain-centric (e.g., maritime/ship cybersecurity), or framework-modeling lenses. These works are valuable baselines for positioning our contribution, yet they do not use a problem-driven viewpoint (P1–P3) to explicitly connect (i) governance needs and constraints, (ii) evidence generation/collection, and (iii) automation mechanisms (evidence-as-code, continuous monitoring, and operational tooling) into a single interpretive frame.

A first close work is the systematic review by Cheimonidis et al. [22] on Dynamic Risk Assessment (DRA) in cybersecurity. The authors motivate DRA as a response to the limits of periodic, "static" risk assessment in rapidly changing threat landscapes, and review 50 DRA models, categorizing them by primary analysis method. They highlight that many DRA approaches are AI/ML-driven, with Bayesian networks being widely used; they also note that DRA models frequently ingest objective, real-time data (often IDS-derived signals) and vulnerability-related inputs. Importantly, they discuss limitations and open challenges, including limited integration of cyber threat intelligence (CTI) and a frequent lack of historical context, pointing to needs around data fusion and trust evaluation—especially in data-rich, zero-trust-like settings. While highly relevant to the continuous/dynamic evidence angle, this line of work tends to prioritize the modeling layer (how to compute risk from streams) more than the governance side (how evidence is packaged, explained, and operationally sustained under real organizational constraints).

Erbas et al. [23] propose an interesting literature review on threat modeling and risk assessment in ship (maritime) cybersecurity. They follow PRISMA and perform a structured search (August 2023) selecting 25 papers, and then build a taxonomy across methods, target environments, and threat types. Their findings are useful here for two reasons: (i) they explicitly analyze how threat modeling and risk assessment are combined in practice within a complex cyber-physical domain, and (ii) they highlight an "automation gap" in tool support, observing inconsistencies and generally weaker support for qualitative methods. However, the scope of their analysis is domain-specific, and the review does not aim to connect continuous evidence generation to broader multi-framework governance processes—so it is best used as a comparative anchor on what happens in high-stakes CPS-like domains, rather than a direct substitute for a cross-domain synthesis.

Finally, Mohd Amin et al. [24] conduct an SLR aimed at modeling a cyber risk assessment framework, using PRISMA-style screening and a three-stage process (planning, conducting, reporting). Their search spans six major databases and uses an exact-match strategy to reduce noise. The outcome is a framework organized around the assessment phases of identification, analysis, and evaluation, extracting variables/dimensions from the reviewed literature, while explicitly acknowledging adoption limits due to resource capability and noting the focus on assessment rather than treatment. This work is helpful in understanding the reasons behind the emergence of "standardization" pressures, but it remains somewhat abstract in terms of continuous monitoring, evidence-as-code, and automation pipelines.

In summary, existing reviews either (a) systematize dynamic/continuous risk computation models (DRA), (b) systematize domain-specific threat modeling and risk assessment methods (maritime), or (c) systematize risk assessment framework components (variables/dimensions). Our review complements these by organizing the body of contributions around the three concrete identified problems (P1–P3), and by explicitly linking governance frameworks, evidence generation, automation mechanisms, and human/organizational constraints into a single narrative that can support both research directions and implementation choices.

4.1. Automation in Cyber Risk Management

A first family of contributions looks at process-level automation across the full risk management lifecycle, rather than at individual detection tools. Sterbak et al. [25] analyse the subprocesses of

information security risk management—asset identification, risk assessment, treatment and monitoring—and conclude that most steps remain manual, time-consuming and heavily dependent on auditor expertise. They observe that organizations typically stitch together multiple heterogeneous tools, which partially automate individual subprocesses but introduce data inconsistencies, learning overhead and substantial manual work, with no comprehensive system covering the entire process. Automation, while present, doesn't effectively address P1 (decision overload) and P2 (drift between interviews and reality). Instead, it exacerbates the fragmentation of tools and data, hindering the delivery of a comprehensive, governance-level view of risk.

AlSadhan et al. [26] approach the problem from an Information Security Continuous Monitoring (ISCM) perspective. They explicitly frame security automation as "indispensable" to cope with the scale and dynamism of modern environments, yet show that major breaches still go undetected, and that even when technical alerts are available, organizations struggle to understand impact and to take timely, risk-based decisions. Their analysis echoes NIST SP 800-137's [27] observation that integrating diverse security products, normalizing data and aggregating metrics across tiers remains a major challenge: ISCM tools can collect more data more frequently, but governance-level users still face difficulties turning this information into actionable risk insights (P1 and P2).

A second cluster of works proposes integrated toolkits and knowledge-based platforms for automated cyber risk assessment. Gonzalez-Granadillo et al. [28] present the AMBIENT toolkit, which coordinates several modules (asset inventory, vulnerability assessment, privacy impact analysis, control evaluation) to automatically compute cyber and privacy risk metrics mapped to standards such as ISO/IEC 27001 and GDPR. While AMBIENT demonstrates that it is feasible to orchestrate multiple data sources and produce semi-automatic risk scores, it still assumes significant initial modeling and configuration effort, and its outputs remain largely metric-centric. It does not explicitly tackle scenario consolidation or human decision fatigue, so it offers only a partial answer to P1 and P2.

Phillips et al. [29] focus on cyber-physical systems (CPS) and propose a knowledge-based framework in which assets, vulnerabilities, threats and controls are represented in a structured model; automated reasoning then enumerates attack paths and evaluates risk based on formal relations rather than ad-hoc interviews. This approach supports more consistent and explainable risk assessments and starts to address compound and cascading behaviours—there is explicit representation of how one compromise propagates along system dependencies. However, the modeling cost is high, and the examples are mostly CPS-specific. Replicating detailed models across all services for a small to mid-large ICT enterprise would be prohibitive, so lighter, scenario-level abstractions are needed.

A third group of studies explores data-driven and predictive models that automate part of the risk estimation step. Kia et al. [30] build a cyber risk prediction model using Common Vulnerabilities and Exposures (CVEs) as the core signal: they cluster CVEs into topics, use time-series features and train machine-learning models (e.g., random forests) to forecast risk levels and alert on emerging vulnerability "themes". This line of work shows that textual and temporal patterns in vulnerability data can be exploited automatically, potentially reducing subjective biases in threat perception and helping organizations anticipate exposure trends, thus partially addressing P2 (alignment with observed exposure). At the same time, the approach operates at the level of global CVE streams, not at the level of enterprise-specific, governance-relevant scenarios; it does not by itself reduce the number of decision objects, nor does it embed control effectiveness or capacity constraints, so P1 and P3 remain largely untouched.

Quantitative cyber risk and insurance literature adds another angle. Carannante and Mazzoccoli's analytical review [31] shows how insurers and actuaries are experimenting with increasingly sophisticated models for cyber risk, distinguishing idiosyncratic, systematic and systemic components and highlighting contagion mechanisms and non-linear dependencies. The review stresses persistent data scarcity, lack of standardized incident datasets and difficulty modeling systemic events, and cautions against over-reliance on complex models with poorly validated parameters. This is directly relevant to P3: the field acknowledges that compound and systemic cyber risks exist and matter, but the proposed solutions are often too data-hungry and opaque to be realistically adopted as everyday governance tools in mid-sized enterprises.

Digital-twin research provides a complementary perspective on model-based automation. Zio and Miqueles [32] survey applications of digital twins in safety analysis, risk assessment and emergency management across multiple industries (nuclear, power grids, construction, manufacturing, transportation). They find that most digital-twin implementations focus on data collection and assimilation, often via IoT and wireless networks, and use the twin to compute safety indicators, support risk monitoring and perform prognostics and forecasting. The review also highlights substantial open challenges: data integration, model fidelity, interoperability and computational cost. Digital twins are emerging as powerful, automation-friendly environments to experiment with risk scenarios and emergency responses, but they are still mostly engineered for specific, high-criticality domains and require substantial modeling effort. This supports using a lighter "digital twin" of an ICT environment as a replication lab, not a replacement for governance-level risk analysis. It also confirms that scalable, explainable treatment of compound risk (P3) remains a research gap.

A separate strand of work deals with maturity models and ontologies, which are relevant for structuring automated assessments. Rabii et al. [33] survey information and cyber security maturity models developed between 2007 and 2018, documenting a large number of frameworks with heterogeneous concepts, scales and procedures and relatively little evidence of real-world implementation and effectiveness. Brezavšček and Baggia [34] extend this picture, showing that recent work tends to align with standards such as ISO/IEC 27001 and the NIS framework but that validation, sector specificity and integration with continuous monitoring remain weak points. In short, the literature confirms that maturity models are ubiquitous but often heavy, overlapping and only loosely connected to operational telemetry.

Ontological approaches attempt to harmonize terminology and support integrated, machine-interpretable risk management. Masso et al. [35] compile an extensive ontology for software risk management, integrating concepts from ISO 31000, ENISA guidance and other risk sources to provide a common vocabulary that can be used across tools and domains. Other ontologies target specific areas such as project risk, operational IT risk or IT-Grundschutz-based catalogs, and are sometimes embedded into tools that automatically generate risk reports or support decision-making. These works show that formalizing risk concepts and their relations is feasible and useful for automation, but they also tend to be domain-specific, with limited evidence of adoption in broad, multi-framework enterprise settings. For our purposes, they mostly justify treating "scenario", "control bundle", "indicator" and "chain" as explicit, reusable objects—without committing to heavyweight ontology engineering.

Taken together, these automation-oriented contributions reinforce the three structural identified problems.

- Regarding P1 (decision overload / alert fatigue): Automation has significantly increased the volume and granularity of available security data, but most approaches still surface rich technical metrics, dashboards and models that require expert interpretation. None of the surveyed works makes reduction of the decision surface (e.g., converging on a small, stable scenario set with few key indicators) its primary objective.
- Regarding P2 (drift between interview-based risk views and operational reality): ISCM and toolkit approaches clearly recognize the need to connect telemetry and governance, and they provide building blocks for continuous evidence collection and aggregation. However, they usually stop at the level of technical control status or vulnerability metrics, with limited attention to how this evidence should be folded back into the organization's scenario catalog, likelihood/impact estimates and treatment planning on a cadence. Alignment is discussed in principle, but only partially realized in practice.
- Regarding P3 (compound / chain risk under crisis conditions): Knowledge-based CPS models, digital-twin applications and cyber-insurance research all acknowledge interdependencies and contagion, and some provide sophisticated tools to model them. Yet these tools are often complex, data-intensive and tailored to specific sectors, which limits their transferability to a mid-large ICT enterprise with limited modeling capacity and strict governance constraints.

Several automation proposals converge on a semantic/ontology-driven pattern: they aim to make

risk-relevant evidence machine-readable, so that heterogeneous signals (assets, vulnerabilities, anomalies, threat intel, controls) can be correlated and reasoned upon through inference rules and shared vocabularies. In this vein, ontology-based risk frameworks model infrastructures and security facts in a formal knowledge base and use predefined rules to support threat modeling and risk estimation [36]. Complementary strands include: (i) reference ontologies that clarify security and risk-treatment concepts and their relations (useful to avoid semantic drift when mapping governance intent to technical measures) [37]; (ii) real-time risk and cyber-situational awareness approaches that combine anomaly and CTI sources with ontology reasoning/rules to infer dynamic risk levels and response options [38]; (iii) ontology-based security risk management architectures that ingest IDS-like outputs and support iterative assessment and control appraisal through a feedback loop [39]; and (iv) context-aware security measurement models that encode large sets of contextual attributes and derived metrics to compute granular-to-overall security effectiveness [40]. From the perspective of this review, these works primarily address P2 by reducing reliance on periodic interviews and ad-hoc correlation, pushing toward telemetry-informed, continuously updateable representations; they can also reduce P1 by externalizing parts of the analyst's reasoning into reusable rules/knowledge structures—although some proposals risk shifting cognitive load into ontology/metric complexity rather than truly simplifying decisions. Finally, practitioner-oriented knowledge bases such as MITRE D3FEND [41] can be framed as a pragmatic "defensive ontology" that helps bridge observed evidence to candidate countermeasures in a more standardized way (risk treatment lens), without requiring a full bespoke ontology stack.

4.2. Risk-based account recovery as micro-level automation

Büttner et al. [42] propose Risk-Based Account Recovery (RBAR) as an extension of risk-based authentication to the password-reset phase, motivated by the fact that recovery workflows are a frequent route to account takeover when recovery methods are weaker than primary authentication. RBAR computes a risk score from contextual and behavioral signals (e.g., IP, user agent, login history) and adapts the recovery flow accordingly: standard recovery for low risk, additional challenges for medium risk, and blocking for high risk—at the cost of potentially rejecting legitimate users.

The authors empirically test five major services (e.g., Amazon, Google) and find evidence of RBAR in only three, implemented heterogeneously (e.g., Google combining stronger factors and background checks; Amazon/LinkedIn relying mainly on CAPTCHA triggers). Based on these observations, they outline an RBAR maturity model ranging from no RBAR to stronger factor-based recovery, and highlight key limitations: implementations behave as black boxes, show inconsistent outcomes under similar conditions, and require many observations to learn a stable "normal" profile—reducing transparency and repeatability.

Relative to our problem lens, RBAR is a useful micro-level case study. For P1, it exemplifies selective escalation intended to reduce friction, but often degenerates into weak challenges that add user effort without proportional security gains. For P2, it shows how telemetry-driven decisions can still fail governance alignment when thresholds and evidence models are opaque and unstable. For P3, it illustrates how weaknesses in a single process (account recovery) can become a systemic failure point that enables downstream scenarios (identity compromise, fraud, or intrusion paths). Overall, the contribution is valuable less as a blueprint and more as empirical evidence that risk-driven automation without transparency and governance integration can remain brittle.

4.3. Human factors: decision overload, alert fatigue and cognitive limits

Classic human-factors research shows that decision-making does not scale linearly with the number of options. Hick's law and the Hick–Hyman formulation [43] model decision time as increasing with the logarithm of the number of alternatives, a pattern confirmed across multiple experimental tasks. In practical terms, enlarging the choice set beyond a modest size tends to slow decisions and increases error rates, especially when options are complex or poorly structured.

The "choice overload" literature refines this insight. Chernev et al. [44] synthesize evidence that

large assortments can reduce the likelihood of choosing and weaken preference strength, with the effect moderated by choice-set complexity, task difficulty and expertise. When options are many, similar and involve trade-offs on multiple attributes, non-expert decision-makers are more prone to deferral or satisficing. From a cyber-risk perspective, a dashboard with dozens of scenarios and hundreds of indicators is structurally similar to a complex assortment: it increases the cognitive effort needed to interpret relative priorities, and makes consistent, high-quality decisions less likely.

In security operations centers, these general decision-science results are mirrored by direct evidence of alert fatigue. Tariq et al. [45] review causes and mitigations of alert fatigue in SOCs, highlighting four main drivers: high false-positive rates, staff and skill shortages, poor prioritization mechanisms, and suboptimal human–tool interaction. They classify mitigation approaches into AI-assisted triage, augmentation and human-AI collaboration, but note that many solutions raise their own challenges (e.g., algorithmic bias, integration complexity, explainability) and that there is a lack of robust metrics for measuring human–AI collaboration effectiveness.

Taken together, these automation-oriented contributions reinforce the three structural problems identified in this paper: current frameworks and tools push organizations towards ever richer sets of controls, indicators and alerts, but they offer limited guidance on how to compress this complexity into a small number of governance-ready scenarios. More telemetry and more dashboards do not automatically lead to better risk decisions; without deliberate scenario aggregation and human-centered design, automation risks amplifying cognitive load.

For this review, the implication is straightforward: any credible proposal for "automation in cyber risk management" must be evaluated not only on its technical sophistication, but on the extent to which it reduces the number and complexity of decision objects presented to boards, risk committees and CISOs.

4.4. Compound and cascading cyber risk: models and limits

Cyber risks rarely materialize in isolation. Identity misuse enables ransomware; third-party outages propagate through supply chains; facility disruptions amplify the impact of cyber events. Research on cascading failures in interdependent systems provides formal tools to think about such chains. Flow–redistribution models show how failures in one network (e.g., power) can trigger overloads and subsequent failures in coupled networks (e.g., ICT, water), leading to non-linear cascades. These models highlight that dependencies and capacity constraints matter at least as much as individual component vulnerabilities.

Within cybersecurity, a substantial body of work uses Bayesian networks and attack graphs to represent conditional dependencies among vulnerabilities, attack steps and defenses. For example, [46] argues that Bayesian networks provide a causal probabilistic model capable of capturing interdependencies among cyber risk factors and of combining scarce empirical data with expert judgment to support more rigorous risk quantification and decision-making. Chen et al. [47] combine Bayesian networks with attack graphs to assess security of power systems, capturing how compromise can spread across components and how control measures affect overall risk.

These models speak directly to Problem 3 (P3): they make compound and cascading behaviors explicit and provide mathematically coherent ways to update beliefs as evidence arrives. However, they come with significant costs. They require detailed system models, relatively rich data to parameterize conditional probabilities, and specialized expertise to build and maintain the graphs. In typical mid-to-large ICT enterprises, where architectures are heterogeneous and constantly changing, maintaining such models at full enterprise scale is rarely feasible.

Quantitative cyber-risk and insurance literature provides another angle on compound risk. Reviews of mathematical models for cyber-insurance pricing and capital management emphasize systematic and systemic components, tail dependence and contagion across insured entities. While these works are important for macro-prudential supervision and product design, they usually rely on assumptions and datasets that are far removed from day-to-day governance in a single enterprise.

The overall picture is that research has rich tools for modeling dependencies, but most of them are

too opaque, data-hungry or maintenance-intensive to be integrated into the simple scenario-based risk registers promoted by ISO 27005, ISO 31000 or NIST ERM guidance. This gap is at the core of P3: organizations need ways to reason about co-occurrence and cascades that are simple enough for governance use yet grounded in evidence.

5. Gaps in current approaches

Taken together, ISO 31000, ISO/IEC 27001, ISO/IEC 27002, and ISO/IEC 27005 define a coherent governance backbone for cyber risk management:

- risk is central and iterative;
- controls are standardized and richly annotated;
- scenarios are a recognized way to structure analysis;
- continuous monitoring and automated detection are encouraged at control level.

However, synthesizing across standards, empirical studies and modeling work, three structural gaps emerge, corresponding to the problems motivating the paper.

5.1. Gap 1 – Decision overload and alert fatigue (P1)

Risk and security standards consistently encourage organisations to use the "best available information", maintain comprehensive control sets and monitor controls continuously. At the same time, decision-science and SOC research show that more information and more options can degrade decision quality when not properly structured.

Current frameworks stop short of addressing this tension. ISO 27001/27005 and ISO 31000 acknowledge uncertainty and bias, but they do not provide operational methods for compressing hundreds of controls, indicators and alerts into a small, stable set of scenarios with clear control implications. Continuous monitoring literature highlights the promises and challenges of risk scoring, yet most implementations still expose large sets of technical findings rather than a reduced, governance-level view. SOC-oriented work proposes AI-assisted triage and visual analytics, but often focuses on analyst-level workflows rather than board-level risk scenarios.

The net effect is that automation frequently increases cognitive load: more tools, more alerts, more dashboards – but no agreed method to fold them into scenario-based risk registers that a risk committee can actually use.

5.2. Gap 2 – Drift between interviews and operational reality (P2)

Standards and guidance documents assume that organizations will maintain structured risk registers, integrate them with continuous monitoring, and update them as conditions evolve. NIST SP 800-137 and NISTIR 8286 explicitly promote security continuous monitoring and cybersecurity risk registers that "roll up" into enterprise risk. ENISA's interoperability work and the NIS 2 Directive [9] push towards more consistent, reusable artefacts and more stringent supervision of risk management.

In practice, most organizations still rely on annual or semi-annual interview-based assessments, supported by spreadsheet-style registers and ad-hoc tooling. Continuous monitoring tooling is often deployed, but its outputs live in separate silos (SIEM, vulnerability scanners, IAM tools, CTI feeds) and are only loosely connected to the risk register. Empirical studies of ISCM adoption and automation possibilities confirm that integrating diverse products, normalizing data and aggregating metrics across tiers is a major challenge, and that agencies struggle to transition from periodic compliance checks to continuous risk-aware decision-making.

At the same time, structured artefacts such as MITRE ATT&CK, CISA KEV, OSCAL catalogs and machine-readable assessments demonstrate that the data building blocks already exist. The missing piece is a stable "glue" between these artefacts and scenario-based risk registers: a way to map changes in telemetry and CTI (e.g., KEV-tagged exposure, credential theft signals) into predictable adjustments

of scenario likelihoods and mitigation-tier indicators, without reverting to opaque black-box models. This misalignment explains Problem 2 (P2): even when telemetry is abundant, governance-level risk pictures lag behind, and the register quickly drifts from what actually happens in systems and networks.

5.3. Gap 3 – Limited, practical support for compound and cascading risk (P3)

Research on Bayesian networks, attack graphs, cascading failures and systemic cyber-risk provides sophisticated ways to represent dependencies and contagion. However, these methods are rarely embedded in mainstream risk-management practice in a way that an ISMS manager or enterprise risk officer can routinely use.

For example, a ransomware attack may occupy the entire IT staff (ill-prepared) in attempting to recover data from infected systems, leaving other threats unattended, such as phishing leading to the theft of privileged identities and the compromise of corporate data. The occurrence of one event increases the likelihood of others, and in some cases, their impacts.

On the governance side, ISO 27005 and ISO 31000 recognise that risks can interact, but treat this at a high level, leaving operational modeling to domain-specific methods. Quantitative cyber-insurance work confronts systemic risk but typically at portfolio or market level, under assumptions and datasets that are inaccessible to individual enterprises. Digital-twin research shows that it is feasible to mirror complex systems and compute safety indicators, but most examples are sector-specific and focus on engineering detail rather than governance dashboards.

The result is a methodological gap: organisations lack simple, explainable indicators for co-occurrence and cascades that can sit alongside traditional likelihood–impact scores. This leaves P3 largely unaddressed in the everyday practice of ISMS and ERM, despite clear evidence from incident reports (e.g., Verizon DBIR 2025) that chained events and third-party dependencies drive many high-impact breaches.

6. Future Directions

Given these gaps, promising research directions cluster around three axes aligned with P1–P3, plus a methodological strand.

6.1. Towards scenario-centered, human-aware automation (P1)

First, automation efforts should focus on scenario-level consolidation, not just on faster data collection. Human-factors evidence suggests that reducing the number of decision objects – while increasing their informativeness – is a precondition for better decisions under time and cognitive constraints. For risk governance, this implies:

- designing small, stable sets of macro-scenarios explicitly tied to business objectives and control families;
- defining a limited number of Key Control Indicators per scenario, kept consistent across tools;
- using automation to keep these indicators up-to-date from telemetry, rather than exposing raw alerts and findings.

In SOC and operations contexts, research on alert fatigue points to adaptive, human-AI collaboration as a way forward: AI-assisted triage, personalized alert presentation, and workload-aware task allocation that respect human capacity constraints. Translating these ideas "upwards" to governance suggests experimenting with role-specific risk views (e.g., board, CISO, system owner) and with metrics for decision effort, not just for incident counts.

A critical weakness in current work – and in many organizational roadmaps – is that "automation" is still often framed as "more data, more dashboards". Future research should make reduction of cognitive load an explicit design objective and measure the impact of automation on decision time and consistency.

6.2. Closing the telemetry - governance loop (P2)

For P2, the central challenge is building a maintainable pipeline between operational evidence and scenario-based risk registers. Existing standards already suggest the ingredients: structured risk registers (NISTIR 8286), continuous monitoring (SP 800-137), machine-readable controls and assessments (OSCAL and related work), and CTI artifacts (ATT&CK, KEV).

Future work could focus on:

- defining reusable mapping patterns from specific telemetry sources (vulnerability scanners, IAM logs, endpoint agents, CTI feeds) to scenario indicators and mitigation-tier changes;
- using simple, auditable rules or lightweight models to update scenario likelihoods based on changes in exposure (e.g., KEV-tagged vulnerabilities present on critical assets) or in control status, rather than relying on opaque ML systems;
- representing assessment and monitoring results in machine-readable formats that can be imported by GRC tools and audited over time, without exposing raw logs.

Predictive work on vulnerability trends (e.g.. Kia et al.'s CVE-based forecasting [30]) indicates that textual and temporal patterns can be exploited to anticipate exposure, but these models currently operate at global CVE-stream level. A realistic direction for mid-to-large enterprises is to use such signals as inputs to scenario-level rules, not as direct replacements for governance judgement.

The weakness here is that very few published systems demonstrate end-to-end alignment between telemetry and risk registers in real organizations. Case-based, longitudinal studies are missing.

6.3. Practical compound-risk indicators and digital-twin testbeds (P3)

For P3, an incremental path is preferable to "big bang" adoption of complex models. Research on cascading failures and Bayesian/attack-graph models can inform governance without being fully exposed at board level. Two complementary directions are promising:

1. Simple co-occurrence indicators at risk-register level, inspired by association-rule mining (support, confidence, lift). Incident and risk-register histories can be mined to identify scenario pairs or triples that frequently co-occur or amplify each other (e.g., "identity compromise + ransomware during data-center incident"). Even basic statistics could highlight combinations that deserve joint monitoring and playbook-level preparation.
2. Deeper models reserved for analysis and simulation, not for everyday dashboards. Bayesian networks and attack graphs can be used selectively for high-risk systems (e.g. core payment platforms, industrial control segments) to explore "what-if" scenarios and stress-test controls. Digital twins provide a natural testbed: a mirrored environment where telemetry and control changes can be experimented with before being reflected in governance metrics.

The key is to separate modeling layers: simple indicators and a small number of compound scenarios for governance; richer models underneath for expert analysis. This helps avoid the common failure mode where organizations either ignore dependencies altogether or adopt models that are too complex to maintain.

6.4. Methodological implications: Action Design Research for automation-assisted governance

Finally, there is a methodological gap. Most works on automation, continuous monitoring and compound risk are either conceptual (frameworks), tool-centric or based on simulations detached from organizational realities. To close this, the literature on Action Research [48] and Action Design Research (ADR) [49, 50] is relevant.

ADR explicitly combines design of artefacts (methods, tools, models) with interventions in organizations and iterative evaluation. Cedergren and Hassel, for example, use ADR to develop and implement

an integrated method for risk assessment and continuity management, refining the method through cycles of application and feedback. This approach is well-suited to automation-assisted cyber risk governance, where:

- artifacts (e.g., scenario taxonomies, mapping rules, dashboards) must fit local constraints and existing frameworks;
- organizational adoption, not just technical feasibility, is critical;
- insights need to be generalized beyond a single case.

Future research on automation in cyber risk management would therefore benefit from multi-year ADR programmes in real organizations, where scenario-based dashboards, telemetry-to-risk mappings and compound-risk indicators are designed, deployed, measured and refined. The risk, if this is neglected, is that proposals remain at the level of conceptual frameworks or isolated tools with little impact on governance practice.

7. Conclusion

This review started from three problems observed in the practice of cyber risk management in mid-to-large ICT environments: P1 – decision and alert overload for human decision-makers; P2 – drift between periodic, interview-based risk assessments and operational reality; P3 – limited support for compound and cascading cyber risks in governance processes.

The analysis of standards, empirical incident reports and academic work shows a mixed picture. On the one hand, frameworks such as ISO 27001/27005, ISO 31000, NIST CSF and NIST IR 8286 provide mature structures for risk-based governance and encourage continuous monitoring and ERM integration. Toolkits, digital-twin platforms, Bayesian/attack-graph models and quantitative cyber-risk models demonstrate that rich automation and dependency modelling are technically feasible.

On the other hand, the last mile from telemetry and models to governance remains weak. Automation often increases the number of alerts and metrics without reducing cognitive load; risk registers are rarely updated in lockstep with operational evidence; compound and systemic risks are acknowledged in theory but only partially represented in everyday tools. Human-factor research and SOC studies underline that cognitive bandwidth is a hard constraint, not an afterthought, and that ignoring it undermines the value of automation.

The contribution of this paper is to structure the landscape around the three problems, identify where existing work already offers useful building blocks, and highlight where evidence and methods are missing. It deliberately stops short of prescribing a specific technical solution; instead, it outlines directions for scenario-centered, human-aware and evidence-driven automation, and suggests ADR as a suitable methodological lens for future projects.

From a forward-looking perspective, the main risk in our reasoning would be to overestimate how much can be achieved by tooling alone. The literature consistently shows that automation needs to be co-designed with governance processes, cognitive constraints and organizational incentives. The next step, beyond this review, is therefore not another model, but carefully designed field work: piloting lightweight automation loops in real organizations, measuring their impact on decision quality and workload, and iterating until risk governance becomes both more informed and less exhausting.

Acknowledgments

This work was partially supported by the project SERICS (PE00000014) under the NRRP MUR program funded by the EU - NGEU.

Declaration on Generative AI

During the preparation of this work, the author(s) used GPT-5 and Grammarly solely for grammar checking, spelling correction, and phrasing refinement to enhance readability. After using these tool(s)/service(s), the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

References

[1] R. Ross, V. Pillitteri, K. Dempsey, M. Riddle, G. Guissanie, Protecting controlled unclassified information in nonfederal systems and organizations, Technical Report, National Institute of Standards and Technology, 2019.

[2] E. Bergström, C. Welch, A. Nolte, M. Rajanen, A. Islind, H. Hult, A. Ravarini, Tool supporting information security risk management in practice., in: STPIS, 2023, pp. 146–159.

[3] SANS Institute, Siem fatigue isn't a technology problem, it's a strategy problem, Available at https://www.sans.org/blog/siem-fatigue-technology-problem-strategy-problem, 2025. Accessed: 2025-12-17.

[4] N. Rastogi, S. Pant, D. Dhanuka, A. Saxena, P. Mairal, Too much to trust? measuring the security and cognitive impacts of explainability in ai-driven socs, arXiv preprint arXiv:2503.02065 (2025).

[5] ENISA, ENISA Cybersecurity Threat Landscape Methodology, Available at https://www.enisa.europa.eu/publications/enisa-cybersecurity-threat-landscape-methodology, 2025. Accessed: 2025-12-17.

[6] Google, M-Trends 2025 Report, Available at https://cloud.google.com/security/resources/m-trends, 2025. Accessed: 2025-12-17.

[7] Microsoft, Microsoft Digital Defense Report 2025, Available at https://www.microsoft.com/en-us/corporate-responsibility/dmc/en-us/corporate-responsibility/cybersecurity/microsoft-digital-defense-report-2025/, 2025. Accessed: 2025-12-17.

[8] Verizon, 2025 Data Breach Investigations Report, Available at https://www.verizon.com/business/resources/reports/dbir/, 2025. Accessed: 2025-12-17.

[9] ENISA, Executive Order on Strengthening the Cybersecurity of Federal Networks and Critical Infrastructure, Available at https://www.enisa.europa.eu/topics/cybersecurity-policy/nis-directive-new, 2023. Accessed: 2025-12-17.

[10] Risk management — Guidelines, Standard ISO 31000:2018, International Organization for Standardization, Geneva, CH, 2018. URL: https://www.iso.org/standard/65694.html.

[11] Information security, cybersecurity and privacy protection — Information security management systems — Requirements, Standard ISO/IEC 27001:2022, International Organization for Standardization, Geneva, CH, 2022. URL: https://www.iso.org/standard/27001.

[12] Information security, cybersecurity and privacy protection — Information security controls, Standard ISO/IEC 27002:2022, International Organization for Standardization, Geneva, CH, 2022. URL: https://www.iso.org/standard/75652.html.

[13] Information security, cybersecurity and privacy protection — Guidance on managing information security risks, Standard ISO/IEC 27005:2022, International Organization for Standardization, Geneva, CH, 2022. URL: https://www.iso.org/standard/80585.html.

[14] Interoperable EU Risk Management Framework, Standard, European Union Agency for Cybersecurity, Attiki, GRC, 2022. URL: https://www.enisa.europa.eu/publications/interoperable-eu-risk-management-framework.

[15] Guide for Conducting Risk Assessments, Standard SP 800-30 Rev. 1, National Institute of Standards and Technology, Gaithersburg, USA, 2012. URL: https://csrc.nist.gov/pubs/sp/800/30/r1/final.

[16] Managing Information Security Risk: Organization, Mission, and Information System View, Standard SP 800-39, National Institute of Standards and Technology, Gaithersburg, USA, 2011. URL: https://csrc.nist.gov/pubs/sp/800/39/final.

[17] B. E. Strom, A. Applebaum, D. P. Miller, K. C. Nickels, A. G. Pennington, C. B. Thomas, Mitre att&ck: Design and philosophy, Technical report (2018).
[18] America's Cyber Defence Agency, Known exploited vulnerabilities catalog, Available at https://www.cisa.gov/known-exploited-vulnerabilities-catalog, 2025. Accessed: 2025-12-17.
[19] N. I. of Standard, Technology, Open Security Controls Assessment Language, Available at https://csrc.nist.gov/Projects/Open-Security-Controls-Assessment-Language, 2023. Accessed: 2025-12-17.
[20] K. L. Dempsey, L. A. Johnson, M. A. Scholl, K. M. Stine, A. C. Jones, A. Orebaugh, N. S. Chawla, R. Johnston, Information security continuous monitoring (iscm) for federal information systems and organizations (2011).
[21] K. Stine, K. Stine, S. Quinn, G. Witte, R. Gardner, Integrating cybersecurity and enterprise risk management (ERM), volume 10, US Department of Commerce, National Institute of Standards and Technology ..., 2020.
[22] P. Cheimonidis, K. Rantos, Dynamic risk assessment in cybersecurity: A systematic literature review, Future Internet 15 (2023) 324.
[23] M. Erbas, S. M. Khalil, L. Tsiopoulos, Systematic literature review of threat modeling and risk assessment in ship cybersecurity, Ocean Engineering 306 (2024) 118059.
[24] Z. M. Amin, N. Anwar, M. S. M. Shoid, S. Samuri, A systematic literature review for modeling a cyber risk assessment framework, Environment-Behaviour Proceedings Journal 9 (2024) 189–195.
[25] M. Sterbak, P. Segec, J. Jurc, Automation of risk management processes, in: 2021 19th International Conference on Emerging eLearning Technologies and Applications (ICETA), IEEE, 2021, pp. 381–386.
[26] T. AlSadhan, J. S. Park, Enhancing risk-based decisions by leveraging cyber security automation, in: 2016 European Intelligence and Security Informatics Conference (EISIC), IEEE, 2016, pp. 164–167.
[27] Information Security Continuous Monitoring (ISCM) for Federal Information Systems and Organizations, Standard SP 800-137, National Institute of Standards and Technology, Gaithersburg, USA, 2011. URL: https://csrc.nist.gov/pubs/sp/800/137/final.
[28] G. Gonzalez-Granadillo, S. A. Menesidou, D. Papamartzivanos, R. Romeu, D. Navarro-Llobet, C. Okoh, S. Nifakos, C. Xenakis, E. Panaousis, Automated cyber and privacy risk management toolkit, Sensors 21 (2021) 5493.
[29] S. C. Phillips, S. Taylor, M. Boniface, S. Modafferi, M. Surridge, Automated knowledge-based cybersecurity risk assessment of cyber-physical systems, IEEE Access 12 (2024) 82482–82505.
[30] A. N. Kia, F. Murphy, B. Sheehan, D. Shannon, A cyber risk prediction model using common vulnerabilities and exposures, Expert Systems with Applications 237 (2024) 121599.
[31] M. Carannante, A. Mazzoccoli, An analytical review of cyber risk management by insurance companies: A mathematical perspective, Risks 13 (2025) 144.
[32] E. Zio, L. Miqueles, Digital twins in safety analysis, risk assessment and emergency management, Reliability Engineering & System Safety 246 (2024) 110040.
[33] A. Rabii, S. Assoul, K. Ouazzani Touhami, O. Roudies, Information and cyber security maturity models: a systematic literature review, Information & Computer Security 28 (2020) 627–644.
[34] A. Brezavšček, A. Baggia, Recent trends in information and cyber security maturity assessment: A systematic literature review, Systems 13 (2025) 52.
[35] J. Masso, F. García, C. Pardo, F. J. Pino, M. Piattini, A common terminology for software risk management, ACM Transactions on Software Engineering and Methodology (TOSEM) 31 (2022) 1–47.
[36] N. Maunero, F. De Rosa, P. Prinetto, Towards cybersecurity risk assessment automation: an ontological approach, in: 2023 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech), 2023, pp. 0628–0635. doi:10.1109/DASC/PiCom/CBDCom/Cy59711.2023.10361456.
[37] Í. Oliveira, T. P. Sales, R. Baratella, M. Fumagalli, G. Guizzardi, An ontology of security from a risk treatment perspective, in: International conference on conceptual modeling, Springer, 2022, pp.

365–379.
[38] C. Sánchez-Zas, V. A. Villagrá, M. Vega-Barbas, X. Larriva-Novo, J. I. Moreno, J. Berrocal, Ontology-based approach to real-time risk management and cyber-situational awareness, Future Generation Computer Systems 141 (2023) 462–472.
[39] O. T. Arogundade, A. Abayomi-Alli, S. Misra, An ontology-based security risk management model for information systems, Arabian Journal for Science and Engineering 45 (2020) 6183–6198.
[40] M. Khaleghi, M. R. Aref, M. Rasti, Context-aware ontology-based security measurement model, Journal of Information Security and Applications 67 (2022) 103199.
[41] P. Kaloroumakis, M. Smith, Toward a knowledge graph of cybersecurity countermeasures (2020).
[42] A. Büttner, A. T. Pedersen, S. Wiefling, N. Gruschka, L. Lo Iacono, Is it really you who forgot the password? when account recovery meets risk-based authentication, in: International Conference on Ubiquitous Security, Springer, 2023, pp. 401–419.
[43] W. E. Hick, On the rate of gain of information, Quarterly Journal of experimental psychology 4 (1952) 11–26.
[44] A. Chernev, U. Böckenholt, J. Goodman, Choice overload: A conceptual review and meta-analysis, Journal of Consumer Psychology 25 (2015) 333–358.
[45] S. Tariq, M. Baruwal Chhetri, S. Nepal, C. Paris, Alert fatigue in security operations centres: Research challenges and opportunities, ACM Computing Surveys 57 (2025) 1–38.
[46] V. Ramakrishnan, Cyberrisk assessment using bayesian networks, ISACA J 5 (2016).
[47] L. Chen, T. Zhang, Y. Ma, Y. Li, C. Wang, C. He, Z. Lv, N. Li, A bayesian-attack-graph-based security assessment method for power systems, Electronics 13 (2024) 2628.
[48] K. Lewin, Action research and minority problems, Journal of social issues 2 (1946) 34–46.
[49] R. L. Baskerville, Investigating information systems with action research, Communications of the association for information systems 2 (1999) 19.
[50] A. Cedergren, H. Hassel, Using action design research for developing and implementing a method for risk assessment and continuity management, Safety science 151 (2022) 105727.
[51] The NIST Cybersecurity Framework (CSF) 2.0, Standard, National Institute of Standards and Technology, Gaithersburg, USA, 2024. URL: https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.29.pdf.

A. Appendices

Table 1
Problem-driven keyword sets used in targeted searches (illustrative, not exhaustive).

Problem lens	Query focus	Example keywords/phrases	Typical sources surfaced
P1 – Overload and human factors	Decision/alert overload, prioritization, cognitive constraints	alert fatigue; decision fatigue; information overload; choice overload; cognitive load; SOC triage; prioritization; human-AI collaboration; resource capability	SOC studies; decision science; human factors in security operations; usability of risk dashboards
P2 – Evidence-to-governance drift	Continuous monitoring, evidence pipelines, risk registers and ERM linkage	continuous monitoring; ISCM; continuous controls monitoring; evidence-as-code; telemetry-to-risk; risk register update; control effectiveness; GRC automation; OSCAL; policy-as-code; compliance-as-code; dynamic risk assessment	ISCM and control-monitoring literature; governance toolkits; standards guidance on monitoring; machine-readable control/assessment artefacts
P3 – Compound and cascading cyber risk	Dependencies, cascades, systemic/-portfolio risk and interdependent infrastructures	compound risk; cascading failures; systemic risk; interdependencies; attack graphs; Bayesian networks; contagion; tail dependence; supply-chain propagation; risk propagation	Attack-graph/Bayesian modelling; interdependent-system cascade models; cyber-insurance/systemic-risk modelling; supply-chain risk studies

Table 2
Thematic clusters used for synthesis and their relationship to problems P1–P3.

Cluster	Scope	Primary problem linkage	Examples in this paper
T1 – Governance frameworks	Standards, frameworks and regulatory/institutional guidance that define risk processes, vocabularies and governance expectations.	Mostly P2; touches P1 and P3 when complexity and risk combinations are acknowledged.	ISO/IEC 27001/27002/27005; ISO 31000; NIST CSF; NISTIR 8286; NIS 2; ENISA interoperability work
T2 – Continuous monitoring and evidence-as-code	Methods and artifacts that operationalise continuous control monitoring and machine-readable evidence flows.	Primarily P2; secondary P1 (alert/metric volume).	NIST SP 800-137; ISCM work; OSCAL; CISA KEV; MITRE ATT&CK; ENISA CTL methodology
T3 – Process-level automation and toolkits	Integrated platforms, pipelines and toolkits that automate parts of the risk lifecycle beyond individual detection tools.	Primarily P2; may mitigate or amplify P1 depending on decision-surface design.	Sterbak et al.; AMBIENT; knowledge-based frameworks; digital-twin platforms; ontology-based toolchains
T4 – Data-driven risk estimation	Quantitative and predictive models that automate likelihood/impact estimation or scoring from data.	P2 and P3; may create P1 issues if outputs are opaque or too granular.	Kia et al.; quantitative cyber-risk/insurance models; Bayesian updating approaches
T5 – Human factors	Empirical and theoretical work on decision limits, overload and alert fatigue relevant to risk governance design.	Primarily P1.	Hick–Hyman; choice overload (Chernev et al.); SOC alert fatigue (Tariq et al.)
T6 – Compound/systemic risk models	Formal models of dependencies, cascades and systemic behaviours across components or organisations.	Primarily P3; also supports P2 when evidence updates are modelled.	Cascading-failure models; attack graphs; Bayesian networks (Ramakrishnan et al.; Chen et al.); digital-twin dependency studies

Table 3
Mapping of cited contributions to thematic clusters (T1–T6) and governance problems (P1–P3). ■ = $substantive focus$;Δ = acknowledged/partial; — = not a focus.

Source	Type	Cluster(s)	P1	P2	P3
[11]	Standard	T1	–	■	–
[12]	Standard	T1, T2	Δ	■	—
[13]	Standard	T1	Δ	■	—
[10]	Standard	T1	Δ	■	Δ
[14]	Report	T1	—	■	Δ
[5]	Methodology	T2	■	Δ	—
[9]	Regulation	T1	—	■	—
[15]	Guidance	T1	—	■	Δ
[?]	Guidance	T1	—	■	Δ
[27]	Guidance	T2	Δ	■	—
[21]	Guidance	T1	—	■	Δ
[51]	Framework	T1	—	■	Δ
[17]	Knowledge base	T2	—	■	—
[18]	Catalogue	T2	—	■	—
[19]	Data format	T2	—	■	—
[25]	Paper	T3	Δ	■	—
[26]	Paper	T2	Δ	■	—
[28]	Paper/tool	T2, T3	—	■	—
[29]	Paper	T3, T6	—	■	■
[30]	Paper	T4	—	■	Δ
[31]	Survey	T4, T6	—	Δ	■
[32]	Survey	T3, T6	—	■	■
[33]	Survey	T1, T3	Δ	Δ	—
[34]	Survey	T1, T3	Δ	Δ	—
[35]	Paper	T3	—	■	—
[36]	Paper	T3	—	■	Δ
[41]	Knowledge base	T2	—	■	—
[37]	Paper	T3 Δ	■	Δ	
[40]	Paper	T2, T4	Δ	■	Δ
[38]	Paper	T2, T3	Δ	■	Δ
[39]	Paper	T3	Δ	■	—
[42]	Paper	T4, T5	Δ	■	Δ
[43]	Foundational	T5	■	—	—
[44]	Survey	T5	■	—	—
[45]	Survey	T5	■	—	—
[46]	Paper	T4, T6	—	Δ	■
[47]	Paper	T6	—	Δ	■
[50]	Paper	T3	—	Δ	—
[22]	Review	T2, T4	—	■	Δ
[23]	Review	T3, T4	—	Δ	Δ
[24]	Review	T1, T3	Δ	—	—