



(RESEARCH ARTICLE)



GOVERNING CROSS-DOMAIN RISK IN AI-DRIVEN CRITICAL INFRASTRUCTURE

Vladimir Bunic *

Converged Security Institute, Barcelona, Spain.

* Corresponding Author

ORCID Details

Vladimir Bunic : <https://orcid.org/0009-0006-6209-629X>

International Journal of Science and Research Archive, 2026, 21(01), 296–302

Publication history: Received on 03 September 2026; revised on 10 October 2026; accepted on 12 October 2026

Article DOI: <https://doi.org/10.30574/ijrsra.2026.21.1.1877>

Copyright © 2026 Author(s) retain the copyright of this article. This article is published under the terms of the Creative Commons Attribution License 4.0

ABSTRACT

Modern critical infrastructures (CI) have become a socio technical system, in their modern attributes, but with a distance that is becoming ever decreasing. We have entered a new era of socio technical systems in which "formation of the digital, physical, industrial and human domains" has come together in a new way to make our CI in the modern day. AI, digital twins, Internet of Things (IoT) ecosystems and autonomous operational technology (OT) are now actively involved in decisions, leading to new instances of cross domain risk. Incidents that begin in one domain may expand to involve other domains and have system impacts outside of the traditional security perimeter. This paper seeks to analyse and describe the converged risk phenomenon by conducting a structured multi scenario analysis and showcase a single Governance framework for a Security Operations Centre (SOC) and Global Security Operations Centre (GSOC). The methodology uses the qualitative scenario reconstruction approach to capture propagation pathways with telemetry corruption, digital twin drift, autonomy escalation, autonomous error-classification or misclassification by AI, authentication instability and multi sector cascading failures. Results of this show that converged incidents need governance mechanisms that are able to bring together the following four areas: telemetry validation, autonomy monitoring, prediction model alignment and systemic risk monitoring. The findings suggest change in the manner SOC and GSOC leaders should develop: Incident response to Governor of socio technical stability. This is supported by the Converged Security Framework, which provides the structural basis for this evolution for responsible application of AI and critical infrastructure.

Keywords: Converged Security; Critical Infrastructure; Digital Twins; Systemic Risk; SOC; GSOC; Autonomy Governance; Cyber-Physical Systems.

1 INTRODUCTION

The incorporation of artificial intelligence (AI), digital twins, Internet of Things (IoT) ecosystems and autonomous operational technology (OT) has fundamentally transformed modern critical infrastructure. The technologies have erased the previous barriers between digital networks, physical mechanisms, industrial control systems and human operators. These now constitute "tightly converged" systems in which digital and physical processes interact,

continually influencing each other - a characteristic noted in contemporary research on cyber physical systems (Lee et al., 2015).

As reliability on AI for decision making grows, it gives rise to novel vulnerabilities in the ecosystem. ENISA (2023) highlights three major risks that can be cascaded among domains, namely: a telemetry corruption, model drift and autonomy escalation. The risks are increased because today's infrastructure is tightly converged, meaning having a problem and implementing countermeasures in one component might result in problems and countermeasures in other components.

This paper aims to explore the converged risk in a structured multi scenario analysis and present the Converged Security Framework, as a new global governance model for Security Operations Centres (SOC) and Global Security Operations Centres (GSOC). The author argues that SOC and GSOC leaders need to grow into governors of systemic risk; people who are able to sense, understand and keep stable the complex and unpredictable interactions between various domains and between humans and socio-technical systems.

2 METHODOLOGY

The author uses a qualitative, scenario based analytical methodology to analyse converged risk and its application to AI driven critical infrastructure. Scenario analysis is a known method to understand complex socio technical interactions, especially when digital, physical, industrial and human systems are interdependent layers. This is in line with the cutting-edge cyber physical systems research that stresses cross domain propagation pathways (Lee et al., 2015).

There are three parts to the methodology.

2.1 Cross domain incident reconstruction

All scenarios simulate a situation starting from a single failure event – corrupted telemetry, drift of digital twins, autonomy escalation, classification error of AI, instability in authentication or cascading failure across multiple sectors. The reconstruction is followed in digital, physical, industrial and human domain to reflect propagation pathways through the incident.

2.2 Governance pathway identification

The author uses a suitable governance approach to stabilise the system for each scenario. These are actions such as Telemetry Validation, Autonomy Override, Digital Twin Synchronisation, Human Machine Teaming, and Systemic Risk Containment. This is in line with ENISA's (2023) recommendation for the governance of AI enabled systems to be capable of responding to systemic vulnerabilities.

2.3 Framework alignment

The scenarios are mapped to the structural elements of the Converged Security Framework. This mapping shows how the framework offers a common operational model that can be used for cross domain incidents.

This methodology can be used to fulfil the purpose of this paper and details the need for governance in a SOC/GSOC with AI-driven operations.

3 MULTI-SCENARIO ANALYSIS

3.1 SCENARIO I: TELEMETRY CORRUPTION IN A SMART ENERGY GRID

In the first case, a cluster of IoT sensors, spread throughout an energy network, start to become inconsistent. While these discrepancies may seem minor, they add up to create distortions in the data flowing through the grid's AI-powered anomaly detection systems over time. Corrective load balancing actions are created by autonomous controllers according to the corrupted telemetry data without any approval from humans. This generates variations in voltage which cause physical access devices at remote substations to reboot unexpectedly.

When trying to get back on top of the situation, operators are faced with dashboard delays and lack of integration. This is a scenario in which digital irregularities can spiral out of control in several domains. One of the leading causes of AI misclassification and escalation in autonomy is corrupted telemetry, according to ENISA (2023).

- **Governance requirements** - Isolation of corrupted telemetry sources, returning autonomous control units to human supervision, realignment of digital twins with validated data, and stabilisation of physical access under controlled conditions.
- **Framework alignment** - This scenario turns on the Cross Domain Telemetry and Validation Layer, AI and Digital Twin Analytics Layer and Autonomous Decision and Control Layer.

3.2 SCENARIO II: DIGITAL TWIN DRIFT IN WATER TREATMENT OPERATIONS

The second case is when a water treatment plant has a progressive discrepancy between the digital twin and actual operational conditions. The twin starts to forecast incorrect dosing of chemicals no longer consistent with the data from the sensor. Setting chemicals with autonomous dosing systems, responding to twin projections, cause chemical levels to be set incorrectly. The condition of the water changes and operators' step in when they detect water treatment disparities between the proposed and the actual.

Digital twin drift is a known issue in cyber physical systems, especially in the absence of a synchronisation process (ISO/IEC 30147:2022, 2022). This scenario shows how the misalignment of the predictive models can cause operational risks.

- **Governance requirements** - Resetting digital twin, suspending autonomous dosing, validating all sensor inputs, restoring human led control.
- **Framework alignment** - This scenario turns on the AI and Digital Twin Analytics Layer, Autonomous Decision and Control Layer, and Human Machine Teaming Layer.

3.3 SCENARIO III: AUTONOMOUS OT ESCALATION IN TRANSPORT INFRASTRUCTURE

The third scenario plays in the field of a transport network in which an autonomous inspection robot uncovers anomalies within the structure of a rail tunnel. The robot then, acting within its pre-programmed limits, raises the problem to the OT controllers, whereupon the ventilation and lighting systems shut down automatically. The shutdown has impacts on physical access, environmental stability and impact to digital monitoring systems.

An action to correct is taken but fails due to a conflict of autonomy. The National Academies (2019) note that, in complex environments, autonomy without boundaries that are governed may lead to safety risks.

- **Governance requirements** - Overriding autonomous actions, manual restoration of environmental controls, re-establishment of digital monitoring and validation of the robot's anomalous detection logic.
- **Framework alignment** - This activates the Autonomous Decision and Control Layer, the Human Machine Teaming Layer and the Systemic Risk Monitoring and Response Layer.

3.4 SCENARIO IV: AI MISCLASSIFICATION IN HEALTHCARE DIAGNOSTICS

The fourth scenario examines an AI-assisted diagnostic system operating within a healthcare environment that experiences a misclassification event. The problem lies with medical Internet of Things (IoT) transmittal of corrupted telemetry information. Patient state digital twins start to forecast inaccurate treatment pathways. These projections are used to provide optimised dosages of the medication from an autonomous dispenser and therefore pose a direct compromise to the safety of patients. Clinicians only intervene once they recognise that there is a discrepancy between AI and clinical observations.

This scenario is representative of how digital instability can directly have an impact on human wellbeing. AI misclassification has been reported in other scenarios where telemetry integrity is at risk, especially when autonomous systems rely on unvalidated data in their operations (ENISA, 2023). The propagation pathway, from corrupted telemetry and digital twin drift to autonomous dosing, demonstrates the systemic nature of converged incidents.

- **Governance requirements** - Suspension of autonomous medication dispensing; validation of all telemetry sources; resetting and re synchronising digital twin models; human led diagnostic review.
- **Framework alignment** - This scenario enables the following layers: Cross Domain Telemetry and Validation Layer, AI and Digital Twin Analytics Layer, and the Human Machine Teaming Layer.

3.5 SCENARIO V: PHYSICAL SECURITY FAILURE TRIGGERED BY CYBER INSTABILITY

The fifth situation focuses on instability in an authentication server that is in charge of the authorization of people who are using badges for access control on a site. Badge readers go bad at the same time, which puts devices in the physical access control system into fail open. Unauthorised movement is detected by operational technology (OT) systems and

automatically triggers lockdown procedures. Staff are unable to move about safely within the facility, digital, physical and industrial systems are conflicting.

In this situation, a digital authentication system has failed, and physical security has come into crisis. Failure to achieve interoperability in the systems in which IoT has been introduced can lead to cascading failures in dependent domains (ISO/IEC 30147:2022,2022). The incident exemplifies the physical and industrial implications which the cyber instability can elicit and it also illustrates wherein humans' safety is at risk.

- **Governance requirements** - Authentication server stabilisation, physical access systems resetting, automatic lockdown systems overriding and doing a cross-domain movement audit.
- **Framework alignment** - In this scenario, the Systemic Risk Monitoring and Response Layer, the Cross Domain Telemetry and Validation Layer and the Architectural Integration Layer are activated.

3.6 SCENARIO VI: MULTI-SECTOR CASCADING FAILURE ACROSS ENERGY, WATER, AND TRANSPORT

The last scenario is the most complicated converged risk. This "synchronised failure cascade" starts in the energy grid, because of a corrupted telemetry. Compensatory efforts by digital twins in water and transport are made with inaccurate predictions. In all three sectors, the actions are escalated at the same time from the autonomous controllers, resulting in a multi sector disruption which overwhelms the operators and destabilises critical operators.

In this example, systemic risk of converged environments is fully illustrated. Complicated interdependencies between sectors make for cascading failures, especially when AI-driven systems rely on shared or correlated telemetry (World Economic Forum, 2021). The propagation pathway shows that an instability in one sector can cause complementary actions in other sectors that tend to exacerbate the instability.

- **Governance requirements** - Tripping cross sector containment measures; implementation of human control in all autonomous areas; resetting and synchronising digital twins throughout domains; introduction of unified cross domain monitoring.
- **Framework alignment** - The scenario turns on all levels of the Converged Security Framework, showing how the need for a one governance entity is critical to handling multi sector systemic risk.

3.7 THE CONVERGED SECURITY FRAMEWORK

As described in the previous sections, the scenarios highlight the complexity and interdependence of modern CIs. They also point to the need for a single practical model which can regulate digital, physical, industrial and human spaces as a single socio technical system. This model is available in the Converged Security Framework. It sets in motion a structured architecture, which empowers the Security Operations Centre (SOC) and the Global Security Operations Centre (GSOC) to predict, identify and manage cross domain incidents clearly and effectively.

The six layers are dependent on each other in this framework. Though each layer defines a unique governance requirement in converged environments, all layers are integrated into a single system. It is designed to follow the current developments in the field of cyber physical systems (CP systems) research (Lee et al., 2015), in research on AI governing bodies (ENISA, 2023) and in research on the interoperability of IoT systems (ISO/IEC 30147:2022, 2022).

3.8 ARCHITECTURAL INTEGRATION LAYER

This layer provides the basic infrastructure on which E-Systems based on information technology (IT), operational technology (OT), Internet of Things (IoT), Physical security systems, and Human machine interfaces (HMIs) are integrated into an understandable operational model. Such domains cannot be independent domains in converged environments; they need to be known and managed as part of one system.

It maintains a seamless governance structure for the operation of telemetry, control logic and human oversight. It offers the building blocks needed for cross domain situational awareness, which allows for the interpretation of what's happening beyond an incident rather than through various operational silos in the SOC and GSOC.

3.9 CROSS-DOMAIN TELEMETRY AND VALIDATION LAYER

Telemetry is the building block of AI models, Digital twins, Autonomous controllers, and Human operators. In the event that telemetry is corrupted, out of sync or even inconsistent, a socio technical system can be subject to systemic failure.

This layer sets up the confidentiality of data traversing the system. It needs ongoing cross-collaboration, validation, and correlation between sensors, digital twins and/or AI controllers. Telemetry becomes a vital compliance-related tool and hence governed in the same way as the ensured reliability in the tool.

According to ENISA (2023), corrupted telemetry is one of the basic sources of the escalation of autonomy and misclassification of AI. This layer directly mitigates that risk since data fed into predictive and autonomous systems are trusted systems.

3.10 AI AND DIGITAL TWIN ANALYTICS LAYER

AI and digital twins are at the heart of today's critical infrastructure. They offer predictive analysis, detection of anomalies and simulation functions, which help in the decision-making process within the operations. Yet this can only be done when they are matched to proven telemetry and real-world conditions.

This layer is used for predictive models, anomaly detection engines or simulation environments. It makes sure that an AI system is explainable, traceable, and in line with a validated telemetry. Digital twins are continually updated with real-world conditions to ensure they don't become a hazard in operation.

The National Academies (2019) also highlight the importance of governance systems to preserve the integrity and accountability of systems using AI. This layer is reflective of those requirements.

3.11 AUTONOMOUS DECISION AND CONTROL LAYER

In critical infrastructures, ANSs are now directly involved in operational decision making. They can help to hold things together in a system, or they can help to cause disruptions when boundaries are not clearly drawn.

This layer defines the extent of machine authority. It establishes policies, procedures and decision traceability for human override. The ability to be autonomous becomes a controlled rather than a uncontrolled operating force.

This is in line with ENISA's (2023) advice on the need to regulate autonomy through well-established oversight mechanisms.

3.12 HUMAN-MACHINE TEAMING LAYER

In converged environments, human operators are still playing a crucial role. Autonomous systems can't replace their ability to draw inference from context, make judgements and intervening during instability. But their functions are no longer those of direct manipulation, but of supervision of governance.

This layer strengthens the importance of the human operators as supervisory authorities. It provides situational awareness, cognition, and procedural authority of operators to intervene effectively when autonomy misbehaves.

3.13 SYSTEMIC RISK MONITORING AND RESPONSE LAYER

The final layer is for the ability to identify propagation paths, predict cascading failures, and to coordinate cross domain containment. It shifts the SOC and GSOC to a proactive approach to systemic risk governance from incident management.

When AI enabled systems are based on common or correlated telemetry, there is a greater risk of cascading failures due to cross sector interdependencies (World Economic Forum, 2021). This layer tackles that challenge head on by offering simple monitoring and response in the digital, physical, industrial and human space.

4 DISCUSSION

This paper shows that the scenarios contain several key learning points about the behaviour of converged socio technical systems. First, events that start in digital spaces often spread into physical, industrial and human space. Propagation pathways such as telemetry corruption, digital twin drift, autonomy escalation and authentication instability are all examples of attacks beyond the traditional security perimeter. This is in line with findings of cyber physical systems research, which highlights the close relationship between the digital and physical processes (Lee et al., 2015).

Second, the scenarios show that autonomous systems amplify stability and instability. Autonomy optimises operational efficiency when working on validated telemetry and synchronised digital twins. But by taking action on corrupted data or unsynchronised data, autonomous systems can escalate incidents out of human control. The duality illustrates the need to have structured governance mechanisms to address possible systemic vulnerabilities as per ENISA's (2023) conclusion that AI enabled systems need governance.

Third, the scenarios highlight that human operators continue to play a crucial role in a converged world. They also are capable of reading the situation, making a judgment and intervening when the situation is unstable, which cannot be

replicated by autonomous systems. However, their function has evolved from direct operational control to governance oversight. The resulting change demands new types of situational awareness, cognitive bandwidth and procedural authority.

Fourth, the Converged Security Framework offers a single operational model that can control cross domain incidents. It has a layered structure that resonates with current work on systemic risk management (ISO/IEC 30147:2022, 2022), IoT interoperability (National Academies, 2019), and AI governance (National Academies, 2019). The framework shifts SOC and GSOC from a reactive incident response into proactive systemic risk governance.

Lastly, the scenarios underscore the compliance and documentation. Telemetry failures, autonomous actions, digital twin behaviour and impacts across domains must be reconstructed in detail for each incident. This requirement resonates with the focus of National Academies (2019) on accountability and traceability in AI systems.

4.1 IMPLICATIONS FOR SOC AND GSOC GOVERNANCE

4.1.1 Operational authority.

SOC and GSOC leaders are expected to always hold the top-level decision-making power on autonomous systems. This authority should be documented in a way that includes delegated authority guidelines, escalation procedures and decision trace documents. Autonomous systems must operate within clearly defined governance boundaries, supported by documented escalation procedures, override mechanisms, and human intervention protocols where necessary.

4.1.2 Situational awareness.

Situational awareness is cross domain in converged environments. All of SOC and GSOC operators should be able to view telemetry and digital twin projections, autonomous actions and the physical system behaviour as part of a single operational story. This is the need for domain integration, a single monitoring system and a unified domain to incorporate domain correlations.

4.1.3 Autonomy governance.

There must be structured systems of supervision for autonomous systems. This governance provisions includes input validation for telemetry, digital twin synching, insight to AI decision flows and autonomy boundary enforcement. Autonomy should be understood as an operable capability, not as an operating force.

4.1.4 Cross domain coordination.

Incidents are frequently extended into digital domain, physical domain and industrial domain as well as human domain in converged environments. Coordinating procedures across domains are thus required to be part of SOC and GSOC operations. These protocols guarantee that if containment is used in one domain it will not be used in another domain in such a way as to cause instability in that domain.

4.1.5 Systemic risk management.

A systemic attitude towards risk management must be embraced by SOC and GSOC leaders. This method covers propagation pathway monitoring, predicting cascading system failures, and cross domain containment. Systemic risk management changes the operational change from ways of dealing with incidents as they happen to more of a governance approach.

4.2 Limitations

This study is conceptual and scenario-based in nature. The scenarios are intended to illustrate potential cross-domain risk propagation pathways and governance challenges rather than predict specific operational outcomes. Future research should validate the proposed Converged Security Framework through empirical case studies, simulations, and operational implementation within critical infrastructure environments.

5 CONCLUSION

Modern critical infrastructure systems are becoming tightly coupled socio technical systems and the convergence of digital, physical, industrial and human areas is changing them. Operational decisions are now being shared by artificial intelligence, digital twins, Internet of Things ecosystems, and autonomous operational technology, which are all new areas of cross domain risk. This paper offers a multi scenario analysis that shows incidents emanating from one domain often cross over to another domain, causing systemic effects outside of the scope of traditional security. The propagation

paths in converged environments are complex, as seen in various corruptions like telemetry, drifted digital twins, autonomy propagation, AI misclassification, authentication instability, and cascading failure across multiple sectors.

The author wraps up by arguing that SOC and Global-SOCs are no longer just dealing with incidents but are now becoming governors of systemic risk. The changes need cross domain situational awareness and autonomy governance, predictive model alignment and systemic risk monitoring. To meet this evolution, the Converged Security Framework offers a single operational model that can support it. The multi-layered architecture is aligned with current research on cyber-physical systems, AI governance, and IoT interoperability, providing a structured framework for the responsible governance of modern critical infrastructure.

This is an ongoing, evolving process for critical infrastructure and therefore for the complexity of operational risk. However, when properly managed through structured governance and operational processes, convergence leads to resilience rather than fragility. With the Converged Security Framework, SOC/GSOC principals can offer a clear, certain, and methodical defence for modern systems.

STATEMENTS ON COMPLIANCE WITH ETHICAL STANDARDS

Acknowledgments

The author acknowledges the work of the researchers and practitioners in the domains of cyber physical systems, artificial intelligence governance and critical infrastructure protection. They are a vital basis in understanding converged risk and for the creation of structured governance models. The authors would also like to thank the other colleagues in the field for their important suggestions while preparing this manuscript.

Disclosure of Conflict of Interest

The author declares that there is no conflict of interest regarding the publication of this article.

Statement of Ethical Approval

This study is based on literature review, conceptual analysis, and scenario-based modelling. It did not involve human participants, animal subjects, clinical trials, or the collection of personal data. Ethical approval was therefore not required.

Statement of Informed Consent

Informed consent was not required because the study did not involve human participants or identifiable personal data.

REFERENCES

- [1] European Union Agency for Cybersecurity (ENISA) (2023) AI Cybersecurity Challenges: Threat Landscape for Artificial Intelligence.
- [2] ISO/IEC 30147:2022 (2022) Internet of Things (IoT) – Interoperability for IoT Systems. International Organization for Standardization, Geneva.
- [3] Lee, J., Bagheri, B. & Kao, H. (2015) 'A Cyber-Physical Systems Architecture for Industry 4.0-Based Manufacturing Systems', *Manufacturing Letters*, 3, pp. 18–23.
- [4] National Academies of Sciences, Engineering, and Medicine (2019) *Safeguarding the Bioeconomy: Protecting the Integrity of Science and Technology in a World of Rapid Change*. Washington, DC: The National Academies Press.
- [5] World Economic Forum (2021) *Shaping the Future of the Digital Economy and New Value Creation*. Geneva: World Economic Forum.