International Journal of Advanced Engineering and Management Research

Vol. 11, No. 06; 2026

ISSN: 2456-3676

Operationalizing Artificial Intelligence Governance for Critical Infrastructure: A Governance Assurance Framework for Trustworthy AI in High-Consequence Environments

Dr. Robb Shawe (Lead Author)[1], Dr. Robert W. Clark (Co-Author)[1]
[1]Department of Sustainability and Department of Cyber-Psychology
Capitol Technology University, 11301 Springfield Road, Laurel, Maryland, USA

doi.org/10.51505/ijaemr.2026.11605 URL: http://dx.doi.org/10.51505/ijaemr.2026.11605

Received: Nov 12, 2026 Accepted: Nov 20, 2026 Online Published: Nov 27, 2026

Abstract
Artificial intelligence (AI) increasingly informs high-consequence decisions in critical infrastructure. Operators need a practical way to establish who may act on an AI recommendation, what evidence supports that decision, and when the system should be restricted or reassessed. Building on the Artificial Intelligence Governance Framework for Critical Infrastructure (AIG-CI), this conceptual study proposes the Artificial Intelligence Governance Assurance (AIGA) Framework. A targeted interdisciplinary synthesis connects six assurance domains to six lifecycle phases through accountable roles, decision records, escalation rules, and seven categories of performance measures. Published cases from wind energy and hospital sepsis prediction illustrate how the framework could be applied to operational decisions and local model changes. The resulting model offers a basis for pilot implementation and empirical evaluation of decision traceability, oversight effectiveness, and corrective action. The proposed controls and indicators require field validation.

Keywords: artificial intelligence governance assurance, AIGA, artificial intelligence governance framework for critical infrastructure, AIG-CI, critical infrastructure protection, governance assurance, explainable artificial intelligence, human oversight, organizational resilience, democratic governance

1. Introduction

Critical infrastructure operators use artificial intelligence (AI) to prioritize maintenance, assess threats, allocate resources, and support time-sensitive decisions. As models, data sources, and operating conditions change, the organization must be able to verify that the authority, controls, and evidence supporting an AI-informed decision remain adequate. When an error can interrupt an essential service or affect safety or individual rights, leaders need to know who can act, what evidence is required, and what warrants returning to established procedures (Tabassi, 2023; Hollnagel, 2014; UNESCO, 2021).

Vol. 11, No. 06; 2026

ISSN: 2456-3676

The National Institute of Standards and Technology AI Risk Management Framework (NIST AI RMF) provides voluntary risk-management guidance through Govern, Map, Measure, and Manage. ISO/IEC 42001 sets requirements for an organizational AI management system and its continual improvement. Both address ongoing governance. AIGA builds on AIG-CI (Shawe & Clark, 2026) and contributes a proposed workflow for critical infrastructure operators that links a consequential decision to its authorized human decision-maker, supporting evidence, exception pathway, and verified corrective action. The distinctive design element is a linked, reviewable record across decision authorization, exception response, and corrective action. An operator can incorporate it into existing risk and management systems. The comparison is conceptual and makes no claim of superior performance or conformity with either reference framework (ISO/IEC, 2023; Tabassi, 2023).

The guiding question is how a critical infrastructure operator can turn AI governance commitments into observable decisions and auditable evidence across the lifecycle. Sections 3–6 connect six domains of responsibility to six lifecycle phases and define candidate performance measures. Section 7 applies those mechanisms to two published sector cases.

2. From Governance Architecture to Governance Assurance

AIG-CI establishes the governance architecture. AIGA specifies how operators can gather evidence that this architecture is functioning as intended during deployment, routine operations, and operational change. The unit of assurance is an AI-supported decision or control event within a defined use case, such as approving a consequential action, authorizing a model release, or closing an exception.

2.1 Conceptual Development and Theoretical Foundation

Framework development used a targeted narrative synthesis. Sources were selected for their relevance to the relationship among operator authority, the reliability of an AI-supported workflow, and public accountability. The synthesis grouped recurring requirements as obligations to govern, operate, verify, comply, learn, and account to affected parties. These obligations informed the six assurance domains. The authors then organized the resulting controls in the order an operator would authorize, deploy, monitor, review, assess, and revise a use case. The comparison with the NIST AI RMF and ISO/IEC 42001 clarified the operational scope of the proposal. The synthesis is conceptual and does not establish the completeness or effectiveness of the proposed controls.

Governance scholarship informs the assignment of executive responsibility and multidisciplinary review (Ansell & Gash, 2008; Bevir, 2013). Socio-technical systems theory informs the joint review of model behavior, staff workload, operating procedures, and decision authority (Carayon, 2006; Trist & Bamforth, 1951). Resilience engineering provides the basis for monitoring, response, fallback, and learning after an exception (Hollnagel, 2014). Explainable AI and work on human interaction with automation inform the requirement for usable decision explanations, calibrated reliance, and a practical ability to override a recommendation (Miller, 2019; Parasuraman et al., 2000). AI ethics and democratic accountability inform lawful

Vol. 11, No. 06; 2026

ISSN: 2456-3676

authority, proportionality, and context-appropriate transparency and redress (Floridi & Cowls, 2019; UNESCO, 2021).

3. Artificial Intelligence Governance Assurance (AIGA)

Each domain covers an enduring assurance responsibility. The domains can contribute to several lifecycle phases, and a single phase may draw evidence from several domains.

3.1 Strategic Governance Assurance

An executive sponsor ratifies the use case, assigns an accountable service owner, sets risk tolerances, and approves the resources, operating limits, and conditions for suspension or expansion.

3.2 Operational Governance Assurance

Operations personnel apply decision rules and maintain the approved fallback. The service owner tests whether operators can evaluate and challenge recommendations within the time available and can reconstruct consequential decisions from the record.

3.3 Technical Governance Assurance

The technical lead documents data provenance, the approved model version, local evaluation, relevant subgroup or operating-condition checks, and monitoring limits. Material changes require reassessment before use.

3.4 Legal and Ethical Governance Assurance

Compliance, legal, or privacy personnel review decision authority, proportionality, privacy, retention, and applicable reporting and review obligations. Where a use case affects individual rights, the operator identifies the applicable notice and review mechanisms before launch.

3.5 Organizational Governance Assurance

The program owner verifies staff readiness and assigns corrective actions. Changes to procedures, training, and controls are tested and documented before findings are closed.

3.6 Societal Governance Assurance

The responsible stakeholder or public accountability lead establishes feedback and complaint channels where appropriate, identifies potentially affected groups, and ensures that reportable events reach the relevant internal or external authority. Public communication balances accountability with privacy and operational security.

3.7 Integrated Governance Assurance

A common use-case identifier connects the six domains. For example, when technical monitoring identifies degraded performance, operations restricts the workflow as needed, compliance assesses any reporting or rights implications, and the sponsor reviews the evidence supporting resumption. The action record documents the responsible owner and verified closure.

Vol. 11, No. 06; 2026

ISSN: 2456-3676

3.8 Organizational Implementation Considerations

Start by selecting one bounded decision-support workflow and documenting its current performance and risks. Inventory the model, vendor, data sources, approved purpose, prohibited uses, and affected parties. The executive sponsor designates the service owner and the roles that can approve a release, stop use, and authorize resumption. Before go-live, the technical and operational leads verify local performance, staff readiness, audit records, and fallback through a controlled test or shadow deployment.

During the pilot, the service owner reviews exceptions at the intervals specified in planning. The sponsor uses documented results to decide whether to expand, restrict, revise, or discontinue the use case. Reuse incident, work-order, and audit systems where they can link records to the correct use case and model version.

Legacy systems may lack versioned data or reliable event logs. The pilot can begin with a bounded manual record, provided the required evidence is complete and retrievable. Vendor opacity can be addressed through agreed access to version notices, evaluation evidence, and incident information; if evidence is inadequate to assess a material risk, narrow or defer the affected use. Understaffed teams can pool review expertise and test the burden of additional approvals during the pilot. Risk-based review and scheduled drills can reduce the pressure to bypass human oversight during time-critical operations.

4. Operational Governance Assurance Lifecycle

Figure 1 shows the progression from planning to organizational learning. Operational monitoring continues during use, and human decision validation precedes any consequential action requiring authorization.

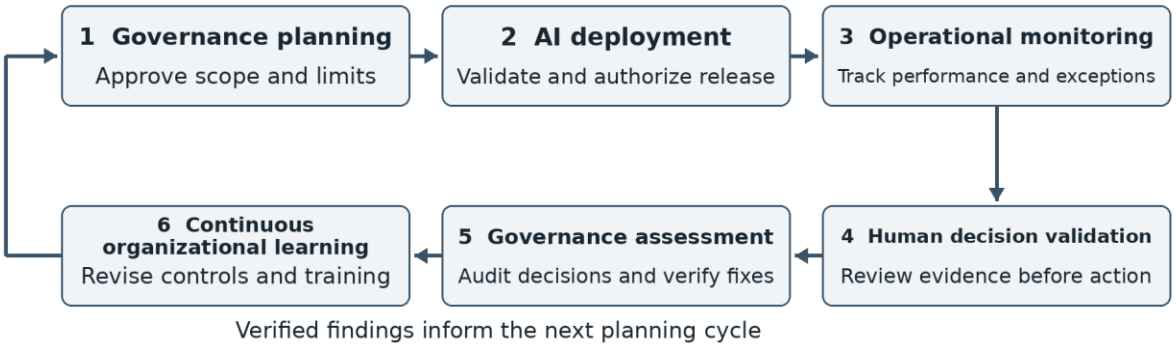


Figure 1. Operational Governance Assurance Lifecycle

Note. Authors' conceptual framework. The return arrow shows how verified findings inform the next planning cycle.

An exception detected during monitoring or decision review can trigger an immediate pause and reassessment. The organization can return to planning when findings change the approved purpose, risk limits, or safeguards.

Vol. 11, No. 06; 2026

ISSN: 2456-3676

Table 1 specifies responsible roles, participating domains, and minimum evidence for each phase of a bounded pilot.

Table 1. Assurance Domains, Responsibilities, and Records by Lifecycle Phase

Lifecycle phase	Lead and supporting domains	Accountable role	Minimum assurance record
Governance planning	Strategic; legal and ethical, societal	Executive sponsor	Approved use case, risk limits, authorized roles, baseline, and fallback.
AI deployment	Technical; operational, legal, and ethical	Technical lead validates. Sponsor authorizes release.	Version-specific local validation, release criteria, fallback test, and sponsor authorization.
Operational monitoring	Operational; technical, organizational	Service owner	Performance and exception log with dispatch, escalation, and response times as applicable.
Human decision validation	Operational; legal and ethical, technical	Authorized decision-maker	Output and version, reviewed evidence, decision rationale, approver, and disposition.
Governance assessment	Organizational; strategic, legal and ethical, societal	Assurance lead	Audit findings, any reportable impacts, corrective owner and deadline, and closure evidence.
Continuous organizational learning	Organizational; strategic, technical, societal	Service owner	Updated procedures or training, verification results, and sponsor decision on material changes.

Note. Roles and records are proposed for local adaptation. Other domains may contribute when a decision changes the risk to a service, workforce, or affected group.

The linked records allow an operator to trace a decision from approval through monitoring and corrective action. Sections 4.1–4.6 describe how those records are created and used.

4.1 Phase One: Governance Planning

The executive sponsor approves the use case, risk limits, designated approvers, and escalation triggers. The service owner records the baseline workflow, including how decisions are made when the AI system is unavailable.

4.2 Phase Two: AI Deployment

The technical lead tests the approved version against locally defined criteria and records known limitations, operating conditions, and rollback steps. The sponsor authorizes release after the service owner confirms operator readiness and applicable reviews are complete.

Vol. 11, No. 06; 2026

ISSN: 2456-3676

4.3 Phase Three: Operational Monitoring

The service owner tracks data quality, model performance, staff workload, and accumulating exceptions against the agreed limits. The technical lead documents suspected drift or unexpected changes. The on-duty authority can restrict or suspend use and invoke the fallback while the exception is assessed.

4.4 Phase Four: Human Decision Validation

Before a consequential action, the authorized decision-maker reviews relevant source evidence, limitations, uncertainty, and feasible alternatives. The record links the model version and recommendation to the human decision, rationale, time, and any escalation. Local procedures identify decisions requiring a second review and specify the fallback when review cannot be completed in time.

4.5 Phase Five: Governance Assessment

The assurance lead reviews a risk-based sample of decisions and all designated critical exceptions. Findings name the accountable owner, corrective action, deadline, and evidence required for closure. A qualified reviewer checks closure independently of the person completing the action where practicable.

4.6 Phase Six: Continuous Organizational Learning

The service owner converts verified findings into revised training, monitoring, or procedures. Material changes return to planning and release review. The executive sponsor records the decision to expand, restrict, resume, or discontinue the use case.

5. Human Oversight, Organizational Accountability, and Decision Integrity

Decision Integrity is the proposed capacity to make a consequential AI-supported decision reviewable for evidence, authority, proportionality, and follow-through. A reviewer should be able to reconstruct what the system recommended, what the decision-maker knew at the time, and why the action was authorized, modified, deferred, or declined.

5.1 Human Authority and Escalation

Assign one accountable decision-maker for each covered action and distinguish that role from technical, legal, and operational advisers. To counter automation bias, require reviewers to record corroborating or conflicting evidence and permit them to override without penalty for a justified disagreement. Where practicable, a second reviewer should independently assess selected high-risk decisions. A rapid fallback and an on-duty alternate help preserve review under time pressure (Miller, 2019; Parasuraman et al., 2000).

Missing decision evidence, out-of-bounds performance, a suspected security or privacy incident, or an unauthorized model change triggers escalation under the approved risk rules. The designated on-duty authority may pause the affected workflow, invoke the fallback, and notify the incident owner. Resumption requires documented remediation or containment, revalidation proportionate to the change, and authorization from the designated release authority.

Vol. 11, No. 06; 2026

ISSN: 2456-3676

5.2 Emerging Risks and Evidence

Monitoring should account for distribution shift, unexplained performance differences across relevant groups or operating conditions, and model or supplier updates. Generative systems introduce additional concerns such as inaccurate outputs, unverified citations, inappropriate disclosure, and manipulation of inputs. Operators should establish permitted uses and independently corroborate material claims before relying on generative outputs in covered decisions. Version records, change reviews, access controls, and appropriate incident reporting support investigation and remediation (Autio et al., 2024; Tabassi, 2023).

Decision records should capture the applicable authorization and enough evidence for review, subject to local access and retention rules. For public-facing applications, the responsible lead reviews relevant complaints and communicates material findings through the authorized channels. Disclosure should be proportionate to the consequences and respect applicable privacy and security restrictions.

6. Governance Performance Measurement and Organizational Assurance Metrics

The seven measurement categories evaluate outcomes and control execution across the six assurance domains. Several categories may draw on the same record. For example, a decision record can support both an oversight measure and an audit-readiness measure. Each pilot should define what counts as an eligible event, who records it, and when the owner reviews results.

Vol. 11, No. 06; 2026

ISSN: 2456-3676

Table 2 gives one candidate measure for each category, together with the record, owner, and proposed review cadence and response.

Table 2. Candidate Governance Assurance Measures

Category	Candidate measure	Record and owner	Review and response
Governance effectiveness	Use cases with current signed authorization ÷ active use cases	Use-case register Executive sponsor	Monthly. Restrict use when authorization expires or scope changes.
Explainability and transparency	Sampled decisions with reconstructable evidence, AI output, version, and rationale ÷ sampled decisions	Decision log Assurance lead	Monthly. Escalate missing critical evidence and correct record capture.
Human oversight effectiveness	Actions with required human approval recorded before execution ÷ actions requiring approval	Decision and action logs Service owner	Per critical event and monthly. Escalate any action taken without required authorization.
Audit readiness and accountability	Corrective actions closed on time with verified evidence ÷ corrective actions due.	Action register Assurance lead	Monthly. Escalate overdue high-risk actions to the sponsor.
Organizational resilience	Time from restriction trigger to fallback activation, reported with event counts	Incident log and drill records Service owner	After each event or drill, investigate missed recovery limits.
Stakeholder confidence and public trust	Eligible complaints due in the period and resolved on time ÷ eligible complaints due in the period	Complaint register Stakeholder lead	Quarterly. Review overdue cases and recurring concerns.
Governance maturity	Scheduled controls with an assigned owner and current operating evidence ÷ controls due for review	Control inventory Program owner	Quarterly. Assign a corrective action where ownership or evidence is missing.

Note. Report counts with rates and record zero eligible events as not applicable. Review intervals and triggers are examples for local calibration. They are proposed evaluation criteria, not validated thresholds.

Reconcile the decision log with the operational action log to identify covered actions that lack an approval record. The technical lead can also track the share of deployed model versions with current local validation, and the compliance lead can track the share of covered changes with completed legal or privacy review before release. Use the deployment and change registers to establish the eligible populations.

Before launch, agree on local tolerances using the baseline, consequence of failure, and staffing capacity. Record exceptions even when aggregate performance meets its target. Compare pilot

Vol. 11, No. 06; 2026

ISSN: 2456-3676

and baseline rates using the same eligibility criteria and assess the burden of review alongside any improvement in traceability where samples are small, present counts, and examine each consequential event.

Override frequency should be paired with a review of the rationale and outcome. A high rate may identify appropriate challenge or a model problem, and a low rate may reflect either reliable recommendations or uncritical acceptance. Complaint closure and stakeholder communications provide evidence of follow-through. Direct assessment of public trust requires additional evidence, such as a separately designed stakeholder survey.

7. Cross-Sector Implementation of Artificial Intelligence Governance Assurance

The two case studies use published empirical findings to illustrate how an operator could apply AIGA. The source studies evaluated their own models or methods. The assurance workflows proposed here apply the framework and require separate evaluation.

7.1 Energy: Wind Turbine Condition Monitoring

Castellani et al. (2024) examined an industrial wind turbine gearbox case in which established diagnostic methods failed to identify serious degradation early enough to avoid prolonged downtime. Their retrospective analysis combined supervisory control and data acquisition (SCADA) and vibration measurements and identified the fault earlier than the methods evaluated in the case. The proposed assurance question is when that additional evidence should change a maintenance decision.

In a bounded pilot, the asset manager authorizes AI recommendations for inspection prioritization, with the duty engineer retaining authority over maintenance decisions under the approved procedure. The technical lead checks that the sensor data and version are within the approved envelope. If an alert is credible, the engineer reviews corroborating condition data and records the rationale in the work order. If sensor data are missing or the model output contradicts other evidence, the engineer can revert to the existing inspection process and escalate the discrepancy. The service owner reviews authorization coverage, alert burden, and findings from completed inspections before requesting any expansion.

7.2 Healthcare: Sepsis Risk Alerts

Wong et al. (2021) found limited performance for the original Epic Sepsis Model in an external hospital validation. A later multicenter study of an updated version reported better discrimination than the original alongside substantial variation across institutions (Wong et al., 2026). The two evaluations illustrate why version-specific and local evidence matters when a model is used to inform a time-sensitive workflow.

In an AIGA pilot, the clinical sponsor authorizes a local shadow evaluation before approving a live alert workflow. Clinical informatics verifies the deployed model version and evaluates missed cases, alert burden, and performance across relevant units or patient groups. The clinician evaluates an alert against the clinical record and continues to use the institution's established pathway. If alert burden exceeds the locally approved limit or a material version change occurs,

Vol. 11, No. 06; 2026

ISSN: 2456-3676

the designated patient-safety authority can suspend the AI alert workflow while the standard pathway continues. Resumption follows documented local retesting and sign-off.

7.3 Transfer Across Sectors

The two cases share the requirement to retain a versioned record, preserve accountable human authority, maintain a fallback, and reassess changes. The authorized roles, review speed, evidence, and thresholds depend on the service and consequences. Transportation, emergency management, cybersecurity, and public administration are candidate settings for further testing of that transfer.

8. Limitations

AIGA is a conceptual operational model. Its performance indicators, proposed roles, and authorization pathways have not yet undergone organizational, longitudinal, or cross-sector validation. The two secondary case illustrations demonstrate how the proposed framework can be applied analytically. They establish neither AIGA's feasibility under routine conditions nor its effect on operational or public outcomes.

The targeted synthesis is selective. Transferability depends on legal authority, availability of vendor evidence, local data quality, staffing, and the workability of fallback procedures. The indicators primarily measure observable governance processes. They should be interpreted alongside safety, workload, rights, and service outcomes.

9. Future Research

Initial research should test the feasibility of recording events and the consistency of the proposed measures. Independent reviewers can classify the same decision and exception records to estimate agreement on eligibility, evidence completeness, and closure. An expert and practitioner Delphi study could then refine role assignments, indicators, and phase transitions before field use.

A prospective pilot could compare matched observation windows in the same workflow before and after implementation, or use a staggered rollout with comparable teams where feasible. Primary outcomes could include pre-action authorization coverage, reconstructable decisions, and timely verified corrective actions. Track time to decision, alert burden, staffing effort, and missed or delayed actions alongside those outcomes—document concurrent staffing, policy, and model changes when interpreting observed differences.

Simulation and tabletop exercises can test fallback activation, override quality, response under time pressure, and evidence required to resume operations. Longitudinal and cross-sector studies can then assess whether the measures behave consistently across different services and organizational capacities. These studies can test whether improved documentation corresponds to better decision quality and service outcomes.

The next article in the series proposes the Artificial Intelligence Governance Maturity Model (AIGMM). The concluding article develops the Integrated Artificial Intelligence Governance

Vol. 11, No. 06; 2026

ISSN: 2456-3676

Capability Framework (IAIGCF). Those contributions extend the research program into maturity and enterprise capability. Field evaluation of AIGA remains a distinct research objective.

10. Conclusion

AIGA translates the AIG-CI governance architecture into a proposed assurance workflow for high-consequence AI-supported decisions. It connects the six assurance domains to lifecycle phases through accountable roles, evidence, escalation, and corrective action. Candidate measures specify how an operator could test the resulting workflow, while the energy and healthcare cases show why authorization, local evaluation, and fallback must be adapted to the service.

The framework's scholarly contribution is an operational definition of governance assurance centered on observable decisions and evidence. Its effectiveness can be tested by evaluating whether an implementation improves decision traceability, meaningful oversight, and verified corrective action while preserving service continuity and manageable review demands.

Author Note

Dr. Robb Shawe, PhD, is a scholar-practitioner, United States Navy veteran, and Chief Executive Officer of New York Security Consulting Professionals L.L.C. His research focuses on artificial intelligence governance, behavioral threat assessment, critical infrastructure protection, cybersecurity governance, resilience engineering, socio-technical systems, organizational decision-making, and interdisciplinary governance frameworks. As lead author, Dr. Shawe conducted the research, synthesized the scholarly literature, developed the conceptual and analytical framework, wrote the manuscript, implemented revisions, and prepared the work for publication.

Dr. Robert W. Clark, PhD, currently serves as Deputy Mayor of Public Safety for the City of Los Angeles and previously served as an Assistant Special Agent in Charge with the Federal Bureau of Investigation. His research focuses on behavioral threat assessment, mass violence prevention, cyberpsychology, intelligence analysis, and public safety governance. Dr. Clark's doctoral scholarship and extensive practitioner experience provided foundational domain expertise that informed the interdisciplinary development of the Artificial Intelligence Governance Assurance (AIGA) Framework.

The authors affirm that this manuscript constitutes original scholarly work developed through the interdisciplinary integration of their respective doctoral research programs and professional expertise. The governance assurance model and analytical framework presented herein substantially extend beyond either dissertation individually while building upon the Artificial Intelligence Governance Framework for Critical Infrastructure (AIG-CI) introduced in the preceding article.

The authors declare no conflicts of interest related to this research. This manuscript has not been published previously and is not under consideration by another publication outlet. Copyright

Vol. 11, No. 06; 2026

ISSN: 2456-3676

shall remain with the authors unless transferred pursuant to a publisher's publication agreement upon acceptance.

Authorship Statement

This manuscript represents a collaborative scholarly effort by **Dr. Robb Shawe** and **Dr. Robert W. Clark** and is informed by the complementary foundations of their respective doctoral research programs.

Dr. Robb Shawe served as the lead author and was responsible for research, literature synthesis, conceptual integration, development of the **Artificial Intelligence Governance Assurance (AIGA) Framework**, manuscript writing, editing, revisions, formatting, and preparation of the final submission. Dr. Shawe also developed interdisciplinary governance concepts that integrate governance science, resilience engineering, socio-technical systems theory, explainable artificial intelligence, cybersecurity governance, behavioral threat assessment, critical infrastructure protection, and democratic governance into a unified operational assurance framework.
Dr. Robert W. Clark contributed the foundational subject-matter expertise derived from his doctoral research and extensive professional experience in law enforcement leadership, behavioral threat assessment, intelligence operations, public safety governance, and mass violence prevention. His practitioner insights and conceptual contributions informed the development, refinement, and operational applicability of the governance assurance framework. Both authors are responsible for the scholarly integrity and accuracy of the work.

The Artificial Intelligence Governance Assurance (AIGA) Framework represents the original conceptual contribution of this manuscript and was developed through the authors' interdisciplinary synthesis of governance theory, resilience engineering, socio-technical systems theory, explainable artificial intelligence, organizational decision-making, democratic governance, behavioral threat assessment, and organizational learning.

Statement on Authorship and AI-Assisted Editorial Technologies

The authors developed the AIGA framework, including its six-domain and six-phase architecture. They affirm that the intellectual conception, theoretical development, governance assurance framework, analytical reasoning, interpretation of the scholarly literature, interdisciplinary synthesis, operational governance model, and conclusions presented in this manuscript represent their original academic work and professional expertise.

During manuscript preparation, contemporary digital writing and editorial technologies may have been used in a limited capacity to support language refinement, grammatical correction, formatting consistency, and editorial organization. These technologies were not used to develop the manuscript's research questions, governance models, analytical arguments, theoretical foundations, methodological approaches, operational assurance lifecycle, governance performance framework, or scholarly conclusions.

All substantive intellectual contributions—including the development of the Artificial Intelligence Governance Assurance (AIGA) Framework, the Operational Governance Assurance

Vol. 11, No. 06; 2026

ISSN: 2456-3676

Lifecycle, the Governance Performance Assessment model, and the accompanying analytical framework—were made by the authors and were critically reviewed, verified, and approved prior to submission.

The authors further acknowledge that current automated AI-authorship detection systems rely on probabilistic methodologies that may incorrectly classify highly structured academic writing, technical governance documents, or consistently formal scholarly prose as machine-generated. Such classifications do not constitute evidence of authorship and should not be regarded as definitive indicators of a manuscript's origin.

Accordingly, the authors accept full responsibility for the originality, scholarly integrity, accuracy, and intellectual content of this work and certify that the manuscript constitutes their independent contribution to advancing interdisciplinary scholarship in artificial intelligence governance, governance assurance, critical infrastructure protection, organizational resilience, cyberpsychology, and democratic governance.

Copyright Notice

© 2026 Dr. Robb Shawe and Dr. Robert W. Clark. All rights reserved.
This manuscript constitutes an original scholarly work developed through the collaborative integration of the authors' respective doctoral research, professional expertise, and interdisciplinary scholarship. No portion of this manuscript may be reproduced, distributed, or transmitted in any form or by any means without the prior written permission of the copyright holders, except as permitted by applicable copyright law or publisher agreement.

The views and opinions expressed herein are solely those of the authors and do not necessarily reflect the official positions, policies, or endorsements of any governmental agency, academic institution, employer, or affiliated organization.

References

Ansell, C., & Gash, A. (2008). Collaborative governance in theory and practice. *Journal of Public Administration Research and Theory, 18*(4), 543–571.

Autio, C., Schwartz, R., Dunietz, J., Jain, S., Stanley, M., Tabassi, E., Hall, P., & Roberts, K. (2024). *Artificial intelligence risk management framework: Generative artificial intelligence profile* (NIST AI 600-1). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.600-1

Bevir, M. (2013). *A theory of governance*. University of California Press.

Carayon, P. (2006). Human factors of complex sociotechnical systems. *Applied Ergonomics, 37*(4), 525–535.

Castellani, F., Natili, F., Astolfi, D., & Vidal, Y. (2024). Wind turbine gearbox condition monitoring through the sequential analysis of industrial SCADA and vibration data. *Energy Reports, 12*, 750–761. https://doi.org/10.1016/j.egyr.2024.06.041

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review.*

Vol. 11, No. 06; 2026

ISSN: 2456-3676

Hollnagel, E. (2014). *Safety-I and Safety-II: The Past and Future of Safety Management*. Ashgate.

International Organization for Standardization & International Electrotechnical Commission. (2023). *Information technology — Artificial intelligence — Management system* (ISO/IEC Standard No. 42001:2023). https://www.iso.org/standard/42001

Miller, T. (2019). Explanation in Artificial Intelligence: Insights from the social sciences. *Artificial Intelligence, 267*, 1–38.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics, 30*(3), 286–297.

Shawe, R., & Clark, R. W. (2026). Artificial intelligence governance for critical infrastructure protection: A human-centered conceptual framework. *International Journal of Advanced Engineering and Management Research, 11*(6).

Tabassi, E. (2023). *Artificial intelligence risk management framework (AI RMF 1.0)* (NIST AI 100-1). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-1

Trist, E. L., & Bamforth, K. W. (1951). Some social and psychological consequences of the longwall method of coal-getting. *Human Relations, 4*(1), 3–38. https://doi.org/10.1177/001872675100400101

UNESCO. (2021). *Recommendation on the Ethics of Artificial Intelligence*.

Wong, A., Currey, D., Schwinne, M., Park-Egan, B., Meyer, S., Gutting, A., Cao, J., Khan, S., Dantes, R., Pan, T., Buchman, T., Singh, K., Bhavani, S. V., Lyons, P. G., Sjoding, M. W., & Tarabichi, Y. (2026). Multicenter prospective validation of an updated proprietary sepsis prediction model. *JAMA Network Open, 9*(2), e260181. https://doi.org/10.1001/jamanetworkopen.2026.0181

Wong, A., Otles, E., Donnelly, J. P., Krumm, A., McCullough, J., DeTroyer-Cooley, O., Pestrue, J., Phillips, M., Konye, J., Penoza, C., Ghous, M., & Singh, K. (2021). External validation of a widely implemented proprietary sepsis prediction model in hospitalized patients. *JAMA Internal Medicine, 181*(8), 1065–1070. https://doi.org/10.1001/jamainternmed.2021.2626