International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal

A VALIDATION-FIRST FRAMEWORK FOR EVALUATING CYBER-INTELLIGENCE GOVERNANCE: DECISION ADMISSIBILITY, FAIR COMPARISONS, AND AN INDONESIAN APPLICATION PROTOCOL

Yudy, Mandega Raswono[1], Agus, Suradika[2] and Evi, Satispi[3]
[1] Doctoral Program in Public Administration, Faculty of Social and Political Sciences, Universitas Muhammadiyah Jakarta, Jakarta, Indonesia
Email: yudy.mandega@umj.ac.id[1]

Abstract: Comparisons of cyber-intelligence governance can produce precise numerical rankings without establishing that the compared workflows represent credible institutional alternatives. This methodological study develops a validation-first framework for comparing hierarchical, participant-governed, administrative-broker, and dual-hub arrangements. It separates communication topology from decision protocol, operational capability, resource allocation, and failure exposure. An action-specific admissibility rule distinguishes accumulated evidence from mandatory information, current authority, and release-time validity. Four analytical diagnostics address compensatory evidence thresholds, architecture–protocol confounding, finite-campaign queue overload, and incomplete accounting of redundancy. A deterministic verification implementation enumerates all 64 evidence subsets in a six-class example, checks a threshold condition across 101 thresholds, and reproduces exact queueing and recovery identities. At the inherited 0.72 threshold, eight logical states qualify; four lack an item required by the explicitly illustrative action specification. These counts establish a logical counterexample, not an operational error rate. An Indonesian application protocol translates the framework into observable process records, practitioner challenge, calibration, and held-out exercise evaluation. No operational incidents, practitioner responses, measured policy effects, or new architecture-performance rankings are reported. The contribution is an auditable specification of what must remain comparable, what must be measured, and what evidence is needed before simulation outputs can support institutional recommendations.

Keywords (English): *Cyber-intelligence governance; network governance; simulation credibility; decision admissibility; comparative evaluation; model verification.*

INTRODUCTION

Cyber-intelligence coordination involves more than moving an alert between organizations. A response may require technical interpretation, knowledge of the affected service, assessment of consequences, an accountable decision, and execution by an authorized party. The institutional problem lies in connecting these activities without assuming that all participants share the same responsibilities or decision rights (Omolere, 2026). Research on cybersecurity and security politics describes the interaction between complex technologies and fragmented governance (Dunn Cavelty & Wenger, 2020). NIST's incident-response guidance likewise situates response within wider cybersecurity risk management rather than treating it as an isolated technical sequence (Lee, 2024).

International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal

Governance architecture is therefore a relevant object of study, but an architectural label is not a complete experimental intervention. A hierarchy can use standardized information, distributed analysis, and delegated decisions. Participant governance can use parallel endorsements rather than a serial meeting. A separate coordinator can reduce some hand-offs while adding staffing and maintenance obligations. A dual-hub design can incur replication costs while preserving state during a failure. Network-governance theory distinguishes organizational forms and their contingent effectiveness; it does not assign universal communication speeds (Petukh, Bashkatov, & Shvyreva, 2023), approval times (Imran, 2026), or failure probabilities to those forms (Provan & Kenis, 2008).

This creates a particular risk in simulation-based policy research. When the preferred architecture receives better communication, faster approval, additional redundancy, and easier decision criteria, a favorable output cannot identify which change produced the advantage. Conversely, an apparently rigorous comparison can disadvantage a design by charging its costs while excluding the circumstances in which its benefits arise. Matching random seeds or server counts does not resolve these problems. The definition of the intervention and the conditions under which its output is evaluated must also be comparable.

A second difficulty concerns what the output represents. Receipt of a specified amount of information is not necessarily a correct interpretation of that information. A decision made before a deadline is not necessarily authorized. Delivery of instructions is not containment of a cyber incident. A weighted evidence score may be useful as a prioritization device, but it cannot silently replace requirements that a specific action must satisfy. Scientific interpretation depends on keeping these constructs distinct throughout the model, the analysis, and the conclusion.

The present study addresses these issues through a methodological framework and analytical demonstrations. Its purpose is not to establish the best cyber-intelligence architecture for Indonesia. It asks instead: RQ1, which features must be specified or held comparable before a governance contrast has an interpretable meaning? RQ2, how can a model distinguish evidence accumulation from an admissible, currently authorized decision? RQ3, what verification and empirical evidence are needed to move from a computational demonstration to an institutional claim? Here, *admissible* means satisfying an explicitly stated action specification (Egbedayo, 2026); it does not mean that software determines legal compliance (Ikhalia, 2026).

The development case is an uncalibrated coordination prototype and its documented methodological audit. Rather than replace its questioned percentages with different synthetic percentages, this article uses the identified problems to derive explicit comparison requirements and checkable counterexamples. The resulting contribution comprises a multidimensional comparison specification, an action-dependent evidence and authorization gate, four analytical diagnostics, and an observable evaluation protocol. The identities used in

International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal

the diagnostics are elementary and are not presented as newly discovered queueing or reliability theory. Their role is to expose failures of interpretation that a large simulation experiment can otherwise conceal.

The scope follows the distinction between implementation verification and validation for an intended application (Sargent, 2010). Reporting guidance and model-description protocols support transparency, but they cannot supply missing institutional observations (Monks, et al., 2019). Accordingly, the completed work reported here consists of specification, derivation, and deterministic software verification. The empirical application remains prospective. Section 2 positions the framework in related research. Section 3 describes its evidence base and Indonesian scope. Section 4 develops the evaluation specification. Section 5 presents analytical results and completed checks. Section 6 sets out the application protocol. Sections 7–9 discuss contributions, limitations, and conclusions (Grimm et al., 2020).

LITERATURE REVIEW

Governance forms and operational mechanisms

Provan and Kenis distinguish participant-governed, lead-organization, and network-administrative arrangements (Provan & Kenis, 2008). The later synthesis by van den Oord and colleagues emphasizes governance modes, contingency factors, and network-level tensions rather than a universally preferred structure. These distinctions guide the organizational alternatives considered here. They do not make a communication graph a direct measurement of a governance mode: two networks with similar edges can allocate decision rights differently, while the same governance arrangement can operate over different technical channels (Van den Oord, Kenis, Raab, & Cambré, 2023).

Collaborative-governance scholarship adds process and institutional conditions. Ansell and Gash examine the circumstances and processes through which collaboration develops (Ansell & Gash, 2008), while Emerson, Nabatchi, and Balogh connect collaborative dynamics with capacity for joint action. In a computational representation, it is therefore important to identify which part of collaboration is actually implemented. A reduction in clarification messages may represent a schema intervention. Shorter approval processing may represent a particular delegation procedure. Neither is, by itself, a measure of trust, legitimacy, motivation, or institutional learning (Emerson, Nabatchi, & Balogh, 2012).

We retain four reference configurations from the development case: G1, a hierarchical collector; G2, a participant-governed coordination process; G3, a separate administrative broker; and G4, a dual-hub arrangement with lateral exchange. These are configuration families, not four fully specified treatments. Each requires a graph, routing rules, an approval protocol, staffing, and an admissibility specification. A claim about one implemented G2 protocol should not be generalized to participant governance as a class. The framework makes that distinction explicit rather than embedding the label directly into an outcome equation.

International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal

Interoperability and the meaning of a decision

Public-sector interoperability spans policy, management, and technology (Pardo, Nam, & Burke, 2012). Similarly, guidance on cyber threat information sharing addresses goals, sources, sharing scope, distribution rules, and the use of information (Johnson, Badger, Waltermire, Snyder, & Skorupka, 2016). These perspectives motivate a separation between transport, interpretation, and permission. A delivered message may have an uncertain source or an obsolete asset identifier; a technically complete dossier may be unavailable to the role that must approve the action; an approval may expire before release (Fu, Sun, Xu, & Yue, 2025).

The proposed admissibility rule is a research specification for these distinctions. It combines compensatory support, where substitutable evidence may contribute to a score, with non-compensatory requirements, where a missing item cannot be offset by unrelated information. Which requirements are non-compensatory depends on the incident and action. Attribution may be optional for a reversible protective measure yet necessary for a different decision. The framework therefore rejects both a universal weighted threshold and a universal requirement to await every possible evidence class.

Decision correctness is a further construct. A process can comply with its stated evidence and approval rules while reaching a substantively mistaken conclusion. The framework retains correctness as a separately assessed outcome when a defensible assessment rubric or ground truth exists. When it does not exist, the field remains unassessed. Logical compliance is not converted into correctness by adding another arbitrary probability function.

Simulation credibility and reporting

Sargent distinguishes conceptual-model validity, implementation verification, operational validity, and data validity, with judgments tied to the intended use (Sargent, 2010). STRESS supplies a reporting structure for empirical simulation studies. The ODD protocol supports explicit descriptions of model purpose, entities, processes, and assumptions (Grimm et al., 2020). Together, these sources establish why a model description, executable code, and a validation argument answer related but different questions (Katiku, 2023).

This study adapts those distinctions into a practical sequence: specify the claim; identify its required evidence; test the corresponding logic; challenge structural assumptions; and compare model behavior with observations when an application claim is intended. The sequence is an application-specific synthesis, not a replacement reporting standard. Passing a test documents one property of one implementation. It does not certify the remaining properties of that implementation or the realism of its institutional mapping.

International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal

An important distinction is between numerical and structural sensitivity. Varying a service-time multiplier leaves the decision protocol unchanged. Replacing serial approval with parallel role-specific endorsements changes the protocol itself. Varying a per-incident outage probability leaves incident-local failure semantics unchanged. Replacing those semantics with a shared calendar outage changes the modeled system. Both kinds of variation can be necessary, but a numerical parameter sweep should not be described as testing unimplemented model forms.

What the present framework adds

The framework connects organizational constructs to inspectable computational obligations. It requires an explicit action definition, a comparison contract, an evidence-status ledger, and a verification test for each consequential mechanism. Its analytical contribution is the joint use of threshold counterexamples, non-identification examples, finite-campaign queue checks, and a recovery-opportunity check before interpreting architecture rankings. Table 1 locates this contribution relative to the source literature. The comparison is conceptual; it does not claim performance superiority over studies with different problems, data, or models (Grünig & Kühn, 2025).

Table 1 Relationship to the literature and the methodological contribution.

Research strand	Established contribution	Question addressed here
Network governance	Distinguishes forms, contingencies, and tensions	Which graph, protocol, and resource changes define the implemented treatment?
Collaborative governance	Explains institutional processes and capacity for joint action	Which operational mechanism is represented, and which construct remains unmeasured?
Interoperability and sharing	Connects technical exchange with policy and management	Does information receipt satisfy the requirements of the particular action?
Verification and validation	Evaluates correctness and fitness for an intended application	Which claim is supported by each completed check?
STRESS and ODD	Improves reporting and model description	Are protocol, failure, and evidence assumptions traceable to executable checks?

The literature treatment is a focused conceptual synthesis. It is not a systematic review, bibliometric study, or claim of exhaustive coverage. No database-screening count, exclusion flow, or risk-of-bias assessment is represented as having been completed.

International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal

RESEARCH METHODS

A methodological study with a documented development case

The study combines source-based conceptual analysis, a retrospective audit case, formal reasoning, and newly executed deterministic checks. The development materials comprise the original Computation manuscript, a later Systems-formatted prototype manuscript, and a methodological audit. The original and later manuscripts use different model definitions and outcomes; their numerical results cannot be treated as successive estimates from one unchanged experiment. Supplementary Material S1 records the files, their roles, and their hashes.

The audit identified concerns about compensatory evidence, architecture-specific processing burdens, incident-local outages, the incomplete recovery value of replicated state, and the gap between on-time completion and substantive response effectiveness. Those findings motivate the framework. The present article does not report the legacy percentage comparisons as empirical findings or as validation of the framework. Nor does it claim a new full replay or repair of the original architecture simulator. The accompanying implementation is a smaller verification harness for the specified analytical properties (Li J. , 2026).

The distinctions in Table 2 apply throughout the article. A parameter inherited from an uncalibrated prototype remains an assumption, even when it is used in a mathematically exact derivation. An exhaustive state count is a logical result, not a sample from a population. A proposed measurement field does not establish that the corresponding observations have been collected. This evidence classification prevents the application protocol from being mistaken for completed fieldwork.

Table 2 Evidence classes and their use in this article.

Evidence class	**Material**	**Permitted use**	**Not established**
Source-derived context	Research literature and dated public guidance	Define constructs and contextual boundaries	Local processing durations or institutional performance
Development-case evidence	Supplied manuscripts and methodological audit	Identify concrete design and interpretation problems	Independent external validation
Completed analytical work	Proofs, exact enumeration, and deterministic code checks	Establish stated logical properties under specified assumptions	Empirical error rates or architecture rankings
Prospective observations	Proposed exercise and process-record protocol	Specify future calibration and evaluation	Collected data, ethics approval, or demonstrated policy effects

International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal

Indonesian relevance without an assumed national chain of command

Indonesia remains the intended application context. The dated text of Presidential Regulation No. 47 of 2023 includes governance, collaboration, preparedness, and cyber-crisis management; Articles 18–19 address contingency planning and simulation through exercises and role enactment (Republic of Indonesia, 2023). These provisions motivate an exercise-based research route. They do not supply numerical model parameters or validate the specific arrangements proposed here. The interpretation is bounded to the cited text (Wang & Zhang, 2023), not a comprehensive statement that all applicable law or amendments have been reviewed (Idowu, Sydney, Ohunyon, & Adesemoye, 2026).

The development case names BSSN, BIN, TNI Headquarters, the three TNI services, Polri, and KOMDIGI. We preserve that institutional perimeter as the origin of the research question, but use generic functional roles in the analytical examples. Table 3 identifies the questions that must be settled before any role-level model is mapped to a real exercise. A collector in a graph is an information-processing role; it is not evidence that one named institution has authority to command the others.

Table 3 Institutional labels retained from the development case and questions for application.

Development-case label	**Abstract function to examine**	**Question requiring application-specific confirmation**
BSSN	Technical assessment and coordination	Which incident class, escalation step, and coordinating remit apply?
BIN	Intelligence context	Is this contribution required, available, and shareable for the selected decision?
TNI Headquarters	Defense-related decision context	Does the exercise involve a defense responsibility, and what are its limits?
TNI AD, AL, and AU	Operational context	Which affected role participates, rather than assuming all services act in every case?
Polri	Forensic or investigative context	Which evidence or approvals are necessary for this action and stage?
KOMDIGI	Digital-service and sector context	Which service owner, operator, and sector responsibilities must also be represented?
Separate broker	Coordination support	Is the role institutionally feasible, resourced, and accountable?

Actual participation should follow the chosen case rather than a fixed eight-node template. Asset owners, service operators, and relevant sector actors may be essential even though the development prototype abstracted them away. Conversely, a civil service incident need not involve every defense or intelligence role. This study does not prescribe a new national command arrangement, infer a transfer of statutory authority, or claim institutional endorsement.

International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal

The proper role of public cybersecurity data

The original research program considered CISA's Known Exploited Vulnerabilities catalog, FIRST EPSS, NVD, and MITRE ATT&CK (CSIA, 2017). These resources describe exploited vulnerabilities (API, 2015), predicted exploitation, vulnerability metadata, and adversary knowledge, respectively. They can inform technical exercise material or provide traceable scenario attributes. Their availability does not establish inter-organizational approval times, workload arrivals, staffing, decision rights, or correctness of a response (MITRE, 2026).

No vulnerability records are sampled in the completed analytical demonstrations. No download failure is replaced with a synthetic dataset presented as empirical. A later data-enrichment stage should preserve source snapshots, dates, transformations, and limitations, but would still require independent process observations for calibration.

RESULTS AND DISCUSSION

Specify the claim and the full comparison

Let a modeled configuration contain communication topology A, decision protocol P, capabilities K, resource allocation C, failure and recovery specification F, and admissibility specification R. Let X denote the exogenous workload and U the simulator's random inputs. An outcome vector is then written as

$$Y = f(A, P, K, C, F, R; X, U). \quad (1)$$

This notation prevents the architecture label from standing for an unspecified bundle. A topology contrast changes A while keeping other components fixed, where that counterfactual is meaningful and feasible. A protocol contrast changes P on a fixed topology. A capability contrast changes a specified mechanism, such as a message schema or an authorization procedure. A package contrast deliberately changes several components and should be interpreted as a package effect within the model, not as a pure effect of topology.

Before execution, the comparison contract should identify the incident and action, the target outcome, the components changed, the components held fixed, the resource-accounting boundary, the common failure exposure, and the admissibility requirements. It should also identify comparisons that are infeasible. For example, a centralized approval protocol may not be meaningfully implementable under a design whose institutional premise is joint approval. It is better to report a restricted comparison set than to manufacture nominally crossed cells that have no coherent interpretation.

International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal

Capability equality and budget equality are different conditions. Giving all configurations the same number of approval servers holds one input constant, but may leave different work requirements per incident. Doubling servers for every architecture creates a common intervention across architectures; it does not make the one-server and two-server arms equal in resource use. Monetary equivalence requires additional information about staff, infrastructure, and opportunity costs. Without those data, the article should report operational capacities rather than claim an equal-cost comparison.

Define action-specific evidence requirements

For incident i and proposed action a, let M_{ia} be a declared set of mandatory evidence classes. Let $E_i(t)$ contain evidence that has been received and remains acceptable at time t, including the specified provenance, verification, version, and validity conditions. A supplementary support function may aggregate other information. With nonnegative weights summing to one, write

$$S_i(t) = \sum_{k \in E_i(t)} w_k, \qquad w_k \geq 0, \quad \sum_k w_k = 1. \tag{2}$$

Let $L_{ia}(t)$ indicate that the required authorization conditions hold, and let $V_{ia}(t)$ indicate that no specified unresolved contradiction or release-blocking condition is present. The proposed research gate is

$$G_{ia}(t) = \mathbf{1}\{M_{ia} \subseteq E_i(t)\}\, \mathbf{1}\{S_i(t) \geq \theta_{ia}\}\, L_{ia}(t)\, V_{ia}(t). \tag{3}$$

The numerical threshold is optional: an application with no defensible compensatory score can omit that factor and use an explicit evidence rubric. Likewise, the binary authorization term is an interface to an externally reviewed rule, not an algorithmic determination of what the law permits. Requirements must be justified through the action specification and applicable procedures. They cannot be chosen retrospectively to improve one configuration's success rate.

The action index is important. A reversible protective step, forensic collection, public attribution, and a service-restoration decision need not share the same mandatory set. Requiring attribution for every protective action would introduce another unsupported delay. Allowing every action to bypass attribution, ownership, or approval requirements would create the opposite error. The admissibility specification must make the distinction before outcomes are compared.

Preserve temporal validity at approval and release

The first time the evidence portion of the gate qualifies is evidence-ready time. Approval request, approval-service start, approval completion, decision release, instruction receipt, and execution completion are separate events. Their differences identify waiting and service contributions. Information arriving later must not improve the recorded state of an earlier decision.

International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal

A frozen snapshot is necessary for auditability but is not sufficient for release-time validity. Evidence can expire, an asset identifier can be corrected, an approval can be revoked, or a new contradiction can arise during an approval queue. The framework therefore associates approval with an evidence-and-policy version and rechecks the mandatory conditions at release. An unchanged, still-valid snapshot can be released. A changed or expired state must be handled by a declared rule: reapproval, an approved update procedure, escalation, or deferral. The choice is part of the treatment, not an unrecorded implementation detail.

A decision log should retain both the approved snapshot and the state used at release. This avoids two opposite errors: retrospective qualification by later information and release on the basis of invalidated information. The verification harness tests only a conservative example in which a changed signature requires reapproval. It does not establish that this rule is optimal or required for every incident type.

Make service demand and resource allocation visible

For resource pool r, a useful load diagnostic is

$$\rho_r = \frac{\lambda_r \, \mathrm{E}[B_r]}{c_r}, \tag{4}$$

where λ_r is the rate of requests to that pool, B_r its service requirement, and c_r its parallel capacity. An incident arrival rate is not automatically an approval-request rate: routing, missing evidence, and prioritization may transform the stream. The ratio is informative only with its request definition and time basis stated.

A participant-governed process may require parallel endorsements, a serial meeting, a rotating coordinator, or approval only from affected roles. These implementations imply different resource demands. A single serial surcharge should not be assigned to the entire governance category without evidence. Similarly, a shared communication pool and independent per-link resources are different modeling assumptions. Neither is universally correct. Resource pooling must follow the modeled service or be treated as a structural alternative.

Priority discipline also matters. FCFS is useful for a hand-checkable benchmark, but it should not be adopted as an operational cyber-response policy by default. NIST SP 800-61 Rev. 3 emphasizes risk-based triage and prioritization rather than handling constrained incident response simply in arrival order (Nelson, Rekhi, Souppaya, & Scarfone, 2025). A calibrated model should distinguish priorities, interruptions, work already in progress, and the circumstances under which a lower-priority task can be deferred. An execution stage also requires explicit capacity when execution completion is an outcome; otherwise the outcome should stop at the last modeled stage (Kishida, 2025).

International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal

Compare equivalent failure exposure and account for recovery

Physical availability should be represented on a shared timeline. Let $Z_r(t)$ indicate whether resource r is available. An outage interval applies to every incident using that resource at the affected time. Sampling independent availability for each overlapping incident instead represents case-specific access or eligibility, not a physical outage. Either abstraction can be meaningful, but they should not share the same interpretation.

A failure comparison must specify what is lost: a communication endpoint, a coordinator, staff availability, substantive analytical capability, authority, or stored state. Losing a non-analytical broker is not equivalent to losing a node that also produces required evidence. Matched comparisons can disable the same functional capability, or deliberately compare role-specific failures; the latter should remain a package comparison. Required evidence and authorization do not disappear merely because the responsible role becomes unavailable. Substitution requires a specified and defensible rule.

Redundancy evaluation should include both its costs and opportunities to provide benefit. The comparison should specify state replicated, timing and durability of copies, failure detection, authorized takeover, recovery work, and common-mode exposure. A model that charges for backup traffic but never permits recovery from an interrupted incident evaluates replication overhead, not full resilience. Section 5 supplies a minimal analytical identity to make this omission testable without claiming to simulate an entire recovery system.

Table 4 Outcomes and evidence required for their interpretation.

Outcome	**Measurement**	**Interpretive limit**
Evidence readiness	First time the declared evidence rule is met	Does not itself establish authorization or correctness
Approval waiting and service	Request-to-start and start-to-completion intervals	Requires separate timestamps; their sum is not evidence-fusion time
Admissible decision by horizon	Decision meets action-specific requirements at release before a fixed horizon	A rubric-defined process outcome, not automatic cyber containment
Decision correctness	Independent assessment against a declared scenario or operational rubric	Remains unassessed when no adequate reference exists
Instruction and execution completion	Verified receipt and completion of a specified task	Cannot be inferred from merely scheduling a duration
Coordination and recovery workload	Role-time, messages, retries, reapproval, state restoration	Lower workload is not better if mandatory work is omitted

International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal

A single composite effectiveness score is not the primary outcome. Combining speed, legality, information completeness, and containment with arbitrary weights would conceal the very distinctions the framework seeks to preserve. Trade-offs should instead be reported on their natural scales, with any utility function justified separately. For prospective statistical analysis, simulation replications, exercise teams, and individual incidents must not be treated as interchangeable independent units.

Analytical Results and Completed Verification

A scalar threshold can admit missing mandatory evidence

Proposition 1: omission condition. Consider a finite evidence universe with nonnegative weights totaling one. For evidence class j, a weighted threshold can accept a set omitting j if and only if

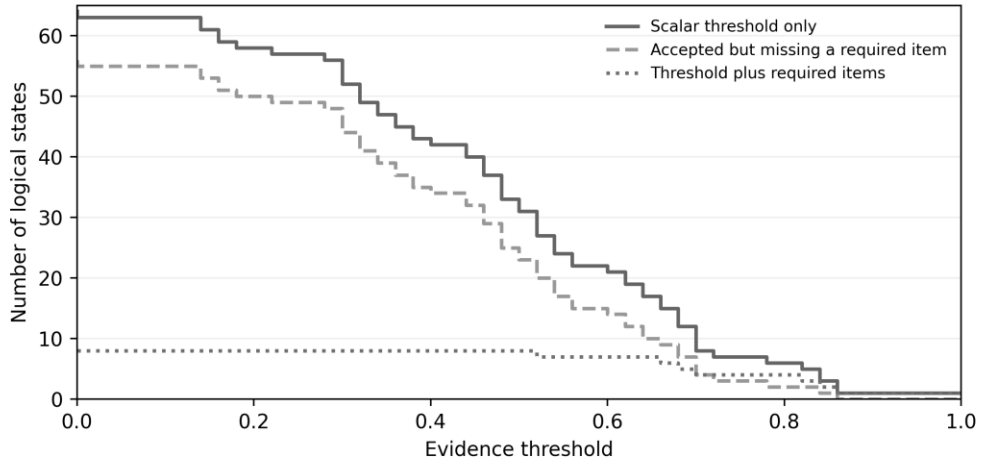
$$\theta \leq 1 - w_j. \tag{5}$$

Proof. The set containing every class except j has weight $1 - w_j$. It therefore supplies an accepted omission whenever the inequality holds. Conversely, nonnegativity implies that every set omitting j has weight at most $1 - w_j$; none can meet a larger threshold. This proves both directions. For a nonempty mandatory set M, a scalar threshold rules out omission of any mandatory item only when $\theta > \max_{j \in M}(1 - w_j)$. Such a threshold may also force collection of optional information. The explicit set condition in Equation (3) expresses the requirement directly.

For a concrete check, we retain the development prototype's six weights: technical 0.22, attribution 0.18, legal/forensic 0.14, asset context 0.16, command context 0.14, and operational context 0.16. At threshold 0.72, the omission condition holds for every individual class. Thus no class is guaranteed solely by the threshold. This is a logical statement about the specification; it does not say how frequently any omission occurs in an operating system.

The new illustrative action specification requires technical, legal/forensic, and asset-context evidence. This set is chosen to demonstrate the method, not asserted as a universal Indonesian requirement. Exhaustive enumeration of all $2^6 = 64$ subsets produces eight threshold-accepted states. Four lack at least one of the illustrative required items. Adding the mandatory-set condition retains four states. For example, technical, attribution, asset context, and operational context sum exactly to 0.72 but omit legal/forensic evidence. By contrast, omitting attribution alone leaves support 0.82 and is admissible under the illustrative set. The distinction demonstrates why a required-item gate is different from simply raising the scalar threshold.

International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal



Picture 1 Exact evidence-state enumeration across thresholds. Counts refer to the 64 possible subsets, with no probability distribution over them

Proposition 2: diagonal non-identification. Suppose two observed simulator means correspond to (A_0, P_0) and (A_1, P_1), with other inputs fixed. Their difference does not determine the topology contrast at fixed protocol P_0 unless the missing counterfactual or an identifying restriction is supplied.

Proof by construction. Let the two observed means be 10 and 8 abstract time units. One completion of the response surface sets $m(A_1, P_0) = 9$, while another sets it to 12. Both reproduce the observed means, but the fixed-protocol topology contrasts are respectively -1 and $+2$. Thus the observed diagonal comparison is compatible with opposite topology effects. These numbers are a constructed example, not fitted results.

Writing $m_{ab} = m(A_a, P_b)$, the algebraic decomposition makes the missing information visible:

$$m_{11} - m_{00} = (m_{10} - m_{00}) + (m_{11} - m_{10}). \quad \textbf{(6)}$$

This identity is not an additive-effects assumption: the second term is evaluated at A_1. A different path may allocate interactions differently. It follows that equal server counts, common random numbers, and narrow confidence intervals do not identify a topology effect when the implemented decision protocols differ. They can still support an accurately labeled package contrast. The verification harness checks the constructed pair of response surfaces without presenting either as a model of actual institutions.

This result also clarifies the status of architecture-specific surcharges in the development case. A serial quorum burden or a fixed synchronization delay is a treatment component, not an empirical discovery. Its consequences may be worth studying, but its interpretation must remain conditional until its size and form are justified. Protocol alternatives should be compared explicitly rather than relabeled as different degrees of collaboration.

International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal

Finite-campaign overload can generate large apparent effects

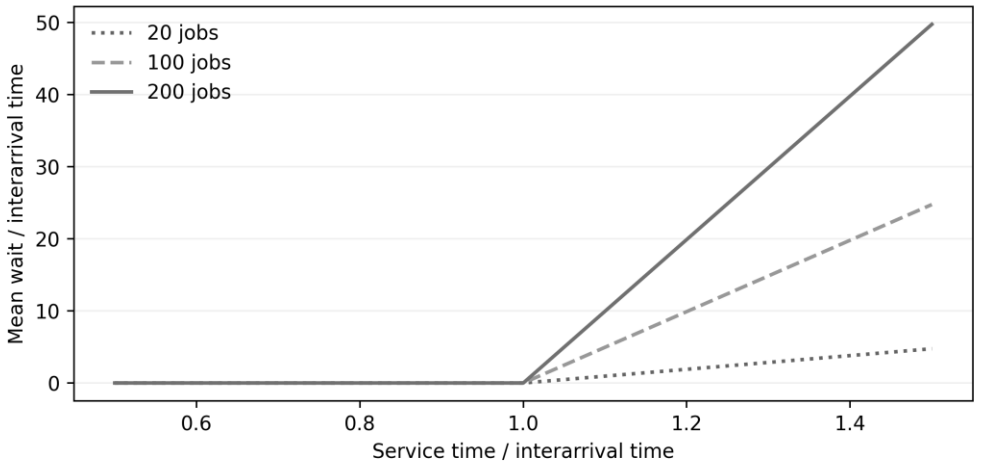
Proposition 3: deterministic queue benchmark. In an initially empty single-server FCFS queue with arrival interval $a > 0$ and constant service time $b \geq 0$, the mean waiting time for $N \geq 1$ jobs is

$$\overline{W}_N = \frac{N-1}{2}\max(0, b-a). \quad (7)$$

Proof. With $W_1 = 0$, the waiting-time recurrence is $W_{n+1} = \max(0, W_n + b - a)$. If $b \leq a$, induction gives zero waiting for all jobs. If $b > a$, induction gives $W_n = (n-1)(b-a)$. Averaging this arithmetic progression yields Equation (7).

In normalized units with $a = 1$ and $N = 100$, service $b = 0.8$ creates no queue, whereas $b = 1.2$ produces mean waiting 9.9 and last-job waiting 19.8. These are exact consequences of the fixture, not empirical timings. The same small change in processing demand can therefore create a large finite-campaign output difference near the capacity boundary. Increasing campaign length further enlarges that difference under overload.

Picture 2 Exact initially empty D/D/1 benchmark for three campaign lengths.



The development prototype is more complex: it has stochastic processing, multiple resources, and transformed request streams. Equation (7) is therefore not a substitute prediction for its archived waiting times. It is a diagnostic for an explanation that attributes large differences to governance without examining service demand and campaign length. A defensible application should report utilization-related quantities, backlog trajectories, initial conditions, priority rules, and sensitivity to observation duration, rather than treating one overloaded finite-campaign mean as a stable organizational characteristic.

A further deterministic check corrects a specific interpretive issue in the development case. Under $q(n) = \lceil 0.75n \rceil$, eight and seven participants both require six qualifying participants. Loss of one participant changes membership and may change traffic or service demand; it does not reduce this required quorum count. The correction illustrates the

International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal

importance of checking a claimed mechanism against the implemented rule rather than inferring it from the direction of a result.

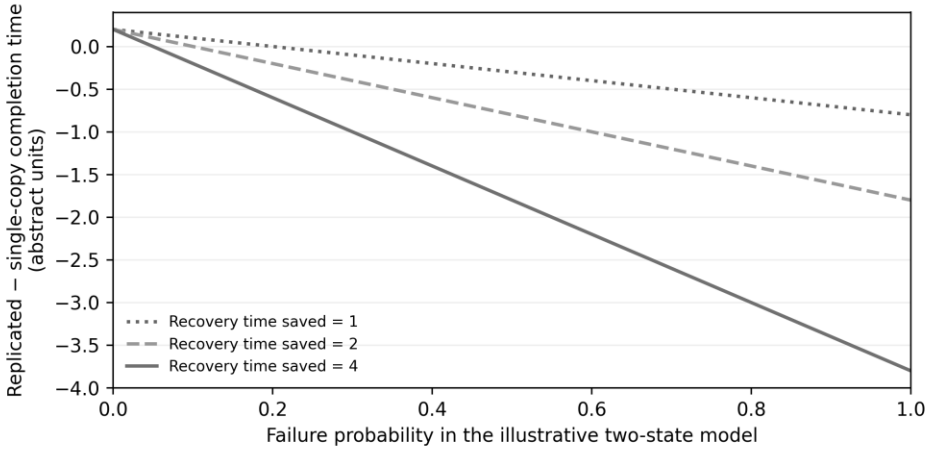
Redundancy requires a recovery opportunity

Proposition 4: minimal recovery accounting. Consider a coupled two-state comparison with common baseline completion time, a constant replication overhead $h \geq 0$, a common failure event of probability p, and a conditional expected recovery saving $\Delta \geq 0$ when that event occurs. Assume no additional queueing, priority, or shared-capacity effects. The expected replicated-minus-single-copy completion difference is

$$\mathrm{E}[T_{\mathrm{rep}} - T_{\mathrm{single}}] = h - p\Delta. \quad (8)$$

Proof. Without failure, the difference is h. With failure, its conditional expectation is $h - \Delta$. Weighting these states gives $(1-p)h + p(h-\Delta) = h - p\Delta$. If $\Delta > 0$, the break-even probability is h/Δ, provided it lies in the feasible interval $[0,1]$. When $p = 0$ or $\Delta = 0$, the model can show only the overhead term.

With illustrative overhead 0.2 and recovery saving 2 abstract time units, differences at probabilities 0, 0.1, and 0.3 are 0.2, 0, and -0.4. No probability is estimated from incidents. The calculation establishes that both favorable and unfavorable outcomes are compatible with the same accounting identity, depending on declared assumptions. It does not demonstrate that a dual-hub governance structure has either outcome in Indonesia.



Exact two-state redundancy accounting with fixed overhead 0.2

Real applications may violate the benchmark assumptions: replication can consume scarce bandwidth, recovery work can compete with other incidents, and both hubs may share a failure cause. These mechanisms require explicit extensions. The methodological requirement is narrower: the evaluation must not imply that it has tested a recovery benefit that its event structure makes impossible (Omomule & Engelbrecht, 2025).

International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal

Verification scope and executable results

The accompanying Python harness completed 21 unit tests with zero failures or errors. It exhaustively enumerated 64 evidence states, checked the omission condition for six evidence classes at 101 integer-percentage thresholds, and compared exact queue recurrence and closed-form results. The omission-condition grid therefore contains 606 class–threshold checks; these are not independent replications. Exported queue and recovery figures each use 303 exact calculation rows. Table 5 summarizes the evidence produced.

Table 5 Completed verification and its scope.

Check	Completed result	What it establishes
Evidence-subset enumeration	64 subsets; 8 scalar-qualified states at 0.72; 4 retained by the illustrative mandatory gate	A checkable construct-validity counterexample
Threshold omission condition	606 class–threshold checks agree with Proposition 1	Exact behavior on the stated grid
Architecture–protocol example	Same observed diagonal, opposite fixed-protocol topology contrasts	Non-identification without the missing comparison
Queue benchmark	Recurrence agrees exactly with the closed form on the tested grids	Implementation agreement for the deterministic limiting case
Recovery benchmark	Exact overhead-minus-recovery identity and sign changes	Complete accounting in the restricted two-state case
Temporal and failure checks	Expiry, future evidence, authority, changed versions, and shared outages handled as specified	Local logic checks, not operational authorization or full recovery simulation

The implementation uses integer evidence weights and rational arithmetic for analytical checks, avoiding rounding ambiguity at 0.72. Figure inputs are exported as machine-readable tables. The code rejects unknown evidence classes and invalid fixture parameters rather than silently substituting values. The notebook contains the same verification implementation and can regenerate the checks without downloading external datasets. It was executed locally; hosted Colab execution is not claimed.

International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal

Prospective Indonesian Application and Evaluation Protocol

Select an observable incident–decision pair

The proposed first application is a controlled exercise concerning a suspected compromise of a public-facing digital service and the decision to approve a bounded protective action. It should use fictional service identifiers and authorized exercise material. The choice narrows the observation boundary: from the first qualifying exercise alert to an assessed decision and, only when represented, completion of the specified protective task. It is not a claim that all cyber-intelligence functions can be validated through one exercise.

Before data collection, the exercise owners should define the action, the relevant actors, evidence requirements, authority, permitted alternatives, escalation conditions, reversibility, and consequences of delay or inappropriate action. Deferral or escalation may be a correct decision in some scenarios and must not automatically be coded as failure. In others, unjustified inaction may be a substantive error. The assessment rubric should make these distinctions independently of which governance configuration is being tested.

NIST SP 800-84 offers guidance on designing, conducting, and evaluating test, training, and exercise activities (Grance, et al., 2006). Its stated scope is primarily single-organization events, so the cross-organizational procedures below are a proposed adaptation, not a claim that the publication supplies a ready-made multi-agency evaluation design. The NIST CSF 2.0 provides a complementary outcome vocabulary. Neither source determines a sample size or supplies local processing-time distributions (National Institute of Standards and Technology, The NIST Cybersecurity Framework (CSF) 2.0; NIST CSWP 29; NIST: Gaithersburg, MD, USA, 2024, 2024).

Establish the workflow before calibrating durations

Practitioner challenge should precede numerical calibration. Participants should be able to dispute a proposed collector, reject a serial consensus rule, identify an unrepresented asset owner, or distinguish mandatory legal review from optional forensic enrichment. The resulting workflow should specify role-level requests and services, not merely institutional boxes. Every unresolved substantive disagreement should be recorded as an alternative specification or a limitation rather than averaged into a single assumed parameter.

The four-pillar transformation agenda retained from the development research—strategic coordination, joint operational protocols, shared technology infrastructure, and collaborative capability development—provides an organizing structure for this stage. Each pillar should be linked to an observable intervention and evidence requirement, as in Table 6. An organizational ambition becomes testable only when the procedures and measurements affected by it are specified (Aladini, Almashikhi, & Khalifa, 2026).

International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal

Table 6 Operationalizing the four-pillar agenda without presuming effectiveness.

Pillar	**Candidate intervention**	**Required evidence**
Strategic coordination	Explicit coordinator, escalation path, and accountable approval role	Role validation, decision-right record, escalation trace
Joint operational protocols	Shared evidence definition and action-specific admissibility rules	Requirement rubric, clarification records, omitted-item assessments
Shared technology infrastructure	Authorized exchange, version control, and recoverable state	Delivery traces, access failures, state-restoration and resource records
Collaborative capability development	Training or rehearsal of the selected protocol	Observed predeclared skill/process measures and a defensible comparison design

Training effects are not simulated as an automatic rise in trust. They require their own study, with measures appropriate to the intended skill or process. Similarly, greater message volume may reflect improved inclusion or redundant work; its interpretation requires the underlying purpose of those messages, not only a count (Avlona & Shklovski, 2025).

Collect process traces and independent decision assessments

The minimum record should distinguish incident, scenario, team, role, event, and resource identifiers. For each analysis and approval task, record request, service start, and completion. For information exchange, record sender, recipient, class, version, receipt, clarification, and validity interval. For decisions, retain the proposed action, required evidence, assessed admissibility, approval basis, approved version, release time, and independent correctness assessment where feasible. Outages and resource availability require shared calendar records, not separate retrospective guesses for each case.

Timestamps require a common clock definition and a stated resolution. Apparent negative waiting times should trigger data review, not automatic correction. Missing data should be distinguished from events that did not occur, actions that were not applicable, and values withheld for security reasons. A case without an observed correctness assessment remains unassessed; it is not assigned a favorable outcome to complete a dataset. Supplementary Material S1 supplies empty collection templates and a proposed data dictionary. No populated template is presented as a collected dataset.

Decision assessors should use a predeclared rubric and, where practicable, review de-identified decisions without the configuration label. Multiple assessors can document disagreement and adjudication. Agreement is useful evidence about the measurement

International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal

instrument, but cannot establish the truth of an ambiguous scenario. The rubric should distinguish procedural admissibility, substantive appropriateness, and execution evidence rather than collapse them into one subjective overall score.

Separate pilot, calibration, and held-out evaluation

A pilot should test the workflow, recording burden, measurement resolution, scenario difficulty, and feasibility of the comparison. It can inform the eventual sample size and a smallest practically relevant difference. It must not be presented as an untouched confirmatory sample after the protocol has been adapted using its outcomes. The final design should freeze scenario definitions, primary outcomes, treatment components, exclusions, and analysis before evaluating the held-out trials.

Where teams experience multiple conditions, assignment order and scenario variants should be counterbalanced to reduce learning and fatigue effects. Repeating identical injects can reveal solutions, so equivalent variants require explicit construction and review. The unit of assignment, the unit of observation, and the unit of statistical inference should be stated separately. Many messages generated by one team do not supply many independent team-level trials. A feasible design may compare two carefully specified protocols before attempting all four configuration families.

Calibration should estimate stage-specific quantities from the calibration sample, preserve dependencies where supported, and distinguish measured distributions from elicited judgments. A small dataset does not justify selecting a complex distribution simply because it fits the same observations closely. Candidate models should be compared using held-out predictive performance and process diagnostics. Exercise-derived estimates remain exercise-derived; transfer to operational settings requires a further argument and additional evidence.

Prespecify outcomes and a validation decision

A candidate primary exercise outcome is the proportion of assigned team–scenario trials with a rubric-admissible decision by a task-specific horizon. Substantive correctness and unjustified action should be reported separately. The horizon must follow the exercise decision requirements, not be adjusted until completion rates look plausible. Secondary outcomes include readiness attainment, approval waiting, reapproval, clarification, workload, and recovery after a specified interruption. Results should show the distribution and non-attainment pattern rather than only a grand mean.

For a paired design, effects should be estimated at the level at which pairing and assignment occur. In a later simulator, repeated random seeds quantify Monte Carlo error conditional on the selected model. Those intervals must remain distinct from uncertainty about measured inputs, model form, and transfer to the application. No universal number of replications, exercise participants, or incidents is prescribed here. Adequacy depends on the target estimand, clustering, variation observed in the pilot, and the required precision (Saleem & Nazar, 2026).

International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal

A held-out validation plan should assess stage-level timing errors, bias, interval coverage when intervals are predicted, outcome calibration when probabilities are predicted, and discrepancies by scenario type. Acceptance thresholds should be set before the held-out outcomes are examined and should reflect the intended decision use. Matching an overall mean while systematically failing interruption or high-workload scenarios is insufficient. Failure of a criterion should result in restricted applicability, model revision followed by new evaluation, or withholding the relevant claim—not silent retuning on the holdout.

Security, ethics, and implementation status

Any practitioner study requires the relevant institutional, ethics, and information-security review before collection. Authorization to participate does not automatically authorize release of organizational process data. Role labels and exact timestamps can still reveal sensitive patterns; de-identification, coarsening, and access controls should follow a documented disclosure review. Public release may therefore include schemas, code, and approved aggregate or synthetic fixtures while restricting actual process records (Zilong & Alobaedy, 2025).

No practitioner recruitment, interviews, exercise observation, operational logging, ethics approval, calibration, or held-out validation was completed for this article. These activities are specified as a prospective protocol. The verification notebook cannot replace them. Its purpose is to make the minimum model properties inspectable before resources are committed to the application study.

Discussion

The contribution is a defensible comparison, not a preferred architecture

The framework changes the question from which organizational label has the highest simulated score to which intervention is being compared, under what decision requirements, with what supporting evidence. This aligns the computational treatment with the institutional question. A shared graph does not imply shared authority. A shared budget does not imply equal work. A shared failure probability does not imply equal loss. These distinctions explain why a seemingly sophisticated simulation can remain weak evidence for a governance recommendation.

The analytical demonstrations make the argument concrete. Proposition 1 identifies the exact condition under which a score can bypass required information. Proposition 2 shows why a bundled diagonal comparison cannot isolate topology. Proposition 3 shows how a small service-demand difference can produce large campaign-dependent delays. Proposition 4 shows why redundancy needs a modeled opportunity to recover value. None proves an architecture superior. Together they define checks that an architecture argument should survive before its numerical ranking is interpreted.

International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal

The contribution is primarily methodological integration and operationalization. Threshold logic, counterfactual completion, deterministic queue recurrences, and conditional expectation are established mathematical tools. Their application here provides a reviewable specification connecting governance constructs, model boundaries, measurement, and computation. Claims of a new general theory of collaborative governance, a new optimization algorithm, or an empirical national performance model would exceed this contribution.

Why the original subject remains worth investigating

Removing unsupported rankings does not remove the research problem. Cyber-intelligence coordination still requires a practical account of how information, responsibility, and action interact under pressure. The Indonesian application remains relevant because the intended research concerns heterogeneous institutional functions and a four-pillar transformation agenda. The revised approach makes those commitments measurable instead of assigning them benefits in advance.

A useful empirical result might show that a simpler approval route reduces waiting without degrading admissibility. Another might show that broader participation catches essential contextual errors while delaying a decision. A recovery test might find that replication pays for itself only during specified interruptions. These are examples of possible findings, not predicted or completed results. The framework is designed to admit unfavorable and mixed findings as well as favorable ones.

A methodological study can be useful before field calibration when its conclusions stay at the level supported by its reasoning and verification. It can reveal that an outcome fails to measure the intended construct or that the comparison does not identify the proposed mechanism. That is different from presenting uncalibrated timing values as estimates and surrounding them with cautionary language. The present contribution is the former.

Reproducibility and responsible use

The package includes the manuscript source, exact verification code, an executed local notebook, figure inputs, collection templates, and a provenance manifest. FAIR principles motivate explicit identifiers, metadata, and machine-readable materials (Wilkinson et al., 2016), but no claim of complete FAIR certification or public repository accession is made. The versioned implementation and its declared scope allow another researcher to inspect the checks without confusing them with an operational tool.

A future extension should retain this separation. New observational data should have their own lineage and access conditions. Changes to mandatory requirements, protocols (Rangini, 2025), or failure semantics should trigger a model-version change and renewed verification (Yilmaz, 2026).

International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal

Limitations

The framework has not been evaluated through practitioner use. Its clarity, feasibility, and usefulness in improving model-development decisions remain empirical questions. The development case comes from one research program and should not be treated as a representative sample of all cyber-governance simulations (Hlazunov, 2024). The focused literature synthesis does not establish that every relevant approach has been reviewed or that no similar methodological synthesis exists (Udrescu, Suduc, Udrescu, & Bizoi, 2026).

The admissibility rule is only as defensible as its requirements and evidence predicates. A formally correct gate can implement an inappropriate rule. The illustrative mandatory set and inherited weights are not calibrated, legally established, or recommended for deployment. Evidence reliability, adversarial deception, strategic withholding, conflicting organizational objectives, and the cost of erroneous action are not modeled by the verification harness. Its signatures and validity flags are logical fixtures, not secure implementations of identity, access control, or digital signatures (Cao, 2026).

The queue and redundancy benchmarks intentionally omit the complexity of the full system. They establish exact results for limiting cases, not numerical approximations whose operational accuracy has been validated. The recovery identity excludes endogenous failure exposure, queue interactions, common-mode effects, and shared recovery resources. The topology non-identification example establishes that missing counterfactuals matter; it does not show that every feasible institutional comparison can be fully crossed.

Developer checks do not provide independent replication. The code verifies the implemented gates and analytical fixtures, not all behaviors of the legacy simulator, and the full architecture model has not been rebuilt in this revision. No new policy-effect percentages or field performance estimates are reported. The prospective protocol likewise offers no assurance that data access, recruitment, or ethical approval will be obtained (Barilla, Caristi, & Ciano, 2024).

Finally, contextual use of a dated regulation is not a comprehensive legal assessment. Before application, the relevant institutions must review the incident class, participating actors, current responsibilities, data-sharing constraints, and decision rights. The study therefore supports a disciplined route to evidence rather than an immediate institutional prescription (Thompson, et al., 2025).

CONCLUSION

This study concludes that credible evaluation of cyber-intelligence governance depends fundamentally on the validity of the comparison rather than on the sophistication of simulation outputs alone. Executable models, precise averages, and narrow numerical intervals cannot establish the superiority of a governance arrangement when the treatment itself remains insufficiently specified or when different organizational mechanisms are

International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal

bundled into a single comparison. The validation-first framework developed in this study addresses this problem by requiring explicit definitions of communication topology, decision protocol, operational capability, resource allocation, failure and recovery conditions, and action-specific admissibility requirements. This approach enables researchers to distinguish between what a computational model demonstrates logically and what can legitimately be interpreted as evidence about institutional performance. The analytical verification further demonstrates four important sources of potential misinterpretation. First, a compensatory evidence score may accept a decision despite the absence of information that is explicitly mandatory for the action. Second, a comparison between bundled configurations may reproduce observed differences without identifying the independent contribution of topology. Third, finite campaign overload can generate substantial timing differences when service demand approaches capacity, making an apparent governance effect partly attributable to resource utilization and campaign duration. Fourth, redundancy cannot be evaluated fairly when replication costs are modeled but the opportunity to obtain recovery benefits is excluded. These findings do not constitute evidence of operational failure or institutional effectiveness; rather, they establish methodological conditions under which governance claims should be interpreted cautiously.

The completed deterministic verification strengthens the auditability of these conclusions. The evidence-state enumeration, threshold checks, queue benchmark, recovery identity, temporal-validity checks, and failure specifications demonstrate that the proposed framework can convert abstract concerns about model credibility into inspectable computational properties. At the same time, the study deliberately recognizes the limits of its evidence base. No practitioner observations, operational incidents, empirical calibration, or held-out validation were completed. Consequently, the Indonesian application protocol should be understood as a structured pathway for future empirical investigation rather than as evidence supporting a preferred national governance architecture. Its principal contribution is therefore methodological: it establishes a disciplined bridge between governance theory, computational representation, measurement, and institutional evidence. This distinction is particularly important for interdisciplinary research because it prevents synthetic precision from being mistaken for empirical certainty. The framework ultimately supports a more defensible research trajectory in which institutional recommendations emerge only after the treatment, decision requirements, resource conditions, failure exposure, and relevant outcomes have been explicitly defined, measured, challenged, and validated.

Recommendations

Future research should prioritize empirical validation of the proposed framework through controlled and ethically authorized cross-organizational exercises rather than immediately attempting to rank alternative cyber-intelligence governance architectures. The first recommendation is to define an observable incident–decision pair before collecting data. The selected exercise should specify the incident context, proposed action, participating functional

International Journal of Interdisciplinary Cultural Studies
ISSN: 2327-008X (Print), ISSN: 2327-2554 (Online)
Volume 21, Issue 2, 2026
https://cgscopus.com/index.php/journal

roles, mandatory evidence, authorization conditions, escalation procedures, reversibility, and consequences of delay or inappropriate intervention. This requirement is essential because different actions may legitimately require different evidence and approval conditions. A reversible protective measure, forensic activity, attribution decision, and service-restoration action should not automatically be evaluated using an identical admissibility rule. The Indonesian application should therefore preserve the distinction between evidence readiness, procedural authorization, substantive correctness, and execution completion.

Second, researchers and participating institutions should establish the workflow collaboratively before assigning numerical parameters. Practitioner challenge should be incorporated into protocol development so that participants can identify missing roles, inappropriate approval sequences, unrealistic resource assumptions, or unsupported information requirements. Process traces should then record requests, service starts, completions, information versions, clarification events, validity intervals, approval states, release times, resource availability, outages, and recovery activities. These records should use a common clock and clearly distinguish missing observations from events that did not occur. Independent assessors should evaluate decisions using a predeclared rubric, preferably without knowledge of the tested governance configuration, so that procedural admissibility and substantive appropriateness remain analytically separate.

Third, calibration and validation should proceed sequentially. A pilot exercise should assess measurement feasibility and recording burden, while calibration data should be used to estimate stage-specific processing and workload characteristics. The final treatment definitions, primary outcomes, exclusions, and analytical procedures should subsequently be frozen before evaluation using held-out trials. Statistical inference should respect the actual unit of assignment and observation rather than treating messages, incidents, and teams as interchangeable independent observations. Simulation replications should also be distinguished from uncertainty in empirical inputs and uncertainty concerning model structure.

Four, Indonesian implementation should integrate security, ethics, legal review, and data-governance safeguards from the beginning. Institutional participation should not automatically imply permission to disclose operational records. Sensitive timestamps, role information, and workflow patterns require controlled access, de-identification, and documented disclosure review. Future studies should also report mixed or unfavorable findings without privileging a predetermined architecture. The four transformation pillars—strategic coordination, joint operational protocols, shared technology infrastructure, and collaborative capability development—should each be translated into observable interventions and measurable outcomes rather than assumed benefits. Most importantly, researchers should refrain from converting the current analytical demonstrations into claims of national institutional effectiveness. The appropriate next step is empirical testing that

determines whether the framework improves comparison quality, measurement validity, and decision-making under realistic conditions. In this way, the Indonesian protocol can develop from a methodological specification into evidence capable of supporting responsible and context-sensitive cyber-intelligence governance recommendations.

ACKNOWLEDGEMENTS

The author would like to express sincere gratitude to the Rector of Universitas Muhammadiyah Jakarta, Prof. Dr. Ma'mun Murod; the Dean of the Faculty of Social and Political Sciences, Universitas Muhammadiyah Jakarta, Prof. Dr. Evi Satispi, S.Pd., M.Si.; and the Head of the Doctoral Program in Public Administration, Prof. Dr. Andriansyah, M.Si., for their valuable support, guidance, and encouragement throughout the research process. The author greatly appreciates the opportunities, direction, and continuous guidance provided from the initial stage to the completion of this research, The author would like to express sincere gratitude to the Promoter, Prof. Dr. Suardika, for his valuable guidance, direction, and advice throughout the research and writing process.

REFERENCES

Aladini, A., Almashikhi, K., & Khalifa, F. (2026). Educational leadership practices for emergency preparedness decision-making and coordination in conflict-affected school systems. *3(1)*, 158–187. doi:10.1108/qea-02-2026-0068

Ansell, C., & Gash, A. (2008). Collaborative governance in theory and practice. J. *Public Adm. Res. Theory, 18*, 543–571. doi:10.1093/jopart/mum032

Avlona, N., & Shklovski, I. (2025). The logics of compliance in data quality: Challenges in creating training data for medical AI systems. doi:10.2139/ssrn.5158998

Barilla, D., Caristi, G., & Ciano, T. (2024). Cost-benefit risk analysis modeling for corporate compliance: evidence from Italy obtained through investment and industry 4.0 tax credit data analysis. *58(5)*, 4459–4478. doi:10.1007/s11135-024-01865-2

Cao, Y. (2026). Digital Accounting Empowers Corporate Governance: Research on the Synergistic Effect of Data-Driven Decision-Making on Budget and Cost Control. *9(1)*. doi:10.30560/jems.v9n1p20

Cybersecurity and Infrastructure Security Agency. (n.d.). Known Exploited Vulnerabilities Catalog. *Available online: https://www.cisa.gov/known-exploited-vulnerabilities-catalog. Contextual source; no catalog records were sampled for this study.* Retrieved from https://www.cisa.gov/known-exploited-vulnerabilities-catalog

Dunn Cavelty, M., & Wenger, A. (2020). Cyber security meets security politics: Complex technology, fragmented politics, and networked science. *Contemp. Secur. Policy, 41*, 5–32. doi:10.1080/13523260.2019.1678855

Egbedayo, O. (2026). Measuring Compliance Latency: A Cross-Sector Analysis of Verification Friction, Institutional Trust Gaps, and the Structural Cost of Compliance Duplication. doi:10.2139/ssrn.6623419

Emerson, K., Nabatchi, T., & Balogh, S. (2012). An integrative framework for collaborative governance. J. *Public Adm. Res. Theory, 22*, 1–29. doi:10.1093/jopart/mur011

FIRST. (2026). Exploit Prediction Scoring System (EPSS). *Available online: https://www.first.org/epss/ (accessed on 15 September*. Retrieved from https://www.first.org/epss/

Fu, Z., Sun, H., Xu, C., & Yue, M. (2025). Downstream Information Sharing in Agricultural Supply Chains with Investment Spillovers. *22(2)*, 147–168. doi:10.1287/deca.2024.0247

Grance, T., Nolan, T., Burke, K., Dudley, R., White, G., & Good, T. (2006). *Guide to Test, Training, and Exercise Programs for IT Plans and Capabilities; NIST SP 800-84; NIST: Gaithersburg, MD, USA, 2006.* doi:10.6028/NIST.SP.800-84

Grimm et al. (2020). The ODD protocol for describing agent-based and other simulation models: A second update to improve clarity, replication, and structural realism. J. *Artif. Soc. Soc. Simul, 23*, 7. doi:10.18564/jasss.4259

Grünig, R., & Kühn, R. (2025). Problem Verification and Analysis. 55–64. doi:10.1007/978-3-662-71620-5_7

Hlazunov, A. (2024). DEVELOPMENT OF BAYESIAN NETWORKS FOR A DECISION SUPPORT SYSTEM DURING INTERNAL CYBER THREATS ANALYSIS. *1(25)*, 103–117. doi:10.28925/2663-4023.2024.25.103117

Idowu, A., Sydney, E., Ohunyon, G., & Adesemoye, A. (2026). Conceptual Advances in AI-Enabled Compliance and Coordination Models for National Emergency Supply Chain Preparedness. *7(4)*, 709–721. doi:10.54660/.ijmrge.2026.7.4.709-721

Ikhalia, J. (2026). The Three Problems of Cyber Regulatory Compliance: Translation, Institutional Continuity, and Evidence Quality: a position paper towards an evidence-led discipline. doi:10.2139/ssrn.7186918

Imran, I. (2026). The Asymmetric Cost of Secrecy: A Bayesian Analysis of Governance, Information Gaps, and Extreme Investment Decisions in Baghdad. *18(2)*, 14–28. doi:10.5296/ajfa.v18i2.19017

Johnson, C., Badger, M., Waltermire, D., Snyder, J., & Skorupka, C. (2016). *Guide to Cyber Threat Information Sharing; NIST SP 800-150; NIST: Gaithersburg, MD, USA, 2016.* doi:10.6028/NIST.SP.800-150

Katiku, S. (2023). Quality Internal Auditing and Management Decision Making in Kenya. *3(2)*, 96–100. doi:10.59413/ajocs/v3.i2.2

Kishida, M. (2025). Risk-Aware Safety Verification and Robustness Analysis of Neural Network. 6951–6956. doi:10.1109/cdc57313.2025.11312594

Kommaraju, R. (2025). Optimizing Data Quality and Compliance Through Integrated Validation Strategies for Clinical Systems. *7*, 299–306. doi:10.37547/tajet/volume07issue08-26

Lee, T. (2024). Integrated Cyber Security Risk Management-Insurance and Investment Cost Analysis. *16(3)*. doi:10.1504/ijdats.2024.10063783

Li, J. (2026). Algorithmic Auditing for Employee Health: Bridging the Verification Gap in GRI 403 Compliance via Topological Data Analysis. doi:10.2139/ssrn.5992894

Li, Z., Zhu, Y., Zhao, L., & Zhao, Y. (2024). Safety decision of autonomous driving behavior based on Bayesian network and its verification method. 145–153. doi:10.1109/qrs-c63300.2024.00029

MITRE. (2026). ATT&CK Data and Tools. *Available online: https://attack.mitre.org/resources/attack-data-and-tools/ (accessed on 15 September*. Retrieved from https://attack.mitre.org/resources/attack-data-and-tools/

Monks, T., Currie, C., Onggo, B., Robinson, S., Kunc, M., & Taylor, S. (2019). Strengthening the reporting of empirical simulation studies: Introducing the STRESS guidelines. J. *Simul, 13*, 55–67. doi:10.1080/17477778.2018.1442155

National Institute of Standards and Technology. (2024). *The NIST Cybersecurity Framework (CSF) 2.0; NIST CSWP 29; NIST: Gaithersburg, MD, USA, 2024.* doi:10.6028/NIST.CSWP.29

National Institute of Standards and Technology. (n.d.). National Vulnerability Database: Vulnerability APIs. *Available online: https://nvd.nist.gov/developers/vulnerabilities. Contextual source; no NVD records were sampled for this study*. Retrieved from https://nvd.nist.gov/developers/vulnerabilities

Nelson, A., Rekhi, S., Souppaya, M., & Scarfone, K. (2025). *Incident Response Recommendations and Considerations for Cybersecurity Risk Management: A CSF 2.0 Community Profile; NIST SP 800-61 Rev. 3; NIST: Gaithersburg, MD, USA, 2025.* doi:10.6028/NIST.SP.800-61r3

Omolere, O. (2026). AI and Machine Learning in Auditing: Enhancing Cybersecurity Risk Management through Real-Time Monitoring, Anomaly Detection, and Predictive Analysis. doi:10.2139/ssrn.5968654

Omomule, T., & Engelbrecht, A. (2025). Sensitivity Analysis of Neural Network Ensembles to Data Quality Issues. 1–8. doi:10.1109/acdsa65407.2025.11165967

Pardo, T., Nam, T., & Burke, G. (2012). E-government interoperability: Interaction of policy, management, and technology dimensions. *Soc. Sci. Comput. Rev, 30*, 7–23. doi:10.1177/0894439310392184

Petukh, A., Bashkatov, V., & Shvyreva, O. (2023). Objectification and verification of financial information. *10(5)*, 22–37. doi:10.26794/2408-9303-2023-10-5-22-37

Provan, K., & Kenis, P. (2008). Modes of network governance: Structure, management, and effectiveness. J. *Public Adm. Res. Theory, 18*, 229–252. doi:10.1093/jopart/mum015

Rangini, V. (2025). Reinventing Data Trust: A Metadata-Driven Framework For Lineage, Quality, And Compliance. doi:10.2139/ssrn.5907542

Republic of Indonesia. (2023). Peraturan Presiden Republik Indonesia Nomor 47 Tahun 2023 tentang Strategi Keamanan Siber Nasional dan Manajemen Krisis Siber; Jakarta, Indonesia, 2023. *Available online: https://peraturan.go.id/files/perpres-no-47-tahun-*. Retrieved from https://peraturan.go.id/files/perpres-no-47-tahun-2023.pdf

Saleem, S., & Nazar, R. (2026). The Impact of Uncertainty and Institutional Quality on Economic Growth: Evidence from Cross-Country Analysis. *3(2)*, 136–151. doi:10.69671/socialprism.3.2.2026.105

Sargent, R. (2010). Verification and validation of simulation models. Retrieved from https://www.informs-sim.org/wsc10papers/016.pdf

Thompson, Y., Li, Y., Kang, Z., Miller, M., Wurgler, R., & Silovsky, J. (2025). Application of the Data Quality Framework to Administrative Data on Child Maltreatment. doi:10.2139/ssrn.5152152

Udrescu, E., Suduc, A.-M., Udrescu, A., & Bizoi, M. (2026). A Bayesian Probabilistic Decision Support System for Cloud Cost Governance in AI-Augmented Agile Software Development. doi:10.2139/ssrn.6826891

Van den Oord, S., Kenis, P., Raab, J., & Cambré, B. (2023). Modes of network governance revisited: Assessing their prevalence, promises, and limitations in the literature. *Public Adm. Rev*, *83*, 1564–1598. doi:10.1111/puar.13736

Wang, S., & Zhang, H. (2023). Institutional Logics of Investment Decision: Analysis of Angel Investment Strategy Based on ML. *2023(1)*. doi:10.5465/amproc.2023.17624abstract

Wilkinson et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data*, *3*, 160018. doi:10.1038/sdata.2016.18

Yilmaz, S. (2026). Data Quality and Utility Label in the EHDS: Building Europe's Health Data Future Beyond Legal Compliance. 1–5. doi:10.1109/dph70309.2026.11650249

Zilong, D., & Alobaedy, M. (2025). Blockchain-Based Healthcare Data Management: Analysis and Evaluation of Security, Scalability, and Compliance for Electronic Health Records (EHRs). 1–7. doi:10.1109/eect64505.2025.10966949