International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(13s) 2026 pp. 726-736
https://doi.org/10.70917/ijcisim-2026-4116
Article

Algorithmic Sovereignty: A Multidimensional Conceptual Framework for National Cyber-Defense Posture in Emerging Economies

Raja Humuntal Manalu [1,*], **Siswo Hadi Sumantri** [1], **Suyono Thamrin** [1], **Pujo Widodo** [1], **Richardus Eko Indrajit** [2]

[1] Indonesia Defense University, Bogor, Indonesia

[2] Pradita University, Tangerang, Indonesia

E-mail: raja.manalu@doktoral.idu.ac.id, siswohs@idu.ac.id , suyono.thamrin@idu.ac.id, pujowidodo78@gmail.com, eko.indrajit@pradita.ac.id.

Abstract: The accelerating embedding of artificial intelligence (AI) into national cyber defense is quietly reconfiguring what it means to be sovereign in cyberspace. Prevailing conceptions of cyber and data sovereignty privilege control over data and physical infrastructure, yet they largely bypass the algorithmic layer—the foundation models, inference pipelines, and machine-learning systems on which contemporary defensive operations increasingly rely. This paper develops algorithmic sovereignty as a multidimensional construct denoting a state's capacity to exercise autonomous and meaningful control over the AI systems embedded across its cyber-defense lifecycle. Integrating the resource-based view, dynamic-capabilities theory, securitization theory, dependency theory, and the literature on cyber power, we theorize six interdependent dimensions—infrastructural, data, model/algorithmic, operational, capability, and governance/normative sovereignty—and cross them with a five-stage maturity continuum to produce a diagnostic matrix. We advance five testable propositions, including the claim that data localization without model-layer control yields a sovereignty illusion, and that absorptive capacity moderates the dependency–vulnerability relationship. Because the contribution is conceptual, we specify a four-stage validation protocol—systematic review, Delphi-based content validation, multi-criteria weighting, and illustrative national instantiation—and operationalize each dimension into a candidate indicator bank suitable for empirical testing. The framework is bounded to middle-power and emerging-economy contexts, where an autonomy–capability trade-off is most acute. We discuss theoretical and policy implications using Indonesia as an illustrative case and outline an agenda for empirical validation.

Keywords: algorithmic sovereignty; cyber sovereignty; cyber defense; foundation models; emerging economies; resource-based view

1. Introduction

Artificial intelligence has moved from the periphery of cybersecurity into its operational core. Machine-learning classifiers triage alerts, large language models summarize incidents and draft response playbooks, and increasingly autonomous agents are trusted to enrich threat intelligence and, in some settings, to initiate containment actions. As these capabilities migrate into the defensive stacks of states, a structural question arises that the existing sovereignty literature is poorly equipped to answer: who actually controls the algorithms on which national defense now depends? For most countries—and acutely for emerging economies—the honest answer is that the decisive components are designed, trained, hosted, and governed elsewhere.

The dependency is not incidental. The frontier of AI is concentrated among a small number of foreign firms and jurisdictions that control the compute, the training data, and the model weights that define what defensive AI can do. Critically, the most capable systems are typically delivered as closed-weight services: their parameters cannot be inspected, their behavior cannot be independently audited, and access to them can be throttled, conditioned, or withdrawn by the provider or by the provider's home government. A defensive posture built on such systems inherits

a vulnerability that no amount of perimeter hardening can remediate, because the vulnerability is not in the network but in the supply relationship itself.

The scholarly and policy vocabulary that states reach for in response—cyber sovereignty, digital sovereignty, data sovereignty—was, however, largely formed before foundation models reshaped the technological landscape. That vocabulary tends to localize the problem at the level of data (where it is stored, under whose jurisdiction it falls) and infrastructure (who owns the cables, the data centers, the chips). These are necessary concerns, but they are no longer sufficient. A state can satisfy every data-localization mandate and still run its security operations center on a foreign model whose weights it cannot see and whose availability it cannot guarantee. Control over the substrate has been mistaken for control over the system.

This paper argues that the missing construct is sovereignty over the algorithmic layer itself, and it sets out to theorize that construct rigorously. We define algorithmic sovereignty as the capacity of a state to exercise autonomous and meaningful control over the AI and algorithmic systems embedded across its cyber-defense lifecycle—spanning data, compute, models, deployment, and governance—such that defensive capability is neither contingent upon, nor coercible through, foreign-controlled algorithmic dependencies. The construct is positioned as an upstream extension of established sovereignty thinking rather than a replacement for it.

Three observations motivate the contribution. First, the literature on digital and cyber sovereignty is conceptually rich but notoriously imprecise, and it has only recently begun to register the algorithmic layer as a distinct object of control [14,15,16,17]. Second, what fragments of model-layer thinking do exist are scattered across disconnected strands—constructivist discourse analysis, staged policy roadmaps, and narrowly military treatments—none of which offers an integrated, operationalizable framework. Third, almost all of this work is framed from the vantage point of technologically advanced producers; the situation of middle-power and emerging economies, which are predominantly consumers of foreign algorithms, is undertheorized despite being where the dependency bites hardest.

Accordingly, the paper addresses three research questions. (RQ1) What are the constitutive dimensions of algorithmic sovereignty in the cyber-defense domain? (RQ2) How can these dimensions be organized into a coherent, theoretically grounded framework that distinguishes maturity levels and supports diagnosis? (RQ3) How can the framework be operationalized and validated for empirical use in emerging-economy contexts?

In answering these questions the paper makes four contributions. It introduces and defines algorithmic sovereignty as a multidimensional construct and locates it relative to adjacent concepts. It develops a six-dimension, five-stage framework—rendered as a diagnostic matrix—grounded in an explicit synthesis of the resource-based view, dynamic-capabilities theory, securitization theory, dependency theory, and cyber-power scholarship. It advances five testable propositions that render the framework falsifiable and connect it to defensive outcomes. Finally, it specifies a concrete validation protocol and a candidate indicator bank, so that the framework is a platform for measurement rather than a closed argument. The remainder of the paper reviews the theoretical background (Section 2), states the research approach for conceptual theory building (Section 3), develops the framework (Section 4), operationalizes it (Section 5), specifies validation (Section 6), discusses implications (Section 7), and acknowledges limitations (Section 8) before concluding (Section 9).

2. Theoretical background

2.1 From cyber sovereignty to the algorithmic layer

Sovereignty in the digital domain is an essentially contested concept. The umbrella terms digital, cyber, technological, and data sovereignty are frequently used interchangeably to denote a claim by a state, organization, or community to meaningful control over the digital infrastructures, data, and software on which it depends [14,15]. Scholars have repeatedly observed that the concept is invoked far more often than it is defined, and that its meaning shifts with the political and strategic interests of the claimant [16,17]. One useful way to organize the field distinguishes a defensive register—protecting national assets and citizens against external dependence and intrusion—from a more assertive register concerned with the projection of regulatory and normative authority over digital activity.

A recurrent intuition in policy-oriented work is that sovereignty develops through stages: a state begins as a user of externally supplied technology, may progress to ownership of domestic infrastructure and the imposition of local rules, and only at maturity becomes a producer that controls core technologies across the supply chain. This staged logic is valuable, and we build on it below, but in its existing form it treats the chain as running from hardware through software without isolating the model layer as a distinct locus of control. As a result, the dominant operational

expression of sovereignty has been data localization and infrastructural ownership—the layers that are most visible and most tractable to regulation, but not necessarily the layers where decisive control now resides.

2.2 The algorithmic turn and foundation-model dependency

Foundation models change the calculus because they introduce a chokepoint that sits above both data and infrastructure. A state may hold its data domestically and run workloads on sovereign hardware while still depending, for the intelligence of its defensive systems, on a model it did not train and cannot inspect. Three properties make this dependency strategically distinctive. The first is opacity: closed-weight models cannot be independently audited for bias, backdoors, or failure modes, so the defender must trust capabilities it cannot verify. The second is non-substitutability: where a defensive workflow is tightly coupled to a specific proprietary model and its interface, switching costs become a source of leverage for the supplier. The third is revocability: access mediated by an application programming interface can be conditioned or withdrawn unilaterally, converting a commercial relationship into a potential instrument of coercion.

These properties explain why an emergent body of work treats control over models—their weights, their auditability, their replaceability—as the true frontier of sovereignty, and why open-weight models have become a focal point for states seeking to reduce strategic exposure. They also explain why genuine autonomy is increasingly framed in terms of the full AI lifecycle—data, compute, models, deployment, and governance—rather than any single layer. The construct we develop takes this lifecycle seriously and treats the model layer as the analytical hinge that prior sovereignty frameworks omit.

2.3 Theoretical lenses

To avoid the criticism that the framework is merely descriptive, we ground each of its components in established theory. We draw on five complementary lenses.

Resource-based view (RBV). The RBV holds that durable advantage derives from resources and capabilities that are valuable, rare, inimitable, and non-substitutable [1,2]. Translated to the level of the state, sovereign control over defensive AI depends on assembling a stock of such resources—compute, data, talent, and, decisively, models. The framework draws directly on the non-substitutability criterion: a foreign closed-weight model on which defense depends is precisely a resource the state cannot freely substitute, which is why dependence on it is a strategic, not merely operational, vulnerability.

Dynamic capabilities and absorptive capacity. Possessing resources is not the same as being able to renew them. Dynamic-capabilities theory describes the capacity to sense opportunities and threats, seize them, and reconfigure the resource base as conditions change [3,4,5]. The related notion of absorptive capacity captures an organization's ability to recognize, assimilate, and apply external knowledge [6]. Together these lenses underpin the dimension that distinguishes states able to adapt and produce defensive AI from those that can only consume it, and they motivate the hypothesis that absorptive capacity moderates how dependency translates into vulnerability.

Securitization theory. Whether algorithmic dependency is treated as a routine procurement matter or as an existential national-security concern is not given by the technology; it is constructed through discourse. The Copenhagen School's securitization theory explains how an issue is elevated to the security agenda through speech acts that designate an existential threat and justify extraordinary measures [7]. This lens grounds the governance/normative dimension and clarifies why algorithmic sovereignty becomes politically actionable only once dependency is publicly framed as a threat to the state.

Dependency theory and digital dependency. The structural position of emerging economies as consumers of foreign algorithms echoes older core–periphery dynamics. Dependency and world-systems scholarship described how peripheral economies remained reliant on the technological core [11,12]; contemporary analyses extend the logic to data and platforms, characterizing the extraction of value from peripheral data as a form of colonization [10]. This lens situates the framework's boundary condition—its focus on middle-power and emerging economies—within a recognized theoretical tradition rather than treating it as an incidental scope restriction.

Cyber power. Finally, sovereignty is ultimately a claim about power. Scholarship on cyber power analyzes how capabilities in and through the digital domain translate into the ability to achieve preferred outcomes and to resist coercion [8,9]. This lens frames the dependent variable of the framework—cyber-defense resilience and resistance to coercion—and links the internal logic of sovereignty to the external behavior it is meant to enable.

A sociotechnical-systems perspective, which insists that technical and social subsystems must be jointly optimized [13], complements these lenses where the framework concerns the interface between automated systems and human operators. We synthesize these traditions not eclectically but functionally: each lens is mapped to a specific dimension or relationship in the framework, as Table 2 makes explicit.

3. Research approach

This is a conceptual paper, and we treat conceptual work as a disciplined research activity with its own standards rather than as unconstrained argument. Following established guidance on theoretical contribution, a contribution is judged by whether it adds value beyond existing accounts—typically by introducing new constructs, reorganizing relationships among known constructs, or specifying conditions under which relationships hold [18,20]. Among the recognized templates for conceptual articles, this paper is best characterized as theory synthesis: it integrates previously disconnected lenses and observations into a single framework that affords a novel way of seeing the phenomenon [19].

The design logic proceeds in four moves. First, we derive the framework's dimensions deductively from the AI cyber-defense lifecycle and abductively from the theoretical lenses, so that each dimension corresponds both to a stage of the technical pipeline and to a recognized source of strategic control. Second, we organize the dimensions against a maturity continuum adapted from staged models of sovereignty, producing a two-axis structure rather than a flat list. Third, we render the framework falsifiable by stating propositions that relate its components to defensive outcomes and specify moderating and boundary conditions. Fourth—and this is what distinguishes the paper from speculative commentary—we specify an empirical validation protocol and an operational indicator bank, so that the framework can be tested, weighted, and applied by other researchers. The paper therefore delivers a theory and the apparatus required to subject that theory to evidence; it does not claim that the theory has yet been empirically confirmed.

4. The algorithmic sovereignty framework

4.1 Construct definition and conceptual boundaries

Algorithmic sovereignty is defined as the capacity of a state to exercise autonomous and meaningful control over the AI and algorithmic systems embedded across its cyber-defense lifecycle, such that defensive capability is neither contingent upon, nor coercible through, foreign-controlled algorithmic dependencies. Three terms in this definition carry analytical weight. Capacity signals that sovereignty is a latent, graded property rather than a binary status. Meaningful control distinguishes genuine command—the ability to inspect, modify, replace, and govern—from nominal control such as a localization rule that leaves decisive capability in foreign hands. Coercibility makes explicit that the ultimate concern is strategic: the test of sovereignty is whether an external actor can degrade or weaponize the defender's dependence.

Two boundaries delimit the construct. Substantively, it concerns the algorithmic layer specifically; it presupposes, but does not subsume, data and infrastructure sovereignty. Contextually, it is theorized for middle-power and emerging economies—states that are predominantly algorithm consumers and for which full autarky is neither feasible nor desirable. For frontier producers the construct would require re-specification, and we make no claim to cover them.

4.2 The six dimensions

The construct is decomposed into six interdependent dimensions arranged as a stack that traces the AI cyber-defense lifecycle from physical substrate to governing norms (Fig. 1). The model/algorithmic dimension (D3) is the analytical hinge that distinguishes this framework from prior sovereignty accounts.

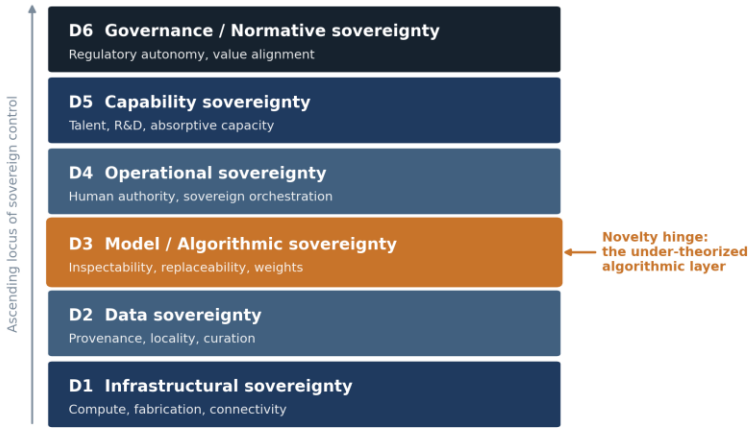

Fig. 1. The algorithmic sovereignty stack: six interdependent dimensions of the AI-enabled cyber-defense lifecycle, with the model/algorithmic layer (D3) as the under-theorized hinge.

D1 – Infrastructural sovereignty. Control over the compute, hardware fabrication, and connectivity that physically host defensive AI. The characteristic failure mode is reliance on foreign-headquartered hyperscalers for sensitive workloads, leaving the defender exposed to jurisdictional reach and single-provider concentration.

D2 – Data sovereignty. Control over the provenance, locality, and curation of the data used to train and operate defensive models. Its failure mode is the inability to assemble domain- and language-specific datasets, or loss of legal control over telemetry that crosses borders.

D3 – Model / algorithmic sovereignty. Control over the models themselves: their inspectability, auditability, portability, and replaceability. Its failure modes are opacity (the inability to audit a closed-weight system), lock-in (operational coupling to a single proprietary model), and revocability (exposure to unilateral withdrawal of access). This dimension is the framework's distinctive contribution.

D4 – Operational sovereignty. Control over deployment and decision: whether human operators retain authority over algorithmic actions, and whether routing, logging, escalation, and authorization remain state-owned functions rather than supplier-determined. Its failure mode is supplier-induced boundary control, where the vendor effectively shapes the limits of permissible action.

D5 – Capability sovereignty. The domestic human capital, research, and industrial base required to adapt and ultimately produce defensive AI. Grounded in dynamic capabilities and absorptive capacity, its failure mode is the absence of a workforce and ecosystem able to assimilate external advances, which freezes a state at the consumer stage.

D6 – Governance / normative sovereignty. Regulatory autonomy over defensive AI, institutional mandate for oversight, alignment of deployed model behavior with national legal and ethical norms, and the leverage exercised through standards and procurement. Its failure mode is normative dependence, in which the values and rules embedded in foreign systems silently override domestic priorities.

Table 2. The six dimensions, their locus of control, theoretical anchors, and characteristic failure modes.

Dimension	Locus of control	Primary theoretical anchor	Characteristic failure mode
D1 Infrastructural	Compute, fabrication, connectivity	Resource-based view	Hyperscaler / single-provider concentration
D2 Data	Provenance, locality, curation	Resource-based view	Loss of legal control over cross-border data

Dimension	Locus of control	Primary theoretical anchor	Characteristic failure mode
D3 Model / Algorithmic	Inspectability, replaceability, weights	RBV (non-substitutability)	Opacity, lock-in, revocability
D4 Operational	Human authority, sovereign orchestration	Sociotechnical systems theory	Supplier-induced boundary control
D5 Capability	Talent, R&D, absorptive capacity	Dynamic capabilities	Frozen at the consumer stage
D6 Governance / Normative	Regulatory autonomy, value alignment	Securitization; cyber power	Normative dependence

4.3 The maturity continuum

Each dimension can be held at different depths of control. Adapting and extending staged models of sovereignty, we specify a five-stage continuum running from dependence to generative capacity (Table 3). The continuum is not a maturity ladder that every state should climb to its summit; as the propositions below make clear, the appropriate position is contingent on a state's resources and on the autonomy–capability trade-off it faces.

Table 3. The five-stage algorithmic sovereignty maturity continuum.

Stage	Label	Descriptor
0	**Dependent**	Capability exists only through foreign-controlled provision; the system cannot be inspected or substituted, and access can be withdrawn.
1	**Regulated**	Foreign provision persists but is bounded by domestic rules—localization mandates, contractual terms, audit clauses. Control is legal rather than technical.
2	**Co-developed**	Capability is jointly developed or adapted with partners; partial inspection and conditional substitution become possible.
3	**Autonomous**	Domestic ownership and operation of inspectable, replaceable systems, with only residual dependency on external inputs.
4	**Generative**	Domestic capacity to produce, evolve, and export the capability; the state sets terms rather than accepting them.

4.4 The diagnostic matrix and configurational fit

Crossing the six dimensions with the five stages yields a diagnostic matrix that characterizes a state's algorithmic-sovereignty posture as a profile rather than a single score (Fig. 2). The matrix is intended as an analytic and policy instrument: it locates where control is strong and where it is brittle, and it makes visible the common and dangerous pattern in which a state achieves advanced data and infrastructure control while remaining dependent at the model layer. We argue that the configuration of the profile—the coherence across dimensions—matters more than the height of any single dimension, a claim formalized in Proposition 5.

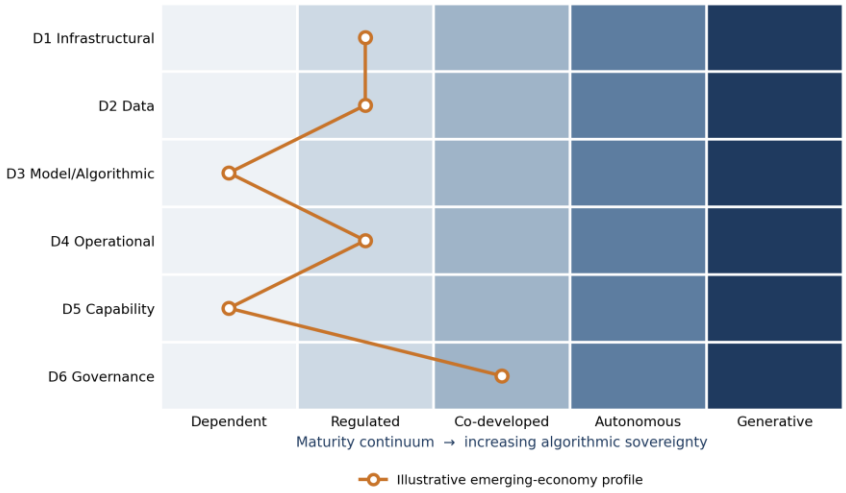

Fig. 2. The algorithmic sovereignty diagnostic matrix (six dimensions × five maturity stages). The overlaid line illustrates a stylized emerging-economy profile in which regulatory and data control outpace model-layer sovereignty.

4.5 Propositions

The framework yields five propositions that connect its components to defensive outcomes and render it testable (Fig. 3).

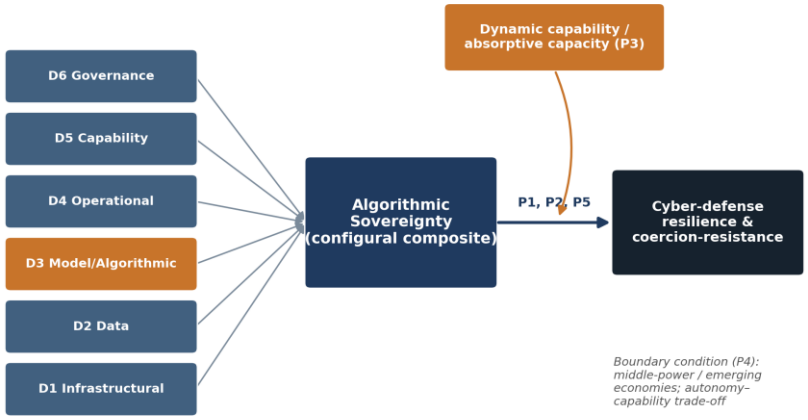

Fig. 3. Nomological model linking the six dimensions, as a configural composite, to cyber-defense resilience and coercion-resistance, with absorptive capacity as a moderator and an emerging-economy boundary condition.

P1. The greater a state's dependence on foreign closed-weight models in its cyber-defense stack (low D3), the more vulnerable its defensive posture is to supplier-induced coercion such as access withdrawal, ceteris paribus.

P2. Sovereignty achieved at the infrastructure and data layers (high D1–D2) without corresponding model-layer sovereignty (low D3) produces a sovereignty illusion: measured compliance increases while genuine strategic exposure is unchanged.

P3. Dynamic capability and absorptive capacity (D5) moderate the dependence–vulnerability relationship, such that states with higher absorptive capacity convert external AI advances into autonomy more rapidly and suffer less from a given level of dependence.

P4. For emerging economies, maximalist strategies that pursue full autonomy immediately reduce short-run cyber resilience by severing access to frontier capability; an intermediate, configurationally coherent posture dominates both autarky and unmanaged dependence (the autonomy–capability trade-off).

P5. Configurational fit across the six dimensions is more predictive of cyber-defense resilience than the maximum value attained on any single dimension; balanced profiles outperform spiky ones at equal aggregate investment.

5. Operationalization: a candidate indicator bank

To move from construct to measurement, each dimension is operationalized through formative indicators—observable facets that jointly constitute the dimension rather than reflecting a single latent cause. Table 4 presents a candidate indicator bank intended as the input to the Delphi validation described in Section 6, not as a finished instrument. Each indicator is scored against the five-stage continuum using behaviorally anchored descriptors (0–4), and dimension scores are aggregated into a profile rather than collapsed prematurely into a single index, preserving the configurational information that Proposition 5 deems decisive.

Table 4. Candidate formative indicators per dimension (Delphi item bank, abridged).

Dimension	Candidate indicator	Illustrative measure / data source
D1 Infrastructural	Share of defensive compute on domestically controlled infrastructure	% of sensitive workloads; national cloud registries
	Dependence on foreign-headquartered hyperscalers for defense workloads	Concentration index; procurement records
	Availability of sovereign / air-gapped compute for sensitive AI	Capacity inventory; expert rating
D2 Data	Provenance and locality of training and operational data	Data-flow audit; classification records
	Legal control over telemetry crossing borders	Statutory review; contract analysis
	National capacity to curate domain- and language-specific corpora	Dataset inventory; expert rating
D3 Model / Algorithmic	Reliance on closed-weight vs. open / inspectable models	% of pipeline by model type
	Ability to inspect, audit, and red-team model behavior	Audit-rights assessment; expert rating
	Model replaceability / portability without operational disruption	Switching-cost analysis; architecture review
	Exposure to unilateral API / access withdrawal	Supplier-risk index; contract terms
	Capacity to fine-tune, retrain, or host models domestically	Capability inventory
D4 Operational	Degree of human authority over algorithmic defensive actions	Process audit; rules of engagement
	State ownership of routing, logging, escalation, authorization	Architecture review
	Explainability and auditability available to national operators	Tooling assessment; expert rating

Dimension	Candidate indicator	Illustrative measure / data source
D5 Capability	Density of domestic AI-security talent	Workforce statistics; per-capita measures
	Defensive-AI R&D investment relative to GDP	Budget data; national statistics
	Absorptive capacity to assimilate and reconfigure external advances	Composite proxy; expert rating
D6 Governance / Normative	Regulatory autonomy over AI used in defense	Legal-framework review
	Institutional mandate for algorithmic-sovereignty oversight	Mandate mapping
	Procurement leverage embedding sovereignty requirements	Procurement-policy analysis

6. Proposed validation methodology

The framework is offered as a validatable theory, and we specify a four-stage protocol that other researchers can execute. The protocol combines qualitative content validation with quantitative weighting and an illustrative application, balancing rigor with feasibility in data-scarce emerging-economy settings.

Stage 1 – Systematic literature review. A review conducted in accordance with the PRISMA 2020 reporting standard [21] refines the dimensions and indicators against the published evidence base, ensuring that the item bank reflects the literature rather than the authors' priors and documenting the search and screening transparently.

Stage 2 – Delphi content validation. A structured Delphi process with a multidisciplinary expert panel—spanning cyber defense, AI engineering, and technology policy—assesses the relevance and clarity of each indicator over two to three rounds [22,23]. Consensus is quantified with Kendall's coefficient of concordance, and each indicator is retained only if its content validity ratio meets the critical value for the panel size [24]. This stage establishes content validity and prunes the item bank.

Stage 3 – Multi-criteria weighting. Because the dimensions are unlikely to contribute equally, a multi-criteria decision method assigns weights. Either fuzzy analytic hierarchy process [25], which accommodates the imprecision of expert judgment, or the best-worst method [26], which reduces the number of pairwise comparisons and improves consistency, is appropriate; the latter is attractive where expert time is constrained. The output is a transparently weighted composite for each dimension.

Stage 4 – Illustrative national instantiation. The validated, weighted framework is applied to a single national case to establish face validity and demonstrate diagnostic utility. The instantiation is illustrative rather than confirmatory: it shows that the matrix can be populated with real institutional evidence and that it surfaces actionable gaps. A configurational analysis using fuzzy-set qualitative comparative analysis [27,28] is proposed as a subsequent step to test Proposition 5 across a sample of states, examining whether balanced profiles are associated with superior defensive outcomes.

Reporting each stage against its established standard—PRISMA for the review, recognized Delphi and content-validity criteria for the panel, and documented multi-criteria procedures for weighting—guards against the principal risk facing conceptual frameworks: that they are asserted rather than earned.

7. Discussion

7.1 Theoretical implications

The framework makes three theoretical moves. It extends sovereignty thinking upward from data and infrastructure to the algorithmic layer, naming a locus of control that prior accounts implicitly assumed away. It integrates five traditions that have rarely been brought together on this problem—RBV, dynamic capabilities, securitization, dependency theory, and cyber power—mapping each to a specific dimension or relationship rather than invoking them rhetorically. And it advances a configurational logic: by treating sovereignty as a profile whose

coherence matters more than its peaks, it reframes the policy question from how high to climb toward how to balance, opening the construct to configurational methods.

7.2 The autonomy–capability trade-off

A central implication, captured in Proposition 4, is that for emerging economies sovereignty is not maximized by autarky. Severing ties to frontier capability in pursuit of immediate independence can lower defensive performance in the short run, because indigenous substitutes may lag the systems they replace. The framework therefore counsels managed interdependence: using regulation, procurement leverage, and open-weight strategies to raise model-layer control where it most reduces coercibility, while continuing to draw on external capability where doing so is safe and substitutable. Sovereignty, on this view, is a portfolio problem, not a switch.

7.3 Policy implications: an Indonesian illustration

Indonesia exemplifies the configuration the framework is designed to diagnose. The country has advanced markedly on the governance and data dimensions—establishing a national cyber and cryptography authority and enacting comprehensive personal-data-protection legislation—yet, like most middle powers, it remains dependent at the model layer, where defensive AI capability is sourced predominantly from foreign closed-weight providers. The framework predicts that such a profile, strong on D1–D2 and D6 but weak on D3, exhibits the sovereignty illusion of Proposition 2: regulatory maturity coexists with unaddressed strategic exposure. The corresponding policy agenda is specific. It includes prioritizing inspectable and replaceable models for sensitive defensive functions, embedding audit-rights and exit clauses in procurement to attack revocability and lock-in directly, investing in absorptive capacity through talent and applied research so the state can climb from the regulated to the co-developed stage, and asserting sovereign orchestration so that authorization and escalation remain national functions. None of these requires frontier-scale model production; each raises sovereignty where coercibility is highest.

7.4 Implications for the Global South

The Indonesian pattern generalizes. Many emerging economies have invested in the visible, regulable layers of sovereignty while the decisive algorithmic layer remained outside the policy frame. By giving that layer a name, a structure, and a measurement strategy, the framework equips these states to diagnose their true exposure and to sequence investment where it most reduces coercibility—and it offers a shared analytic vocabulary for the regional cooperation, such as pooled compute or joint open-model initiatives, through which middle powers can raise sovereignty collectively rather than alone.

8. Limitations and future research

Three limitations bound the contribution. First, the framework is conceptual: its dimensions, propositions, and indicators are theoretically derived and have not yet been empirically validated, which is precisely why Section 6 specifies a protocol for doing so. Second, four of the five theoretical lenses originate in management and security scholarship developed in the global North; applying them to emerging-economy contexts requires reflexivity, lest categories travel poorly, and the dependency lens is included partly to keep that risk in view. Third, the construct addresses a fast-moving technological frontier, so its indicators—especially those concerning model openness and access regimes—will require periodic revision as the AI supply chain evolves. Beyond executing the validation protocol, promising extensions include longitudinal study of how states move between maturity stages, comparative analysis across regions, and integration of the framework with operational measures of cyber resilience to test the outcome relationships directly.

9. Conclusion

As artificial intelligence becomes the intelligence of cyber defense, the question of who controls the algorithm becomes inseparable from the question of who is sovereign. The prevailing language of data and infrastructure sovereignty, formed before foundation models reshaped the field, no longer captures where decisive control resides. This paper has introduced algorithmic sovereignty to fill that gap, defined it as meaningful and non-coercible control over the AI embedded in the cyber-defense lifecycle, and developed it into a six-dimension, five-stage framework grounded in an explicit synthesis of established theory. By advancing testable propositions, operationalizing each dimension, and specifying a validation protocol, the paper offers not a closed argument but a research platform—one calibrated for the middle-power and emerging-economy contexts where the stakes of algorithmic dependence are highest and the path to managed autonomy most consequential.

REFERENCES

1. Barney, J.B., 1991. Firm resources and sustained competitive advantage. Journal of Management 17 (1), 99–120.
2. Wernerfelt, B., 1984. A resource-based view of the firm. Strategic Management Journal 5 (2), 171–180.
3. Teece, D.J., Pisano, G., Shuen, A., 1997. Dynamic capabilities and strategic management. Strategic Management Journal 18 (7), 509–533.
4. Teece, D.J., 2007. Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance. Strategic Management Journal 28 (13), 1319–1350.
5. Eisenhardt, K.M., Martin, J.A., 2000. Dynamic capabilities: what are they? Strategic Management Journal 21 (10–11), 1105–1121.
6. Cohen, W.M., Levinthal, D.A., 1990. Absorptive capacity: a new perspective on learning and innovation. Administrative Science Quarterly 35 (1), 128–152.
7. Buzan, B., Wæver, O., de Wilde, J., 1998. Security: A New Framework for Analysis. Lynne Rienner Publishers, Boulder.
8. Nye, J.S., 2010. Cyber Power. Belfer Center for Science and International Affairs, Harvard Kennedy School, Cambridge, MA.
9. Nye, J.S., 2011. The Future of Power. PublicAffairs, New York.
10. Couldry, N., Mejias, U.A., 2019. The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism. Stanford University Press, Stanford.
11. Cardoso, F.H., Faletto, E., 1979. Dependency and Development in Latin America. University of California Press, Berkeley.
12. Wallerstein, I., 1974. The Modern World-System I. Academic Press, New York.
13. 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.
14. Floridi, L., 2020. The fight for digital sovereignty: what it is, and why it matters, especially for the EU. Philosophy & Technology 33, 369–378.
15. Pohle, J., Thiel, T., 2020. Digital sovereignty. Internet Policy Review 9 (4).
16. Couture, S., Toupin, S., 2019. What does the notion of sovereignty mean when referring to the digital? New Media & Society 21 (10), 2305–2322.
17. Mueller, M.L., 2020. Against sovereignty in cyberspace. International Studies Review 22 (4), 779–801.
18. Whetten, D.A., 1989. What constitutes a theoretical contribution? Academy of Management Review 14 (4), 490–495.
19. Jaakkola, E., 2020. Designing conceptual articles: four approaches. AMS Review 10, 18–26.
20. MacInnis, D.J., 2011. A framework for conceptual contributions in marketing. Journal of Marketing 75 (4), 136–154.
21. Page, M.J., et al., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 372, n71.
22. Okoli, C., Pawlowski, S.D., 2004. The Delphi method as a research tool: an example, design considerations and applications. Information & Management 42 (1), 15–29.
23. Dalkey, N., Helmer, O., 1963. An experimental application of the Delphi method to the use of experts. Management Science 9 (3), 458–467.
24. Lawshe, C.H., 1975. A quantitative approach to content validity. Personnel Psychology 28 (4), 563–575.
25. Chang, D.-Y., 1996. Applications of the extent analysis method on fuzzy AHP. European Journal of Operational Research 95 (3), 649–655.
26. Rezaei, J., 2015. Best-worst multi-criteria decision-making method. Omega 53, 49–57.
27. Ragin, C.C., 2008. Redesigning Social Inquiry: Fuzzy Sets and Beyond. University of Chicago Press, Chicago.
28. Fiss, P.C., 2011. Building better causal theories: a configurational approach to organizational research. Academy of Management Journal 54 (2), 393–420.
29. Timmers, P., 2019. Strategic autonomy and cybersecurity. EU Cyber Direct, Brussels.
30. Roberts, H., Cowls, J., Casolari, F., Morley, J., Taddeo, M., Floridi, L., 2021. Safeguarding European values with digital sovereignty: an analysis of statements and policies. Internet Policy Review 10 (3).