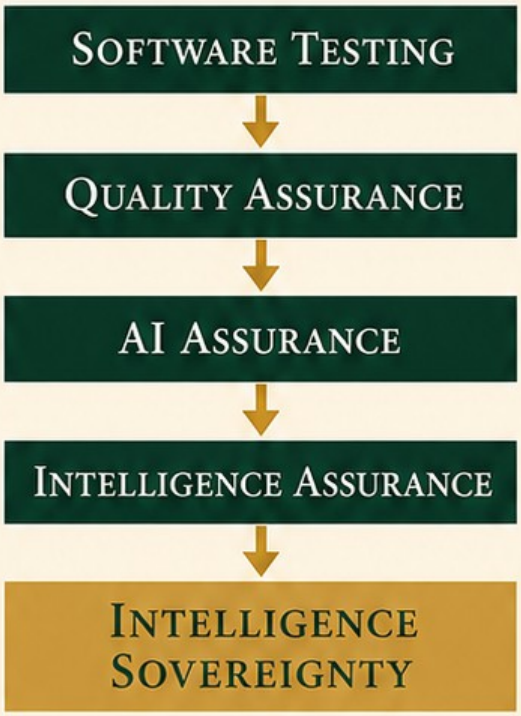
INTELLIGENCE SOVEREIGNTY FRAMEWORK

ARTICLE SERIES

FROM DIGITAL TRUST TO INTELLIGENCE SOVEREIGNTY

Reflections One Year After My ANST Keynote



ONOSOJI O. ONONUGA

CEREARK QUALITY, RISK & ASSURANCE CONSORTIUM

FREEDOM TO GOVERN THE INTELLIGENCE THAT SHAPES OUR FUTURE

INTELLIGENCE SOVEREIGNTY PAPERS

Paper No. 1

FROM DIGITAL TRUST TO INTELLIGENCE SOVEREIGNTY

Reflections One Year After My ANST Keynote

PUBLIC CONSULTATION EDITION v1.1

Publication information

Publication	Intelligence Sovereignty Papers - Paper No. 1
Title	From Digital Trust to Intelligence Sovereignty
Subtitle	Reflections One Year After My ANST Keynote
Author	Onosoji O. Ononuga
Publisher	Cereark Quality, Risk & Assurance Consortium
Identifier	ISF-PAP-001
Edition	Public Consultation Edition v1.1
Status	Candidate for Principal approval; revised public consultation edition
Date	4 August 2026
Classification	PUBLIC - subject to final approval

Status and revision notice

This candidate edition is prepared to supersede Public Consultation Edition v1.0 for current circulation, subject to the Principal's approval and completion of publication quality assurance. Version 1.0 will remain preserved in the publication archive as the first approved public consultation edition.

Version 1.1 strengthens the paper's engagement with technological, digital and AI-sovereignty scholarship; clarifies the Framework's originality boundary; and refines its treatment of interdependence, evaluative capability, resilience, legitimate control and institutional authority. The controlled four-capability definition and the central proposition are retained.

This publication remains a proposal for consultation. It is not government policy, legal advice, regulatory guidance, an accredited standard, a validated assessment method or a certification scheme.

Suggested citation

Ononuga, O. O. (2026). From Digital Trust to Intelligence Sovereignty: Reflections One Year After My ANST Keynote (Public Consultation Edition v1.1). Intelligence Sovereignty Papers, Paper No. 1. Cereark Quality, Risk & Assurance Consortium.

Copyright and disclaimer

Abstract

Artificial intelligence is changing not only how institutions perform work, but the conditions under which they exercise judgement and authority. Existing scholarship has already developed technological sovereignty as agency under interdependence, AI sovereignty as control over strategic inputs and infrastructures, competence-based sovereignty, organisational and control sovereignty, resilience, contestability and managed dependence.

This paper proposes a narrower operational contribution. The Intelligence Sovereignty Framework asks whether a consequential intelligence capability remains independently understandable, evaluable, governable and continuable at the institutional point of reliance. Its unit of analysis is not the national AI stack alone, but the socio-technical arrangement through which intelligence materially shapes a decision, service or strategic function.

The paper distinguishes sovereignty inputs from governing capability; control mechanisms from the competence to exercise them; and service continuity from continuity of institutional judgement. It presents a five-question, non-compensatory screening instrument for consultation and introduces the Institutional Evaluation Paradox as a separate research programme concerning the possible erosion of evaluative capability under sustained cognitive delegation.

The Framework does not claim technological autarky, domestic ownership of every component or immunity from legitimate external constraint. It proposes governable and accountable interdependence: external reliance may be compatible with Intelligence Sovereignty where evidence, challenge, continuity, accountability and practical freedom of action remain proportionate to consequence.

Keywords

Intelligence Sovereignty; Evaluation Paradox; institutional cognitive resilience; intelligent customer; intelligent client; Intelligence Assurance; AI Assurance; sovereign AI; digital sovereignty; operational resilience; third-party concentration risk; evaluative capability; judgement continuity; cognitive delegation; retained competence; proportionality; Nigeria; NITDA.

> *The more an institution delegates intelligence, the less capable it may become of evaluating the intelligence it receives.*

Contents

CHAPTER 01

The Conversation We Need to Have

Artificial intelligence is changing the object of governance: from software and data to the judgement produced through them.

A narrower and stronger claim

The field is crowded, and the paper's claim must therefore be exact. Intelligence Sovereignty does not originate the application of sovereignty to digital technology or artificial intelligence. Nor does it originate sovereignty as agency, capability, competence, resilience, contestability, lifecycle control or managed interdependence.

Prior work has already shown that technological sovereignty concerns governmental agency within an interdependent international system; that AI sovereignty depends on data, compute, models, infrastructure, skills and regulation; that productive competence can be measured through the integration of AI knowledge; and that institutions need contractual, technical, operational and contestable control over AI arrangements.

The proposed contribution is narrower: to integrate evaluative capability, legitimate governing authority and service-level continuation into a non-compensatory assessment of whether a consequential intelligence dependency remains governable in practice. The novelty claim is therefore combinatorial and operational, not ownership of each component idea.

The assurance progression

The professional journey remains cumulative: Software Testing establishes evidence about system behaviour; Quality Assurance examines the conditions that produce quality; AI Assurance addresses the distinctive risks of models, data and intelligent behaviour; Intelligence Assurance asks whether decision-shaping outputs deserve institutional reliance; and Intelligence Sovereignty asks whether that reliance can remain understandable, independently evaluable, governable and continuous under changing conditions.

Digital trust is not an additional rung in that progression. It is one outcome of the progression working well. The paper states that point once and then moves on, because the harder question is what happens when trust is operationally justified but institutional competence is decaying.

The governing question

The paper is organised around one governing question: if an intelligence capability changed, failed, became unavailable or remained available but increasingly opaque to its users, would the institution retain the ability to govern the decisions that depend upon it?

This question places evidence before assurance and assurance before trust. It also adds a capability that conventional continuity planning can miss: the ability to evaluate, not merely to operate.

CHAPTER 02

From Software Quality to Intelligence Governance

Nigeria has moved from consultation toward a national software-quality regime. That makes the next governance question more immediate, not less.

A policy position that moved during drafting

In April 2026, the National Information Technology Development Agency published the National Software Development Guideline, the National Software Testing Guideline and the Software Testing Organisations Licensing Guideline for consultation. On 29 July 2026, NITDA formally launched the National Software Quality Assurance Framework, signed by the Director-General under the NITDA Act 2007, bringing those three instruments into a single national assurance architecture.

The Framework is adopted and signed, but phased rather than yet fully operational. It requires government software to undergo independent third-party testing and certification before going live, and makes compliance a condition of IT Project Clearance. Its three-tier risk model applies the most stringent assurance to Class A systems; this paper uses core banking switches, national identity infrastructure and power-grid control as illustrative high-consequence examples. Full effect is scheduled for the second quarter of 2027, with capacity-building and an Expression of Interest for prospective testing organisations in the interim.

That distinction matters. Describing the instruments as merely draft would now understate the legal and policy change; describing the regime as already fully in force would overstate its present implementation. The precise position is that the Framework has been formally adopted and signed, while commencement, accreditation and enforcement are proceeding through a defined implementation pathway.

What the framework signifies

The important policy shift is that software quality is being treated as a matter of national governance. Systems used by government are no longer left entirely to supplier assurance or project-level judgement. Independent testing, organisational competence and risk-based scrutiny are being drawn into a more formal national architecture.

This direction is consistent with a broader principle: evidence should be proportionate to consequence. Systems supporting national identity, core finance, public safety or essential services require deeper scrutiny than low-impact administrative tools.

The live next frontier

The existence of a software-quality regime does not establish that government acted because of any single speech or professional intervention. Policy emerges through institutional work, consultation and converging advocacy. The relevant observation is more modest: themes raised in the 2025 ANST keynote - national capability, assurance, certification and stronger collaboration - now sit within a concrete policy landscape.

The next frontier follows naturally. As government software begins to classify, predict, recommend and prioritise, independent software testing remains necessary but no longer exhausts the assurance question. The issue becomes whether public institutions can evaluate the intelligence generated, retain the competence to challenge it, and continue governing affected decisions when suppliers, models or access conditions change.

CHAPTER 03

A Crowded Sovereignty Field

Intelligence Sovereignty cannot be credible unless it states where adjacent disciplines already reach - and where its proposed contribution begins.

> *Sovereign AI concerns control of the technology. Intelligence Sovereignty concerns continued authority over the judgement produced through it.*

Digital sovereignty and sovereign AI

Digital sovereignty generally concerns the ability of states, institutions, communities or individuals to exercise meaningful autonomy and control in digital environments. AI sovereignty narrows the focus toward the strategic assets, infrastructures, competencies and rules required to develop, deploy or control AI.

Technology-sovereignty scholarship already frames sovereignty as state agency rather than territorial ownership. Competence-based approaches measure the local mobilisation and integration of productive AI knowledge. KASE and related sovereignty-stack approaches identify data, algorithms, compute, connectivity, electricity, education, cybersecurity and regulation as national enablers. More recent work describes controlled dependence through contractual, technical, model-possession and operational controls.

The ISF accepts that lineage. It does not treat external collaboration as sovereignty failure and does not require a complete domestic AI stack. It asks a later and more granular question: when a ministry, regulator, hospital, firm or critical service relies on consequential intelligence, does the responsible institution retain the capability to understand, independently evaluate, govern and continue that function?

This shifts the unit of analysis from the national stack or innovation system to the consequential intelligence capability: the combined model, data, infrastructure, suppliers, professionals, procedures, evidence and decision rights through which judgement is produced and acted upon.

Responsible AI and AI assurance

Responsible AI and AI assurance focus on whether AI systems are lawful, safe, reliable, transparent, accountable and compatible with rights and institutional values. They provide essential lifecycle controls, impact assessment, monitoring and evidence. Intelligence Sovereignty depends upon those disciplines; it does not replace them.

Their typical unit of analysis, however, is the AI system or its lifecycle. The proposed unit of analysis here is the complete intelligence-producing arrangement through which consequential judgement is created and used: systems, data, suppliers, professionals, decision rights, internal expertise, verification practices and continuity options.

Operational resilience and third-party risk

Operational resilience has already developed mature approaches to critical third parties, concentration risk, vendor lock-in, stressed exits and the continuity of important services. DORA requires tested exit strategies for ICT services supporting critical or important functions. UK regulators similarly require firms to understand dependencies, preserve accountability and plan for severe disruption and stressed exit.

The Framework does not claim to have discovered these risks. Its proposed extension concerns judgement continuity. A firm may preserve the technical service, transfer to another provider and meet an impact tolerance, yet still lack the competence to assess whether the replacement intelligence is sound, biased, mis-specified or inappropriate for the decision.

The intelligent customer function

Public procurement and outsourcing governance already names a capability close to the centre of this paper. The intelligent customer, or intelligent client, function is the retained in-house competence an organisation keeps in order to specify, oversee and critically judge work it contracts out, so that it does not become dependent on a supplier it can no longer question. UK practice has developed the principle for decades: early CCTA guidance applied it to government IS/IT sourcing; the Health and Safety Executive applies it to safety-critical contracting; and the Government Property Function sets current expectations for intelligent-client roles (CCTA, 1994; Health and Safety Executive, n.d.; Government Property Function, 2022).

The parallel is close and must be conceded. The HSE formulation could almost serve as a definition of evaluative capability: an organisation using a contractor for safety analysis must confirm that the method is appropriate, ask the right questions and understand the limitations of the result, without itself possessing the depth of knowledge required to perform the analysis. That is the understand-and-evaluate distinction this paper draws, articulated long before it.

The proposed contribution is therefore not the intelligent customer function but its instability under AI. The classic function assumes a retained in-house competence supervising an external supplier. When the delegated work is cognitive and the supplier is an always-available model embedded in daily practice, relying on it can degrade the very competence the intelligent customer function depends upon. Intelligent-customer capability can no longer be treated as a stable background condition; it becomes an asset that erodes through use unless deliberately preserved. Intelligence Sovereignty treats that preservation as a first-order governance requirement rather than an incidental feature of contract management.

The claimed distinction

The claimed distinction is not sovereignty as capability in the abstract. That position is well established. Nor is it the proposition that contracts, testing, monitoring, portability or fallback matter; neighbouring frameworks already recognise those controls.

The proposed distinction is the treatment of retained evaluative competence and continuation capacity as critical, non-compensatory conditions of effective institutional authority. A country may possess domestic compute, local models and strong regulation while a particular institution remains unable to challenge a consequential output or continue the affected decision function. Conversely, an externally supplied capability may remain governable where evidence access, independent testing, alternatives, retained expertise and tested continuity are credible.

Sovereignty inputs may strengthen agency. Control mechanisms may create leverage. Intelligence Sovereignty asks whether those inputs and mechanisms have produced a live institutional capability at the point where judgement actually depends on them.

Comparison at a glance

Field	Primary question	Principal object
Digital sovereignty	State or institutional capacity and control in the digital environment	Digital infrastructure, data, jurisdiction and agency
Sovereign AI	Ability to build, access or control AI capability	Compute, models, data, infrastructure, talent and national choice
Responsible AI / AI assurance	Whether AI systems are lawful, safe, reliable and accountable	AI-system lifecycle, impacts, controls and evidence
Operational resilience	Whether important services can withstand, recover and exit third-party disruption	Continuity, concentration, substitutability and impact tolerances
Intelligent customer / client	Can the organisation specify, oversee and critically judge externally supplied work?	Retained in-house competence, supplier challenge and corporate memory
Intelligence Sovereignty	Whether consequential judgement remains understandable, evaluable, governable and continuous	Complete intelligence-producing arrangement and retained evaluative capability

CHAPTER 04

The Institutional Evaluation Paradox

The more an institution delegates intelligence, the less capable it may become of evaluating the intelligence it receives.

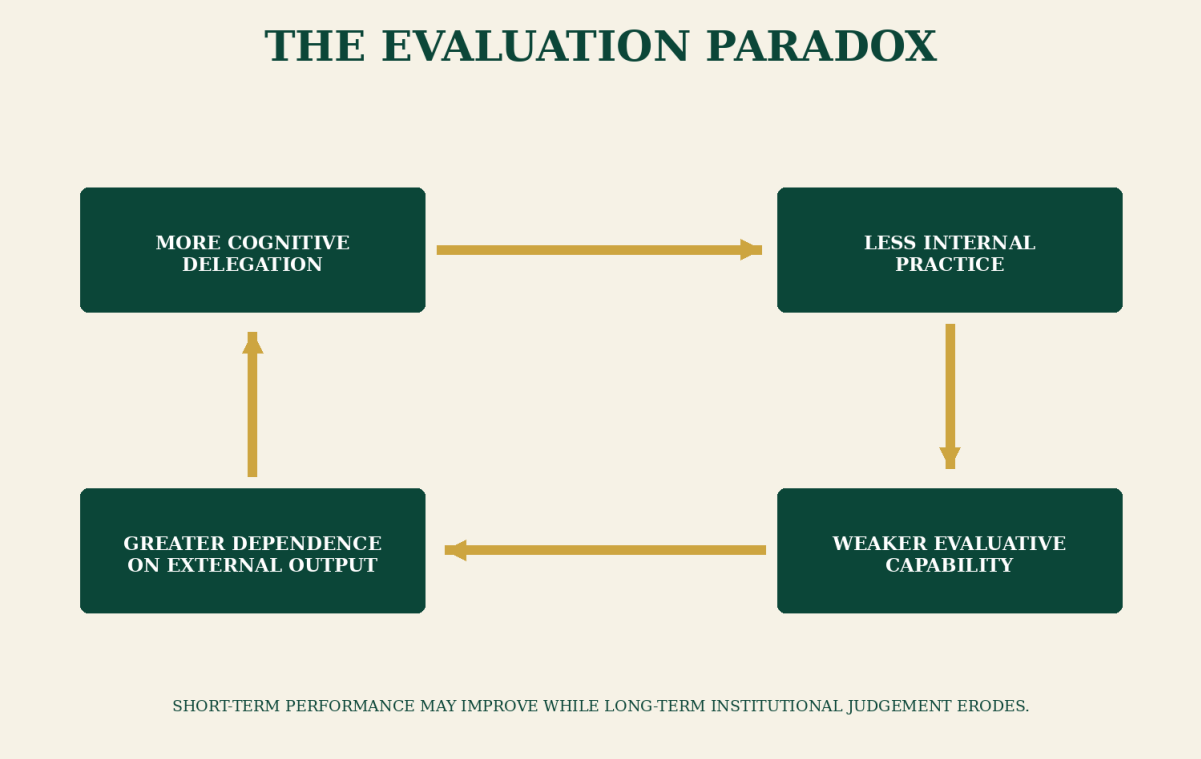


Figure 1. The Evaluation Paradox - a proposed institutional feedback loop.

Delegation changes the evaluator

The broad problem is not new. Human-factors research has long shown that automation can weaken situation awareness, monitoring and takeover performance. Automation-bias and cognitive-offloading research further shows that assisted performance can improve while independent understanding, skill development or calibration deteriorates.

The institutional question is narrower. When cognitive work is delegated repeatedly, the organisation may preserve formal accountability and apparently successful outputs while losing the practised capability required to test, reproduce, contest or replace the intelligence on which it relies.

The label 'Evaluation Paradox' has close prior use in 2026 work on cognitive atrophy and verification in agentic workflows. This edition therefore uses Institutional Evaluation Paradox to identify the distinct scale and research object proposed here.

Evidence from human-automation research

Bainbridge's ironies of automation, the out-of-the-loop performance literature and studies of automation bias establish that people can become poorer monitors and less capable of intervention as automation performs more of the underlying task. Verification complexity, cognitive load, trust and task conditions materially affect whether errors are detected.

Recent AI research extends this lineage. Studies of cognitive offloading and metacognitive calibration report a recurring performance-competence separation: users may complete tasks more efficiently while learning less, preserving less independent capability or becoming less well calibrated about what they can do without assistance. Formal delegation-and-verification models also show that rational incentives may produce over-delegation and reduced verification effort.

These literatures support mechanism plausibility. They do not yet establish the full institution-level claim advanced here: longitudinal depreciation of organisational evaluative capability, moderated by feedback observability, with consequences for continuity and effective authority.

What is established and what remains inferential

Established evidence supports automation bias, out-of-the-loop effects, cognitive offloading, reduced verification effort, organisational forgetting and short-term impairment of independent skill development in some settings. It also supports the proposition that expertise is more reliable where environments provide sufficiently regular feedback and genuine opportunities to learn.

The proposed institutional extension remains inferential. It suggests that sustained delegation may alter staffing, practice, career pathways, knowledge retention and challenge routines until evaluative capability declines faster than the organisation recognises. The effect is expected to be hardest to observe where feedback is delayed, rare, selective, counterfactual or normatively contested.

Paper No. 2 develops that proposition separately through intrinsic and achieved observability, the seeded-versus-real detection gap, measurement reactivity and a simulator-plus-historical-case validation programme. None of those methods is treated as validated in this paper.

The institutional form of the paradox

At institutional scale, the paradox is broader than individual automation bias. Expertise may migrate to suppliers; internal career pathways may disappear; source data may become less familiar; assurance teams may verify process rather than substance; and senior decision-makers may retain formal accountability while losing practical means of challenge.

The critical failure is therefore possible even without provider withdrawal. The service may be available, accurate on average and contractually compliant. Sovereignty still weakens if the institution cannot recognise when the capability is wrong, explain why it is wrong or reconstruct the judgement independently.

Evaluative capability as a protected institutional asset

The Framework treats retained evaluative capability as an asset in its own right. Where its loss would impair essential decisions, public authority or strategic continuity, that asset may also acquire sovereign significance. It includes domain expertise, access to underlying evidence, the ability to reproduce or independently test important conclusions, knowledgeable human challenge and institutional memory of how decisions were made before automation.

This does not require every task to be performed manually. It requires enough practice and expertise to preserve meaningful oversight. Human involvement that cannot detect error, contest assumptions or choose an alternative is ceremonial rather than substantive.

Research hypotheses, identification and falsification

The Institutional Evaluation Paradox remains a research programme rather than a settled causal law. Its distinctive hypotheses concern capability trajectory, not merely verification behaviour: lower intrinsic feedback observability may be associated with faster erosion under sustained delegation; achieved observability may mitigate that risk; and synthetic-probe performance may overstate capability against consequential real failures.

Credible testing must distinguish baseline confounders from mechanisms. Delegation-driven staffing loss, reduced practice and loss of tacit knowledge may be mediators rather than rival explanations and should not automatically be controlled away. Measurement itself can also preserve competence by keeping evaluators alert, so low-frequency measurement must be distinguished from high-frequency competence-preservation intervention.

The theory would be weakened if independent automated evaluators reliably detect consequential failures, if evaluative performance does not decline despite sustained disuse, if observability does not moderate the trajectory, or if observed decline is wholly explained by factors that precede delegation.

Defining Intelligence Sovereignty

A useful definition must discriminate. It should identify the minimum capability whose absence changes the governance condition.

A concise definition for consultation

This edition proposes a tighter definition:

> *Intelligence Sovereignty is the sustained ability of a nation or institution to understand, evaluate, govern and continue the intelligence capabilities on which consequential decisions, services and strategic interests depend, without unacceptable vulnerability to external actors.*

The definition is intentionally shorter than the earlier seven-verb formulation. Develop, access, assure and adapt remain important operational dimensions, but they are not all required in the core sentence. The discriminating work is done by four capabilities: understand, evaluate, govern and continue.

The definition concerns effective and legitimate human institutional agency within interdependence. It does not imply sovereignty possessed by artificial intelligence itself, and it does not make domestic ownership or self-sufficiency a prerequisite.

The referent: arrangement and output

Throughout this paper, an intelligence capability means the socio-technical arrangement that produces and mediates decision-shaping outputs: model, data, infrastructure, suppliers, professionals, procedures, evidence and decision rights. The intelligence an institution receives is the output of that arrangement - for example a ranking, classification, recommendation, prediction, synthesis or judgement. Sovereignty attaches primarily to the arrangement because that is what can be governed and continued; evaluation applies to its outputs because institutional authority depends on deciding whether a particular output merits reliance. Continue therefore means preserving the consequential decision or service and enough of the arrangement to support it, not preserving every model or supplier unchanged.

Understand

An institution understands a capability when it possesses enough technical, domain, operational and evidential knowledge to explain its role, identify material assumptions, recognise limitations and challenge significant outputs. Understanding does not mean complete access to every model parameter. It means sufficient comprehension for the consequence at stake.

An institution can therefore govern without fully developing a system, and it can access a system without understanding it. These are deliberately distinct states.

Evaluate

An institution evaluates a capability when qualified people can independently test, reproduce or contest its consequential outputs in a given case - not merely comprehend the capability in general. Understanding is knowing what the capability does and where it may fail; evaluation is being able to determine whether a particular output should be relied upon. Understanding can survive in documentation; evaluation depends on live, practised competence and is therefore the capability most exposed to the Evaluation Paradox.

Govern

An institution governs when legitimate decision rights, accountability, monitoring, intervention thresholds and change controls remain effective in practice. Formal ownership or contractual rights are not enough if the institution lacks the information or competence needed to exercise them.

Governance includes the ability to limit use, require evidence, suspend reliance, alter human oversight and accept or reject residual risk.

Continue

An institution can continue when it can preserve the critical decision or service through disruption, deterioration, provider withdrawal or transition. Continuity may involve an alternative provider, a simpler capability, a manual process, mutual assistance or a redesigned service. Sovereignty does not require keeping the identical system alive.

The relevant outcome is continued legitimate control over the decision, not technological permanence.

Unacceptable vulnerability

No institution is independent of external actors. The threshold is therefore not dependence but unacceptable vulnerability. Vulnerability becomes unacceptable when external decisions, failure or withdrawal could materially impair a critical function and the institution lacks proportionate means to understand, challenge, substitute or continue it.

The threshold must be determined by consequence, not rhetoric. A low-impact translation tool and a national-identity fraud model should not attract the same sovereignty investment.

Operational dimensions

Four supporting dimensions explain how the core capability is strengthened. Develop concerns selective indigenous or internal capability. Access concerns dependable participation in external ecosystems. Assure concerns evidence proportionate to consequence. Adapt concerns the ability to change models, processes and institutional arrangements as conditions evolve.

These dimensions support the definition; they do not inflate it.

Core capabilities and operational dimensions

Element	Evidence question
Understand	Can the institution explain material assumptions, limitations and evidence well enough to challenge consequential output?
Evaluate	Can qualified people independently test, reproduce or contest consequential outputs, rather than only comprehend the capability in general?
Govern	Can legitimate authority set conditions, intervene, suspend reliance and accept residual risk in practice?
Continue	Can the critical decision or service continue through failure, withdrawal, deterioration or transition?
Develop	Is selective internal or indigenous capability required by criticality and consequence?
Access	Is participation in external ecosystems dependable and sufficiently contestable?
Assure	Is reliance supported by evidence proportionate to consequence?
Adapt	Can the institution change technology, process and governance as conditions evolve?

CHAPTER 06

Confidence Is Not Control

Operational resilience protects service continuity. Intelligence Sovereignty adds the continuity of institutional judgement.

> *Exit capability preserves operations. Evaluative capability preserves judgement.*

What resilience already provides

Financial regulation offers a mature account of dependence. DORA requires comprehensive and tested exit strategies for ICT services supporting critical or important functions. FCA and Bank of England expectations require firms to map third-party dependencies, manage concentration and vendor lock-in, preserve accountability, and plan for stressed as well as orderly exit.

These regimes are directly relevant. Intelligence-dependent institutions should adopt the same discipline: identify critical arrangements, understand supply chains, define tolerances, rehearse disruption and maintain credible exit options.

The gap between service and judgement continuity

The proposed additional question is what must be preserved through the exit. In conventional outsourcing, the answer is the important business service. In an intelligence-dependent arrangement, continuity must also include the ability to judge whether the replacement output is appropriate.

A tax authority might transfer a fraud-detection workload to a new model without interrupting case processing. Yet if officials cannot compare the new model's assumptions, recognise changes in false-positive patterns or explain why risk classifications differ, the service has continued while institutional judgement has weakened.

Exit capability preserves operations. Evaluative capability preserves judgement.

Assurance of performance is not authority

A model may meet accuracy thresholds, pass fairness checks and remain highly available. None of those results alone proves that the institution can govern the decision process. Performance evidence concerns observed behaviour under defined conditions. Authority concerns the institution's practical ability to decide what evidence is sufficient, when confidence should be withdrawn and how the decision should proceed without the capability.

Confidence is therefore not control. The distinction is especially important when a supplier's model becomes the de facto standard against which all alternative judgement is assessed.

A bridge, not a competing universe

The Framework should operate as a bridge across existing disciplines. AI assurance supplies system evidence. Operational resilience supplies continuity, concentration and exit disciplines. Technology- and AI-sovereignty frameworks address strategic inputs, infrastructure and national agency. Public law supplies legitimacy, reason-giving and contestability.

Resilience answers whether the function survives disruption. Intelligence Sovereignty asks whether it remains independently evaluable and legitimately governable before, during and after that disruption. A service can continue while institutional judgement weakens; equally, strong formal control rights may exist without people capable of exercising them.

The Framework therefore distinguishes control mechanisms from control capability. Contracts, local inference, model possession, audit rights, fallbacks and portability are mechanisms. They become effective only where the institution has the evidence, competence, authority and operational knowledge required to use them.

The Institutional Shift

Formal authority is insufficient when operational judgement has migrated beyond the institution that remains accountable.

From automation to delegated judgement

Traditional software automated defined processes. Intelligent systems increasingly shape the interpretation of evidence, the ranking of cases and the selection of action. The institutional change is not that machines have acquired legal authority; they have not. It is that practical influence over judgement can migrate while formal responsibility remains in place.

This is familiar to principal-agent theory and to the growing literature on automated administrative decision-making. Authority may be delegated, information may be asymmetric and accountability gaps may emerge. AI intensifies the problem because the delegated object is not only execution but analytical judgement. Public-administration research identifies recurring trade-offs among efficiency, accountability, fairness, resilience, transparency and the rule of law, while administrative-law scholarship emphasises reviewability of the full socio-technical decision process rather than model explanation alone (Cobbe, Lee & Singh, 2020; Roehl & Hansen, 2024).

Accountability without capability

A public body can remain legally accountable for a decision while lacking practical capacity to interrogate the intelligence that shaped it. Human sign-off does not resolve the problem where the human cannot reproduce the reasoning, detect material error or make a credible alternative decision.

Reviewability requires more than a nominal human sign-off. It requires records, reasons, accessible evidence, contestability and officials capable of making a reasoned departure from machine-generated advice. Recent work on algorithmic administration similarly connects public-sector AI to discretion, the duty to give reasons and proportionality (Pavlidis & Kastanas, 2026). These literatures do not resolve Intelligence Sovereignty, but they establish that evaluative competence and accountable review are existing public-law concerns rather than newly discovered ones.

The institutional objective is therefore meaningful authority: the combination of legal responsibility, access to evidence, competence to challenge, and practical options to intervene.

Why this is not automatically constitutional

Earlier drafts described this as a constitutional shift. That language ran ahead of the analysis. The immediate problem is institutional and administrative: delegation, competence, accountability, continuity and evidence. Constitutional implications may arise where machine-generated intelligence materially shapes rights, public powers or access to remedies, but those implications require jurisdiction-specific legal analysis.

The more proportionate claim is that intelligence-dependent public administration changes the conditions under which legitimate authority is exercised. Institutions must preserve the capacity needed to make their accountability real.

Institutional design implications

Meaningful authority requires more than a human-in-the-loop label. It requires decision rights, escalation thresholds, independent evidence, protected domain competence, access to contestable alternatives and mechanisms through which affected people can challenge consequential decisions.

The test is practical: could the institution recognise material failure, explain the basis for intervention and continue the decision function without surrendering judgement to the supplier?

Effective sovereignty and legitimate sovereignty

Operational capability is necessary but not self-legitimating. Sovereignty claims must disclose for whom control is sought, to what end, who receives the benefits, who bears the risks and how affected people can contest consequential use. Stronger state capability can protect public agency, but it can also insulate surveillance, exclusion or domestic incumbents from challenge.

The Framework therefore supports accountable interdependence rather than immunity from external constraint. External interruption may be arbitrary coercion, ordinary failure, lawful accountability or contested private

governance. The assessment must examine authority, process, proportionality, reciprocity and review rather than treating every interruption as equivalent.

From intelligence to evidence

Institutions must also preserve control over the translation from intelligence into authorised action. A prediction, score or recommendation may be operationally useful without satisfying the evidential standard required for a consequential decision. The institution should distinguish observation, inference, prediction and allegation; preserve provenance and uncertainty; and ensure that an independent reviewer can test the basis of action.

CHAPTER 08

From Proposition to Method

A field centred on assurance must expose its claims to evidence. This paper therefore includes an illustrative instrument rather than referring only to future methods.

The five-question test

The Intelligence Sovereignty Test is presented here as an unvalidated screening instrument. It is not a certification method and should not be used to declare an institution sovereign. Its purpose is to identify where deeper assessment is warranted.

Each question is scored 0, 1 or 2: 0 means no credible evidence; 1 means partial, untested or person-dependent evidence; 2 means established, documented and periodically tested evidence. The aggregate is diagnostic, not comparative, and the dimensions are not fully compensatory. A score of 0 on Evaluate or Continue is a critical flag and requires deeper review regardless of the total score; strength elsewhere cannot compensate for an inability to test consequential outputs or preserve the critical decision or service.

The five questions operationalise the four capabilities in the concise definition. Understand, Evaluate, Govern and Continue map directly to those capabilities; Retain competence tracks Evaluate across time by asking whether the institution is preserving the practised knowledge and judgement needed to evaluate future outputs rather than merely demonstrating present-day capacity.

The numerical screen is supplemented by unscored interpretive checks: whether powers of interruption are legitimate and reviewable; whether affected people can contest the use of intelligence; whether intelligence can be translated into procedurally sufficient evidence; whether benefits and burdens are candidly identified; and whether the governance tier assigned responsibility actually possesses the capability to act.

Illustrative application

Consider the same hypothetical national-identity authority, but an institution that has invested in documentation, governance and continuity while leaving one capability hollow. Its aggregate looks reassuring; a single dimension does not.

The model prioritises identity applications for fraud investigation. It does not make the legal decision, but its ranking materially influences which cases receive scrutiny and delay. The evidence is illustrative and does not assess any real Nigerian institution.

Interpreting the result

A total of 7 out of 10 would, on aggregate, suggest a broadly governed capability. The critical-floor rule overrides that reading. A score of 0 on Evaluate means the institution cannot determine, in any given case, whether the model's ranking is sound - so its documented understanding, its governance thresholds and even its rehearsed fallback all rest on outputs it cannot independently verify. Strong governance without evaluative capability is authority exercised over a black box: the institution can suspend or replace the model, but cannot tell when it should. The floor breach, not the total, is the finding, and it warrants deeper review regardless of the reassuring aggregate.

The appropriate response is not necessarily to build a national model. It is to restore the missing floor: negotiate evidence-access and independent-testing rights, introduce a benchmark dataset, and re-establish reproducible challenge - the specific capability the score of 7 conceals.

Validation requirements

The instrument requires external testing before formal use. Priorities include inter-rater reliability, clarity of evidence criteria, sensitivity to sector and consequence, resistance to gaming, and comparison with established resilience and AI-assurance assessments.

A field becomes credible when methods survive independent use. This illustrative application begins that journey; it does not complete it.

Screening questions

Question	Evidence sought	Score
1. Understand	Can the institution explain the capability, its material assumptions, limitations and role in the decision?	0–2
2. Evaluate	Can qualified people independently test or challenge important outputs using credible evidence?	0–2
3. Govern	Are decision rights, accountability, monitoring and intervention thresholds effective in practice?	0–2
4. Continue	Can the decision or service continue through provider failure, withdrawal, transition or material model deterioration?	0–2
5. Retain competence	Are skills, data access, institutional memory and alternative judgement routes being preserved over time?	0–2

Illustrative score: national-identity fraud prioritisation

Question	Illustrative evidence	Score
Understand	The vendor supplies model documentation, stated assumptions, known limitations and a model card; officials can explain the capability's role and general failure modes.	2
Evaluate	The model is served through a closed API with no case-level explanation, no reproduction environment and no benchmark dataset; the contract does not permit independent testing, so no output can be independently verified.	0
Govern	Decision rights, escalation and suspension thresholds are defined, and were exercised in a live incident.	2
Continue	A manual prioritisation fallback has been designed and rehearsed under a stressed-exit exercise.	2
Retain competence	Experienced fraud analysts remain, but data access is narrowing and no plan protects their practice over time.	1

Illustrative total: 7/10, with a critical-floor breach on Evaluate. This is not a pass/fail threshold and not an assessment of any real institution.

CHAPTER 09

Proportionality, Failure Modes and Misuse

Intelligence Sovereignty is not maximised by eliminating dependence. It is optimised by governing dependence in proportion to consequence.

When sovereignty investment is the wrong choice

Not every intelligence capability warrants substantial internal duplication or national infrastructure. For low-consequence uses, the cost of maintaining alternatives, specialist staff or sovereign hosting may exceed the risk. A mature framework should permit an institution to accept dependency where the decision is reversible, harm is limited and substitution is easy.

The correct objective is not maximum autonomy. It is proportionate freedom of action.

Sustaining evaluative capability against the incentive that erodes it

The Framework's remedy appears to work against its own diagnosis. The Evaluation Paradox is driven partly by efficiency: institutions delegate because doing so is faster and cheaper, and the same pressure can remove the specialists, practice and redundant capacity that preservation requires. Left to local cost-benefit, an institution will tend to under-invest in evaluative capability precisely where a capability is most embedded and most relied upon. A remedy that depends on each institution voluntarily funding competence it hopes rarely to use is therefore unstable by construction.

Safety-critical regulation has confronted this problem by removing retained competence from discretionary cost-benefit for the highest-consequence cases. In the UK nuclear sector, the prime responsibility for safety remains with the licensee, and current ONR guidance expects sufficient licensee core safety and intelligent-customer capability to specify, oversee, question and accept work performed by contractors. The requirement is proportionate to hazard and rests on suitably qualified and experienced people, organisational knowledge, succession and continuing control rather than an assumption that the market will voluntarily preserve client-side expertise (Office for Nuclear Regulation, 2006; Office for Nuclear Regulation, 2024).

Intelligence Sovereignty proposes an analogous treatment for the high-consequence tail of intelligence capabilities. Where a capability materially shapes rights, safety, entitlement or essential services, evaluative competence should be treated as a floor requirement set in proportion to consequence - specified, evidenced and inspected - not left solely to each institution's immediate economics. For reversible, low-impact uses, accepting gradual erosion may remain rational, and mandating retained competence could itself be wasteful. The narrower claim is that for systems whose failure society cannot afford to discover late, evaluative capability is a condition of responsible operation rather than an optional efficiency trade-off.

Protectionism

Sovereignty language can justify market closure, domestic monopoly or insulation from legitimate accountability. The Framework rejects technological nationalism as a default and does not score foreignness as weakness.

External collaboration may increase sovereignty where it builds locally retained competence, plural challenge and credible alternatives. The relevant objective is accountable interdependence: dependencies, interruption powers, evidence rights, accountability obligations and continuity consequences should be transparent, legitimate and governable.

False localisation and sovereignty theatre

A system may be hosted locally yet depend on external updates, proprietary tooling, foreign expertise, concentrated hardware or inaccessible evidence. Domestic branding, model possession and data-centre location are therefore claims requiring examination, not proof of sovereignty.

Prior scholarship has shown that sovereign-AI narratives may recast continuing material dependence as autonomy. The ISF uses sovereignty theatre as an operational failure mode: the production of a credible appearance of autonomy without corresponding evaluative, governing and continuity capability at the point of consequential reliance.

Wasteful duplication

Rebuilding globally available capability without a clear criticality case can divert scarce resources from governance, skills and assurance. Selective internal capability is more credible than a universal build-local mandate.

The Framework should therefore require a reasoned case for investment: consequence, concentration, substitutability, evaluative risk and public value.

Is sovereignty the right frame?

Resilience, assurance, procurement and institutional cognitive resilience each describe part of the problem. The word sovereignty is justified only where the dependency affects practical agency: the institution's capacity to determine its own course, exercise public authority or preserve a critical strategic choice.

The Framework therefore does not use sovereignty as a synonym for quality or robustness. Resilience concerns survival; assurance concerns justified confidence; capability concerns performance; sovereignty concerns whether those conditions preserve effective and legitimate freedom of action.

The naming remains open to consultation, but the case for Intelligence Sovereignty is now narrower: consequential intelligence can become a condition of institutional agency, and the loss of evaluative or continuation capability can hollow out that agency even when formal authority and normal operations remain.

The institutional claim remains an empirical leap

The Framework scales established operator-level mechanisms into a proposed institution-level dynamic. That move is plausible, but direct longitudinal evidence of an AI-induced institutional competence spiral is presently sparse. The paper therefore does not claim that cognitive delegation inevitably produces institutional hollowing-out. It proposes conditions under which such erosion may occur and a research programme capable of distinguishing it from reverse causation, ordinary deskilling and pre-existing weakness.

Independent testing should compare institutions before and after comparable delegation, exploit phased deployments or policy changes where possible, track actual verification performance rather than self-reported confidence, and state null results. A consultation edition may advance the hypothesis; later editions should not elevate it into doctrine unless the evidence survives those tests.

Over-formalisation and self-conferred authority

A further failure mode belongs to the Framework itself. Publication codes, maturity labels and institutional presentation can imply validation before evidence exists. This edition therefore distinguishes internal document control from public authority. It is a proposal under consultation, not a standard, licence or certification scheme.

The language of assurance obliges the Framework to apply the same restraint to itself that it asks of others.

CHAPTER 10

Nigeria and the Next Frontier

Nigeria's software-quality regime creates a practical platform for the next inquiry: how should public institutions assure and retain competence over decision-shaping intelligence?

Build on the regime that now exists

Nigeria is the policy setting for this paper, not evidence that the proposed Framework is valid. No Nigerian institution is assessed here, and the five candidate domains are hypotheses for controlled application rather than claims about current capability. The immediate opportunity is to connect the formally launched National Software Quality Assurance Framework with AI assurance, operational resilience and institutional competence. NITDA's Class A/B/C risk model provides a natural entry point: the more consequential the system, the more deeply government should examine not only whether it works, but whether its intelligence-producing arrangement remains understandable, evaluable, governable and continuous.

A small number of controlled pilots would be more valuable than a national declaration. Candidate domains include identity, tax, health, financial supervision and agricultural decision support. Each pilot should map the intelligence capability, evidence, supplier dependencies, internal expertise and credible continuity routes.

Preserve evaluative capability

The most important policy intervention may be human rather than infrastructural. Public institutions should protect professional pathways through which domain experts continue to practise judgement, verify model output and learn from failure. Procurement should include evidence access, benchmark rights, knowledge transfer and support for independent evaluation.

For high-consequence systems, that protection should move beyond aspiration. A proportionate regulatory, licensing or assurance floor should require evidence that critical evaluative capability is staffed, practised, succession-planned and periodically inspected. The obligation should attach to consequence rather than to AI use in general, preserving lighter treatment for reversible, low-impact applications and guarding against the compliance theatre described in Chapter 9.

The assurance profession has a distinctive role. Testers and quality leaders already know how to ask what evidence supports confidence, what remains uncertain and who accepts residual risk. Their mandate should expand into model evaluation, data quality, decision assurance and third-party intelligence risk.

A research and consultation programme

The revised Framework should now be exposed to external challenge. Priority reviewers include AI-sovereignty scholars, operational-resilience regulators, public-law experts, human-factors researchers, government technology leaders and practising assurance professionals.

Research should test the Evaluation Paradox, refine the concise definition, validate the five-question screen and examine whether Intelligence Sovereignty adds explanatory value beyond institutional cognitive resilience, intelligent-customer capability and operational resilience. Sector applications should publish both positive and negative findings. The objective is not to protect the Framework from criticism, but to discover whether it is useful and whether its chosen name earns the additional claims it carries.

Closing argument

Software Testing established confidence in system behaviour. Quality Assurance broadened that confidence to the processes and institutions that produce quality. AI Assurance addresses the distinctive risks of intelligent systems. Intelligence Assurance asks whether decision-shaping outputs deserve reliance.

Intelligence Sovereignty asks whether that reliance can remain subject to meaningful institutional judgement over time. Its distinctive concern is not possession of the technology, but preservation of the capability to evaluate and govern what the technology tells us.

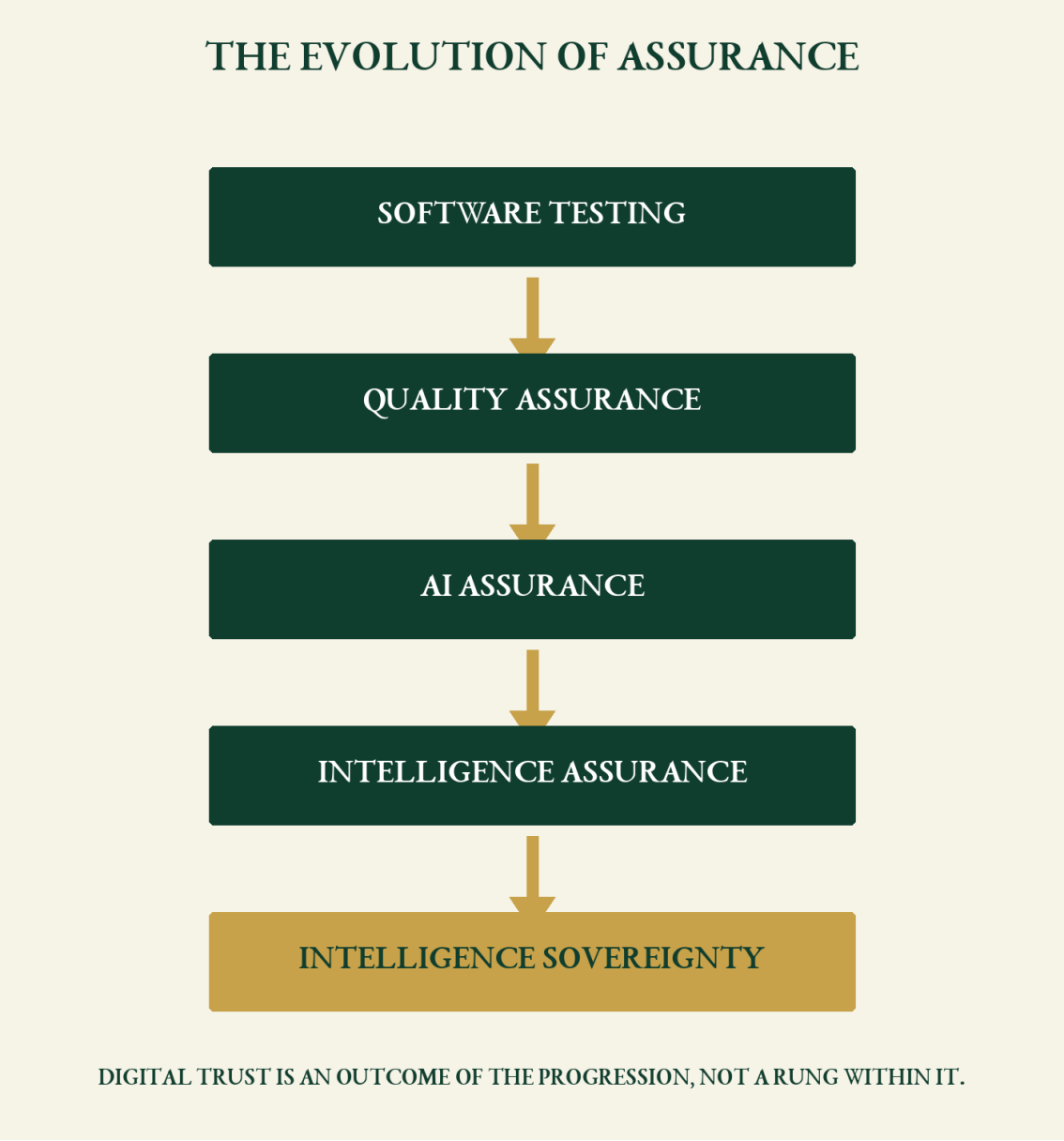
The defining governance question of the AI age is no longer whether humanity can build increasingly intelligent systems. It is whether nations and institutions retain the enduring freedom to understand, evaluate and govern the intelligence on which those systems - and increasingly society itself - depend. This edition calls that challenge Intelligence Sovereignty while inviting scrutiny of whether institutional cognitive resilience is the more precise frame. The label remains provisional. The governance problem does not.

CHAPTER 11

Appendices and Reference Material

The following tools make the revised proposition inspectable without implying validation.

Appendix A – The five-stage progression



The progression is cumulative. Digital trust is an outcome rather than an additional stage.

Appendix B – Intelligence Sovereignty Test

Question	Minimum evidence proposition
Understand	We can explain the role, assumptions, limitations and evidence of the capability.
Evaluate	We can independently test or challenge consequential outputs.
Govern	We can intervene, suspend, change or constrain reliance under

Question	Minimum evidence proposition
	accountable authority.
Continue	We can preserve the decision or service through failure, withdrawal, deterioration or transition.
Retain competence	We preserve the people, practice, data access and institutional memory needed for meaningful challenge.

Scoring: 0 = no credible evidence; 1 = partial, untested or person-dependent evidence; 2 = established, documented and periodically tested evidence. The screen is illustrative and unvalidated. Scores are diagnostic rather than comparative. A score of 0 on Evaluate or Continue triggers deeper review regardless of the aggregate score.

Appendix C - Worked assessment interpretation

The illustrative identity-fraud example produced 7/10, but a score of 0 on Evaluate triggered the critical-floor rule. The aggregate appears reassuring while the institution remains unable to determine whether an individual ranking merits reliance. The floor breach, not the total, is the finding. The proportionate response is to restore independent evaluation through evidence-access and testing rights, benchmark data, reproducible challenge and protected specialist practice - not automatically to build a national model.

Appendix D - Relationship to adjacent fields

Discipline	Governing question	Primary object
Software Testing	Does the software behave as intended?	System behaviour and defect evidence
Quality Assurance	Do processes, governance and evidence justify confidence?	Organisational conditions that produce quality
AI Assurance	Is reliance on the AI system justified across its lifecycle?	Model, data, impact, safety and accountability
Operational Resilience	Can the important service remain within tolerance through disruption?	Service continuity and third-party dependencies
Sovereign AI	Does the actor retain AI capability, agency and choice?	AI stack, capacity, access and control
Intelligent Customer / Client	Can the organisation specify, oversee and critically judge external delivery?	Retained client competence and supplier oversight
Intelligence Sovereignty	Does the actor retain evaluative and operational authority over consequential judgement?	Intelligence-producing arrangement and institutional competence

CHAPTER 12

References and Source Note

This edition deliberately engages the adjacent literature on sovereign AI, operational resilience and cognitive delegation.

Al-Anezi, F. (2026). Generative artificial intelligence in healthcare: Automation bias, deskilling, and cognitive implications - a systematic review. Journal of Healthcare Leadership.

Bainbridge, L. (1983). Ironies of automation. Automatica, 19(6), 775-779. https://doi.org/10.1016/0005-1098(83)90046-8

Bank of England. (2026). Updated outsourcing and third-party risk management supervisory statements for central counterparties and recognised payment system operators.

Central Computer and Telecommunications Agency. (1994). Managing contracts for IS/IT services: The role of the intelligent customer. HMSO. ISBN 0-11-330644-X.

Cobbe, J., Lee, M. S. A., & Singh, J. (2020). Reviewable automated decision-making. Computer Law & Security Review, 39, 105475. https://doi.org/10.1016/j.clsr.2020.105475

Council of Europe. (2024). Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law.

European Parliament and Council. (2022). Regulation (EU) 2022/2554 on digital operational resilience for the financial sector (DORA), especially Article 28(8).

Financial Conduct Authority. (2026). Outsourcing and operational resilience. Updated 14 July 2026.

Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: A systematic review of frequency, effect mediators, and mitigators. Journal of the American Medical Informatics Association, 19(1), 121-127.

Goh, E., Gallo, R., Hom, J., et al. (2024). Large language model influence on diagnostic reasoning: A randomized clinical trial. JAMA Network Open, 7(10), e2440969. https://doi.org/10.1001/jamanetworkopen.2024.40969

Goh, E., Gallo, R. J., Strong, E., et al. (2025). GPT-4 assistance for improvement of physician performance on patient care tasks: A randomized controlled trial. Nature Medicine, 31, 1233-1238. https://doi.org/10.1038/s41591-024-03456-y

Government Property Function. (2022). Intelligent Client Roles: Functional Guidance. Cabinet Office. https://www.gov.uk/government/publications/intelligent-client-roles-functional-guidance

Health and Safety Executive. (n.d.). Intelligent customer capability. https://www.hse.gov.uk/humanfactors/topics/customers.htm

Huang, L., & Vishnoi, N. K. (2026). The geometry of learning under AI delegation. arXiv working paper 2603.02950.

Huang, L., Xiao, W., & Vishnoi, N. K. (2026). Delegation and verification under AI. arXiv working paper 2603.02961.

Internet Governance Forum. (2023). DC-DAIG session report: Can (generative) AI be compatible with data protection?

Lyell, D., & Coiera, E. (2017). Automation bias and verification complexity: A systematic review. Journal of the American Medical Informatics Association, 24(2), 423-431.

National Information Technology Development Agency. (2026a). National Software Development Guideline; National Software Testing Guideline; Software Testing Organisations Licensing Guideline. Consultation instruments published April 2026.

National Information Technology Development Agency. (2026b). National Software Quality Assurance Framework. Formally launched 29 July 2026 under the NITDA Act 2007; phased to full effect in Q2 2027.

National Institute of Standards and Technology. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0).

Office for Nuclear Regulation. (2006). Safety Assessment Principles for Nuclear Facilities: 2006 Edition, Revision 1. Historical edition, superseded by later SAPs. https://www.onr.org.uk/media/aqgedje3/saps2006.pdf

Office for Nuclear Regulation. (2024). NS-TAST-GD-049: Licensee Core Safety and Intelligent Customer Capabilities (Issue 7.2). Published 1 July 2024; review date July 2029. https://www.onr.org.uk/publications/regulatory-guidance/regulatory-assessment-and-permissioning/technical-assessment-guides-tags/nuclear-safety-tags/ns-tast-gd-049

Ononuga, O. O. (2025). AI and the Full Software Testing Lifecycle: Replacement or Revolution? Keynote address to the Association of Nigeria Software Testers.

Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. Human Factors, 52(3), 381-410. https://doi.org/10.1177/0018720810376055

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. Human Factors, 39(2), 230-253. https://doi.org/10.1518/001872097778543886

Pavlidis, G., & Kastanas, I. (2026). Algorithmic administration and the EU AI Act: Legal principles for public sector use of AI. arXiv:2604.22765. https://arxiv.org/abs/2604.22765

Roehl, U. B. U., & Hansen, M. B. (2024). Automated, administrative decision-making and good governance: Synergies, trade-offs, and limits. Public Administration Review, 84(6), 1184-1199. https://doi.org/10.1111/puar.13799

United Nations. (2026). Open Source Week transcripts on AI sovereignty, interoperability, provider dependence, skills and internal governance capacity.

Balston, G. (2026). AI Sovereignty and National Security: Identifying the UK's Dimensions of Control. Centre for Emerging Technology and Security.

Dibiaggio, L., Nesta, L., & Vannuccini, S. (2024). European Sovereignty in Artificial Intelligence: A Competence-Based Perspective. Sciences Po Chair Digital, Governance and Sovereignty.

Edler, J., Blind, K., Kroll, H., & Schubert, T. (2023). Technology sovereignty as an emerging frame for innovation policy: Defining rationales, ends and means. Research Policy, 52(6), 104765.

Schliessmann, C. P. (2026). AI Sovereignty as Control Sovereignty. SSRN working paper.

Scassa, T. (2024). Sovereignity and the Governance of Artificial Intelligence. UCLA Law Review Discourse, 71, 214.

Muegge, D. (2024). EU AI sovereignty: for whom, to what end, and to whose benefit? Journal of European Public Policy.

Bainbridge, L. (1983). Ironies of automation. Automatica, 19(6), 775-779.

Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. Human Factors, 37(2), 381-394.

Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: a systematic review of frequency, effect mediators, and mitigators. Journal of the American Medical Informatics Association, 19(1), 121-127.

Lyell, D., & Coiera, E. (2017). Automation bias and verification complexity: a systematic review. Journal of the American Medical Informatics Association, 24(2), 423-431.

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: a failure to disagree. American Psychologist, 64(6), 515-526.

NITDA source note

NITDA formally launched the National Software Quality Assurance Framework on 29 July 2026. It is adopted and signed, with full effect phased to the second quarter of 2027. At the time of this edition, older constituent documents could still be found within NITDA's draft-instrument pages; that website lag does not alter the status of the signed Framework. This paper therefore distinguishes formal adoption from full operational effect.

About the author

Onosoji O. Ononuga is an assurance and governance practitioner with more than two decades of experience across software testing, quality assurance, programme delivery and regulated transformation. He is the author and originator of the proposed Intelligence Sovereignty Framework presented in this series. Its claimed contribution is the institutional framing of evaluative capability and its connection, where consequence warrants, to continuity of public judgement under external dependence. His current work examines the assurance of cognitive output, retained institutional competence and the governance of intelligence dependencies.

About Cereark

Cereark Quality, Risk & Assurance Consortium develops practical approaches to software quality, evidence-ready delivery, AI assurance and institutional resilience. Intelligence Sovereignty is an emerging research programme within that body of work, issued for external challenge rather than presented as an established standard.

Consultation invitation

Comments are invited from policymakers, regulators, sovereign-AI scholars, intelligent-customer and procurement practitioners, operational-resilience specialists, public-law experts, human-factors researchers, assurance professionals

and institutional leaders. Priority questions are whether sovereignty is the right frame or whether the paper principally describes institutional cognitive resilience; whether the proposed distinction adds anything beyond the intelligent-customer function; whether the institution-scale Evaluation Paradox is empirically identifiable rather than ordinary skill atrophy or reverse causation; whether mandated evaluative-capability floors proportionate to consequence are the right instrument or would ossify into compliance theatre; whether the capability-output referent is sufficiently clear; and whether the five-question screening instrument produces consistent and appropriately non-compensatory decisions.

CHAPTER 13

Publication Control Record

Revision history is separated from the current-edition QA controls so that the controls below describe the present publication rather than accumulate as a changelog.

Revision history

v1.1 candidate (4 August 2026): rebuilt literature positioning; narrowed originality boundary; adopted Institutional Evaluation Paradox terminology; added accountable interdependence, effective-versus-legitimate sovereignty, control-mechanism versus control-capability, intelligence-to-evidence translation and unscored interpretive checks; retained the controlled definition and five-question screen.

Edition	Public Consultation Edition v1.1
v0.9	Rebuilt as a restrained consultation edition following hostile expert review.
v0.9.1	Added the human-factors lineage, explicit Evaluate capability, evidence caveat and critical scoring floors.
v0.9.2	Mapped the five-question screen to the four core capabilities.
v0.9.3	Engaged the intelligent-customer prior, clarified the referent, strengthened Chapter 7 and rebuilt the worked example.
v0.9.4	Added the incentive-and-sustainability mechanism, ONR precedent and explicit framing consultation question.
v0.9.5	Updated NITDA status from consultation instruments to the formally launched and signed national Framework.
v0.9.6	Completed production regression: separated revision history from QA controls, alphabetised references and clarified illustrative Class A examples.

Current-edition QA controls

Status	Candidate for Principal approval; revised public consultation edition
PASS	Publication status is Revised Consultation Edition v0.9.6; the earlier v1.0 remains superseded and not for circulation.
PASS	NITDA status is stated as formally launched and signed on 29 July 2026, with implementation phased to full effect in Q2 2027.
PASS	Core banking switches, national identity infrastructure and power-grid control are identified as illustrative high-consequence examples, not attributed as NITDA's published Class A list.
PASS	Novelty and scope are tested against digital sovereignty, sovereign AI, AI assurance, operational resilience and the intelligent-customer function.
PASS	The Evaluation Paradox is situated in the Bainbridge and automation-bias lineage and presented as an institution-scale hypothesis requiring empirical validation.
PASS	The concise definition, capability-output referent and four core capabilities are internally consistent.
PASS	The five-question screen is mapped to the four capabilities; Retain competence is the longitudinal safeguard for Evaluate.
PASS	Critical scoring floors and the 7/10 illustrative case demonstrate why aggregate scores are non-compensatory.
PASS	Proportionality, protectionism, false localisation, wasteful duplication, incentive failure and compliance-theatre risks are addressed without adding further caveats.
PASS	Administrative-law, human-factors, intelligent-client and nuclear intelligent-customer sources are

Status	Candidate for Principal approval; revised public consultation edition
	represented in the relevant argument.
PASS	References are alphabetised; the CCTA 1994 bibliographic date and Government Property Function 2022 publication date have been checked.
PASS	The paper makes no claim of validated standard, certification, licensing authority or completed external validation.
PASS	Reader-facing edition number, suggested citation, publication metadata, running matter and control record are consistent at v0.9.6.
PASS	Figures, tables, headings, page flow and reference pages have been rendered and visually inspected.
PASS	Accessibility audit completed with no unresolved high-, medium- or low-severity findings.

INTELLIGENCE SOVEREIGNTY PAPERS

Public Consultation Edition v1.1

A proposed assurance and governance discipline for preserving institutional judgement under intelligence dependence.

FREEDOM TO GOVERN THE INTELLIGENCE THAT SHAPES OUR FUTURE