THE
kyndryl
institute

Trend topic: **Sovereignty**

By

Audrey Tang[1]

Senior Accelerator Fellow, Institute for Ethics in AI, University of Oxford

Audrey Tang is a globally recognized technologist and civic innovator working at the intersection of AI, ethics, and governance. She is a Senior Accelerator Fellow at the Institute for Ethics in AI at the University of Oxford, where her work focuses on designing accountable, human-centered approaches to AI deployment. As a software engineer, Audrey is a leading contributor to open-source communities and shapes international discussions on digital trust, participatory systems, and responsible technology. In 2023, TIME named her in its inaugural list of the 100 Most Influential People in Artificial Intelligence.

[1]Image illustration based on photo at audreyt.org, licensed CC BY-NC-SA 4.0 Kaii Chiang

Explore more from The Kyndryl Institute
kyndryl.com/institute

Safeguarding sovereignty with data as soil

In times of profound global uncertainty, when geopolitical fault lines shift, supply chains strain, and regulatory borders harden, the reflex of the multinational enterprise is to consolidate.

The goal becomes what I call the corporate singleton: a single cloud-based AI layer into which all proprietary global data must be funneled, promising a God's-eye view that can predict and manage the chaos. Hyper-centralization is not resilience. It is a single point of failure.

For years, the technology sector has repeated the mantra that "data is the new oil," an inert resource to be extracted, pumped to a central refinery, and monetized. We must abandon this metaphor. **Data is not oil. Data is soil.** It is the living ground containing the nutrients in which decisions take root. Different soils nurture different crops because they are shaped by different climates, microbes, histories, and patterns of care. Data is similarly relational: tied to the culture, constraints, and realities of the place where it was generated. Extract it from its ecosystem and much of its meaning goes with it.

But even soil is not the whole story. An organization can keep its data beautifully local and still surrender the thing that matters most: judgment. It can pass every residency test while outsourcing interpretation, prioritization, and action of that data to an AI system not designed or governed

for that organization, trained elsewhere, governed elsewhere, and optimized for someone else's risk appetite. Here AI architecture becomes decisive in keeping data sovereign. The architecture determines whether local data is interpreted by a locally accountable model operating within local rules, or by a remote general model whose assumptions travel across jurisdictions without their accountability traveling with them. **The first wave of data sovereignty asked, "Where does data live?" The next must ask, "Where does decision-making live?"** That is what I mean by decision sovereignty: not simply control over data storage, but control over how AI may interpret data, what actions it may trigger, who may challenge those actions, and how authority can be brought back in-house when conditions change.

This frames AI not as the subject of sovereignty, but as the mechanism through which sovereignty is exercised or surrendered.

Challenging the Mainframe Mentality

Many enterprises are approaching AI with a mainframe mentality: one giant enterprise copilot in the cloud, one model connected to every document, workflow, and application, everyone becoming a terminal typing into the same intelligence.

This looks efficient, just as the mainframe once did. But it also creates monoculture risk, central honeypots, hidden jurisdictional dependencies, and organizational atrophy. Local nuance is reduced to prompt engineering. Local authority becomes a permissions spreadsheet. In calm conditions, this feels convenient. In crisis, it becomes a bottleneck. The global model that performs brilliantly on last quarter's patterns becomes a liability the moment a new tariff, a new conflict, or a new regulation changes the local rules overnight.

This frames AI not as the subject of sovereignty, but as the mechanism through which sovereignty is exercised or surrendered.

Having a single point of control that concentrates risk as well as capability is not the answer. Nor is fragmentation. The answer is bounded delegation.

From Kami to Steward: A Model for Decision Sovereignty

In the Japanese Shinto tradition, a Kami (神) is a guardian associated with a particular place: a grove, a river, an old tree. Its role is not universal command. It is care for one ecosystem. Kami is akin to a form of enterprise governance. To make the practical meaning plain, I will call the enterprise equivalent a steward: a locally accountable role, embodied by a person, a team, or a bounded system, tasked with operating an AI agent within a defined mandate for a defined community. The river's guardian does not try to govern the forest. Likewise, the steward of a hospital triage system should not become the steward of payroll, personnel or procurement.

At Oxford's Institute for Ethics in AI, my colleague Caroline Green and I have built a Civic AI governance framework called “the 6-Pack of Care.” The core idea is boundedness: different tasks require different scopes of authority, responsibility, and review.

The organizations that manage mission-critical infrastructure are often already practicing this kind of stewardship. They simply do not yet name it as such.

Jun Murai argued in his recent essay for The Kyndryl Institute that connection is resilience and that effective governance is vital. I agree. But the critical question is not only about effective governance; it’s by whom. That role belongs with the local steward operating under a charter, not with an unbounded model at the center.

Consider what a large enterprise IT operator already does. It serves a bounded community: the patients of a hospital network, the depositors of a regional bank, the commuters of a transit system. It runs under explicit service-level agreements. It maintains audit trails and incident-response protocols. It practices graduated rollout by testing changes on a small scale first, running them safely in parallel, and retaining the ability to reverse course quickly. It hands systems over when contracts end. These are the operational habits of bounded stewardship.

What these organizations often lack is not operational discipline. It is a governance layer that makes AI authority explicit, local, and contestable.

The Sovereignty Stack

For organizations deploying AI in core decision-making, decision sovereignty means being able to determine not only where data resides, but how it is interpreted, what actions may follow, how decisions can be reviewed, and how the system can be paused or replaced. One useful way to think about this is as a stack with five layers:

1. **Residence** — where the data lives.
2. **Access** — who may see or use it.
3. **Inference** — which AI model or analytical system is allowed to draw conclusions from it.
4. **Actuation** — what operational systems are allowed to do because of those conclusions, such as sending messages, approving payments, routing cases, or shutting down a machine.
5. **Exit** — how the organization pauses, migrates, or retires the system without losing continuity, records, or accountability.

A surprising number of current AI programs stop at the first two layers. The greatest enterprise risk now sits in the third and fourth. Data can remain local while a general model elsewhere decides how to triage a patient, sequence a supply chain, investigate an employee, or shut down a machine. Location alone does not protect judgment. And without the fifth layer, exit, an organization that has delegated inference and actuation to a vendor has no credible way to take them back in moments of failure, regulatory change, or crisis, when continued operation is no longer acceptable.

Consider a multinational insurer operating claims-processing systems in Japan, Brazil, and Germany. Under the conventional approach, all three feed a single large foundation model in a US data center: a general-purpose AI model trained broadly and then adapted to enterprise tasks. The model is powerful but generic. It does not understand that Japanese policyholders expect a different tone and cadence in claims correspondence than Brazilian ones. It cannot easily comply with Germany's stricter data-processing requirements without layers of transformation and validation before the model can even operate on the data. When it makes an error, misclassifying a claim or denying coverage that local regulation requires, the appeals process routes through a global team with no relationship to the affected community.

The bounded alternative is different. Each jurisdiction runs a smaller domain-specific model fine-tuned on local claims data, regulatory language, and interaction patterns. These models share a common evaluation framework so global risk teams can compare performance and share lessons. But interpretation, enforcement, and appeals stay local. Each model can be audited, retrained, paused, or retired independently.

Federate Safety, Do Not Centralize It

Bounded stewards must still cooperate without recreating the centralized fragility we set out to escape. Federation is the answer: not one global safety authority, but many local stewards sharing threat intelligence and evaluation methods under common protocols while keeping enforcement local.

This is already happening. The ROOST initiative, Robust Open Online Safety Tools, launched at the Paris AI Action Summit in February 2025, demonstrates the model. Partners such as Bluesky, Roblox, and Discord each train local AI models to detect harm within their own context. Threat signals are shared via federated learning. Safety is tuned to local norms without being dominated by a single corporate policy.

The test of whether such a federation is working is simple: after a disruption, do local teams and partners trust the system more, or less? In a volatile world, whether trust rises or falls after a crisis is a core business-continuity metric.

The Courage to Sunset

Enterprise infrastructure is beset by what I call imperial creep: systems that expand beyond their mandate because they were built to live forever. Here stewardship should mean operating a system within a clearly bounded purpose. When that purpose is never explicit, never reviewed, and never brought to an end, stewardship turns imperial. The assistant that schedules maintenance quietly becomes a procurement agent. The fraud model quietly becomes an employee-monitoring system. The service-desk assistant quietly becomes a behavioral scoring engine. Nobody authorized these expansions. They happened because nobody defined the boundary or the ending.

A bounded enterprise steward is designed with a sunset clause: strict resource caps, a specific mandate, and an expiration or review date. When a temporary joint venture concludes, a post-merger integration finishes, or a specific crisis subsides, the steward retires gracefully. It archives institutional memory, hands over decision traces and evaluation results to its successor, and powers down. This is how an organization preserves accountability without carrying unnecessary sensitive context indefinitely.

An AI agent that cannot be shut down without breaking the enterprise is not a tool. It is a liability. The organization that can hand off gracefully is the organization that clients will trust with the next deployment.

Therefore, I would encourage leaders to do the following:

1. **Give every AI system a charter.** No deployment goes live without a named owner, a published scope, a pause trigger, and a handover plan. Without these, it is ungoverned infrastructure, regardless of how sophisticated the model is.

2. **Reframe data residency as a design asset.** Invest in local fine-tuning. A model trained on the specific patterns of your Japanese insurance claims, your Brazilian payroll regulations, or your German energy telemetry will often outperform a generic model while satisfying residency requirements.

3. **Write charters before RFPs.** Convene the people who will be affected. Use structured deliberation to define scope, red lines, severity classes, and remedies. Make these the basis of the procurement specification. Any scope expansion should require fresh authority.

4. **Join a federated safety network.** Extend cybersecurity information-sharing to cover AI-specific incidents. Share proofs, model cards, and evaluation results, not raw data. The goal is shared defense without shared dependence.

5. **Design exit before scale.** Every deployment needs a sunset clause and a succession plan. If you cannot switch models, clouds, or vendors without losing continuity, you do not have sovereignty. You have managed dependence.

The Plurality of the Enterprise

The assumption that a single global intelligence layer is the answer needs reexamining, especially in fragile or highly regulated environments. The organizations that thrive will not be those that build the thickest walls or the largest central databases. They will be those that treat data as living soil, appoint local stewards to tend it, and measure success by whether trust, coordination, and accountability grow stronger under pressure, not by whether one model can answer every question from every continent.

The resilient enterprise will look less like a single brain and more like a federation of trusted stewards: locally accountable governance roles, each tending its own soil, sharing what it learns, and knowing when to let the next season begin.

References

1. Audrey Tang and Caroline Green, Civic AI — 6-Pack of Care, Institute for Ethics in AI.