A 42 CONSULTING.AI WHITEPAPER

Unknown Data Is Ungoverned Data

Why an AI System Is Only as Governed as the Data It Was Built On

Dr M Maruf Hossain, PhD, GAICD
Chief AI Strategist · 42 Consulting.AI
enquire@42consulting.ai · www.42consulting.ai

42 Consulting.AI
Scaling Intelligence

CONTENTS

Table of Contents

FOREWORD

From the Author.

Every claim an organisation makes about its AI — that it is fair, that it can be explained, that it can be challenged — rests on a quieter claim it rarely examines: that it knows what the system was built on. An AI system is a function of its data; its behaviour, its biases, and its blind spots are learned from the examples it was trained on. It follows that an organisation can account for a system only as far as it can account for that data. Unknown data is not a smaller problem than bias or opacity. It is the ground the others stand on.

This is easy to miss because a system trained on data of unknown provenance does not announce the fact. It runs, produces outputs, passes its functional tests, and looks exactly like a system with fully documented data. The gap is invisible in operation and surfaces only when someone asks a question the organisation cannot answer: where did this data come from, were these people's records used on a proper basis, does this dataset represent the population the system decides about? "We are not sure" is, in governance terms, the same as "no".

This paper is about why that gap is disqualifying rather than merely undesirable. It is the difference between a system the organisation has governed and one it has merely operated — and it shows why a provenance gap cannot be accepted as a managed risk, why the obligation extends to data the organisation did not collect, and why knowing the data is the precondition for every other assurance the framework provides.

It is written for the board and the executive, who have been told their AI is "well governed" and want to know whether that governance extends to the data underlying it. Can the organisation say, for each system that decides about people, where its training data came from and on what basis? If the honest answer for any material dataset is that no one knows, then that system is not governed — it is running, and the organisation is accountable for outcomes it cannot account for.

M Maruf Hossain

SECTION ONE

Executive Summary.

An AI system can be no better governed than the data it was built on. An organisation can assess fairness, generate explanations, and offer contestation — and if the underlying data is of unknown origin, each of those is a claim it cannot actually stand behind. This whitepaper establishes three propositions.

First, unknown data is ungoverned data — a provenance gap is a barrier, not a risk. To govern an AI system is to be able to account for what it does, and an organisation cannot account for a system whose training data it cannot trace. "We do not know where this dataset came from" is not a risk to be rated and accepted. It is a deployment barrier: a system trained on data of unknown provenance must be rebuilt on documented data before deployment, or not deployed at all.

Second, provenance is the precondition on which the other assurances depend. Fairness, explainability, and contestation are not independent of the data — they presuppose it. A fairness assessment measures disparities across a population that the organisation must be able to characterise; an explanation attributes a decision to factors that it can vouch for; a contestation reviews a basis that it can reconstruct. On data of unknown provenance, each looks satisfied on its own gate while being silently void. Data governance is not the fourth control. It is the ground beneath the other three.

Third, the obligation reaches data that the organisation did not collect. Accountability for a system's data does not transfer to whoever supplied it. Where a system is trained on vendor data, or built on a foundation model whose training data is opaque, the organisation is still accountable for the provenance it cannot see — and "the data is the vendor's" is not a discharge of the obligation but a reason to treat the system's provenance as unverified until proven otherwise. A system whose data cannot be traced sits outside the governance perimeter, however it got there.

This paper sets out why governance requires traceable data, why a provenance gap is a barrier rather than a managed risk, what must actually be known before a system is deployed, why the obligation extends to data the organisation did not collect, and why knowing the data is the precondition beneath every other assurance. It ends with five questions a board should be able to answer about the data under its AI — not next quarter, but today.

SECTION TWO

You Can't Govern *What You Can't Trace.*

To govern a thing is to be able to account for it — to say what it will do, why, and within what limits, and to answer for it when asked. An AI system is built from data, and its behaviour is learned from the examples it was trained on. It follows that an organisation can account for a system only to the extent that it can account for that data. Where the data is traceable — its origin known, its basis documented, its quality assessed — the organisation has a foundation on which to make claims about the system. Where it is not, it has none.

This is not a matter of degree that better testing can close. Testing a model tells you how it behaves on the cases you test; it does not tell you what is in the data that shaped it, what populations are missing from it, or whether the people in it consented to its use. A system can pass every functional test and still be built on data the organisation is not entitled to use, that under-represents the people it decides about, or that came from a source no one can now identify. The problems that live in the data are precisely the ones that do not surface in ordinary testing.

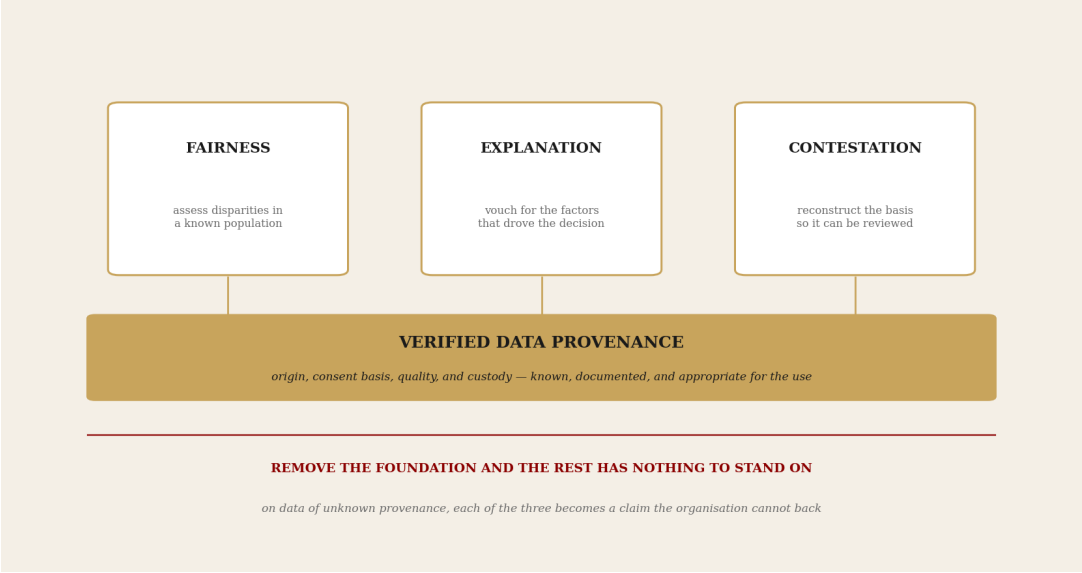

Figure 1 — The precondition. Fairness, explanation, and contestation each rest on verified data provenance; remove that foundation and each becomes a claim the organisation cannot back.

So data provenance is not one governance activity among several — it is the foundation on which the others are built. A fairness assessment is only as trustworthy as the organisation's knowledge of the population the data represents. An explanation is only as honest as its grasp of what actually drove the model, which is a function of what the model learned from. A contestation is only as substantive as the decision record that can be reconstructed, which depends on knowing the

inputs. Each downstream assurance presupposes the one upstream: that the organisation knows its data.

This is what it means to say unknown data is ungoverned data. Not that unknown data is risky — though it is — but that a system built on it cannot be governed at all, because governance is the capacity to account for the system, and that capacity is absent at the root. The organisation can still operate such a system; it cannot govern it. And it remains fully accountable for what the system does, yet lacks the means to answer for it — which is the worst position a governed enterprise can be in.

The rest of this paper follows from this. If a system built on untraceable data cannot be governed, then a provenance gap cannot be a risk the organisation accepts and manages — it must be a barrier. If provenance is the precondition, then the obligation to establish it must reach every material dataset, including those the organisation did not collect. And if the whole edifice of responsible AI rests on knowing the data, then knowing the data is where governance has to begin.

SECTION THREE

The Barrier, *Not the Risk.*

The most consequential decision an organisation makes about data provenance is a categorical one: is a provenance gap something to be managed, or something that stops deployment? Treated as a risk, an unknown-provenance dataset is rated, documented, assigned an owner, and — under delivery pressure — accepted, because the model works and the launch is due. Treated as a barrier, the same gap simply halts the system until it is closed. The framework treats it as a barrier, and the distinction lies between governing the data and merely noting its absence.

The reason a provenance gap cannot be a managed risk is that it is not a known risk. A risk is a possibility whose likelihood and impact can be estimated, and that estimate is what makes it manageable. But an organisation that does not know a dataset's origin cannot estimate what it might contain — whether it embeds a bias, whether it was collected without a lawful basis, whether it misrepresents a population. The uncertainty is not quantifiable because the very information needed to quantify it is what is missing. "Accept the risk" presupposes a risk you can size; a provenance gap denies you that.

So the framework is categorical. "We do not know the provenance of this dataset" is not an acceptable risk acceptance — it is a deployment barrier. A system trained on data of unknown provenance must be retrained on data with documented provenance before deployment, or it is not deployed. There is no rating high enough, and no compensating control strong enough, to convert missing knowledge into governed data. The only remedy is to establish the provenance or to remove the data.

This is stricter than ordinary risk management, and deliberately so. Everywhere else, the organisation weighs and accepts; here it does not, because the thing in question is not a hazard to be traded off but the precondition for weighing anything at all. An organisation cannot responsibly accept a fairness risk it cannot see, an explanation it cannot verify, or a contestation it cannot support — and all three become unassessable the moment the data is unknown. Making the provenance gap a barrier is how the framework refuses to build the rest of its governance on ground it cannot stand on.

SECTION FOUR

What Must *Be Known.*

If a provenance gap is a barrier, then "knowing the data" has to mean something specific enough to test — otherwise every organisation would claim to know its data, and none could be held to it. The framework specifies what must be established, documented, and available for audit before a dataset may be used. Three things must be known, and a fourth assessed.

Where it came from — origin and custody. The organisation must be able to say where a dataset originated, how it was collected, and the chain of custody and transformation from its source to the form on which the model was trained. This is what makes the data traceable: not a vague sense of its source, but a documented lineage sufficient to reconstruct and audit how it became what it is. A dataset whose custody chain has a gap is, to that extent, of unknown provenance.

On what basis — consent for this use. Where a dataset contains personal information, the organisation must be able to demonstrate a consent basis that specifically covers its use for AI training in the domain in which the system operates. Generic consent to collect data is not consent to train a model on it. The basis must be assessed and documented before the data is used, because a lawful-looking dataset used without an adequate basis is a governance failure that no amount of model performance can repair.

Whether it is fit — of sufficient quality for the use. The data must be assessed against the dimensions that determine whether it can support the decisions the system makes: completeness, so it represents the population without material gaps; accuracy, so it reflects what it records; currency, so it is recent enough for use; and representativeness, so it does not systematically under-represent the people the system will affect. Data that fails these is not merely lower-quality — it produces a model whose flaws are baked in and invisible during standard testing.

None of this is discharged once and forgotten. Provenance, basis, and quality can lapse: a consent basis can expire or be withdrawn, a dataset can go stale as the world it describes moves on, a data-sharing right can terminate. So the framework attaches retirement triggers to each — conditions under which a system must be retrained or suspended because the data beneath it is no longer governed. Knowing the data is not a certificate obtained at deployment; it is a state the organisation must maintain for as long as the system operates on that data.

What must be known is set out below. The point of specifying it this precisely is that every organisation would otherwise claim to know its data, and none could be held to it. Each row is a condition that can be tested and evidenced before a dataset is used.

What must be known	The test it must pass	What fails it	What retires it
Origin and custody	A documented lineage from source to the form the model was trained on, sufficient to reconstruct and audit	A vague sense of the source; a custody chain that cannot be reconstructed	A data-sharing right terminates
Consent basis	Consent that specifically covers AI training in the domain the system operates in	Generic consent to collect, treated as consent to train	Consent expires or is withdrawn
Quality and fitness	Completeness, accuracy, currency and representativeness, assessed against the decisions the system makes	Systematic under-representation of the people the system will affect	The dataset goes stale as the world it describes moves

Table 1 — What must be established before a dataset may be used, and what causes it to lapse.

The final column is the part most often missed. Provenance, basis and quality are not discharged once and filed. Each can lapse without anyone acting, which is why the framework attaches a retirement trigger to each rather than a one-time sign-off.

SECTION FIVE

Data You *Didn't Collect.*

The hardest version of the provenance problem is the data the organisation never handled — datasets bought from a vendor and, increasingly, the vast training corpora behind the foundation models on which modern systems are built. It is tempting to treat this as someone else's responsibility. The framework does not allow it: the organisation is accountable for the data governance of every system it deploys, and it cannot discharge that accountability by pointing to whoever supplied the data. Accountability does not transfer with the dataset.

Vendor data is your data for governance purposes. Where a system is trained on data procured from a third party, the organisation must obtain the same provenance, consent basis, and quality assurances it would need for data it collected itself — and where the vendor cannot or will not provide them, the resulting gap is the organisation's gap. "The vendor assured us it was fine" is not provenance; it is a promise standing in for the documentation the framework requires.

Foundation-model data is often opaque by construction. A system built on a foundation model inherits a training corpus whose composition the provider typically does not fully disclose — an opacity the deployer cannot close by diligence alone, because the information is structurally unavailable (Hossain, 2026). This does not make the obligation disappear. It means the organisation must document what is known, assess the risk the unknown provenance poses for the specific use, and obtain accountable sign-off before deploying such a model in any domain affecting individuals — treating opacity as a live governance condition, not a solved one.

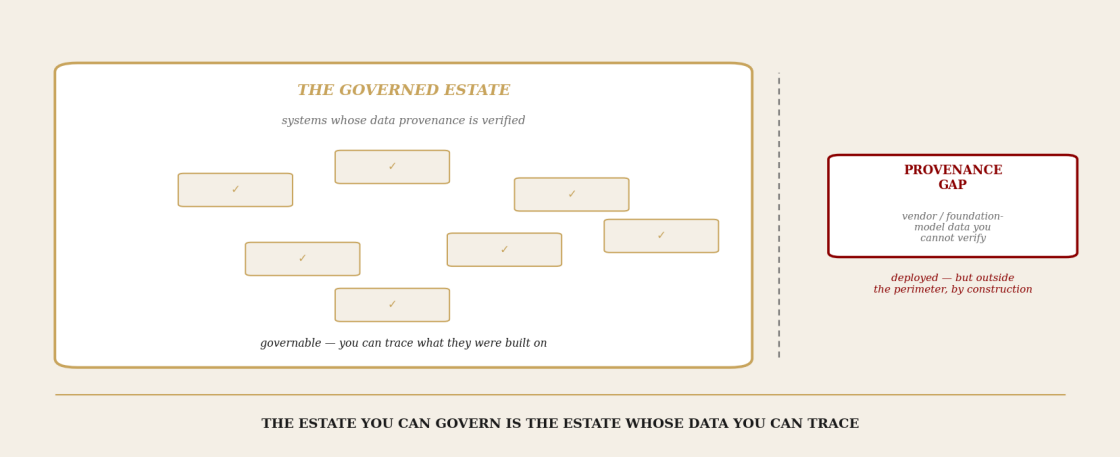

Figure 2 — The governance perimeter. The estate an organisation can govern is the estate whose data it can trace; a system with a provenance gap — vendor or foundation-model data it cannot verify — sits outside the perimeter, however it was deployed.

A provenance gap puts a system outside the perimeter. The set of systems an organisation can genuinely govern is exactly the set whose data it can account for. A system whose provenance cannot be established is, by construction, outside that perimeter — deployed, perhaps, and generating outputs, but not something the organisation can stand behind. The gap between the systems an organisation has deployed and the systems it can actually govern is measured precisely by its unverified data.

Closing the gap is a procurement decision made before purchase. Because the obligation cannot be waived, it becomes a condition of acquisition: a vendor system or foundation model must be able to support the organisation's provenance requirements before it is bought, not after. "We cannot verify its data" is not a constraint the organisation accepts after the fact; it is a reason not to acquire that system for decisions about people — the same discipline the framework applies to fairness and explanation, applied to the data beneath them.

SECTION SIX

The Precondition *Beneath the Rest.*

Every responsible AI obligation — the line that certain uses cannot cross, the fairness threshold someone must own, the explanation owed to the person affected, the right to contest a decision — shares a silent dependency. Each presupposes that the organisation knows the data its systems were built on. Data governance is not the last of the responsible-AI obligations. It is the one the others are quietly standing on.

Trace the dependency, and it is exact. A fairness assessment compares outcomes across groups in a population; if the organisation cannot characterise the population the training data represents, the assessment measures disparities in an unknown — a number without a referent. An explanation names the factors that drove a decision; if the organisation cannot vouch for what the model learned, it cannot vouch that the factors it reports are the factors that truly acted. A contestation reviews the basis of a decision; if that basis cannot be reconstructed from trustworthy inputs, there is nothing solid for the reviewer to examine. Remove verified data, and each of these controls still runs — and each is silently void.

This is why a provenance gap is so much more dangerous than it looks. It does not announce itself as a failure of fairness, explanation, or contestation; those controls keep operating, keep producing their artefacts, keep passing their gates. The failure is one level down, invisible to each of them, and it hollows all three at once. An organisation can hold a clean record on every downstream control and still be governing nothing, because the ground the whole structure rests on was never verified.

So, knowing the data is not one item on the responsible AI checklist — it is the precondition that makes the rest of the checklist meaningful. An organisation serious about fairness, explanation, and contestation has no choice but to be serious first about provenance, because the three cannot be delivered on data it cannot account for. “Unknown data is ungoverned data” is not a slogan about one principle. It is a statement about the foundation on which the other principles are built — and about what happens to all of them when that foundation is missing.

SECTION SEVEN

Held by *Architecture.*

A precondition that depends on people remembering to check it is not a precondition; it is a hope. The framework, therefore, does not leave data provenance to diligence. It wires the requirement into the same machinery that governs the rest of the AI estate, so that no system can reach the decisions it makes about people without its data having been established, documented, and found fit — and so that a lapse in that foundation suspends the system rather than passing unnoticed.

The line runs from constitution to evidence. Data integrity and provenance are adopted at board level as a governing principle; the policy and standard specify exactly what must be known and to what quality; the deployment gate refuses any system whose provenance documentation is incomplete; and the register that tracks each system carries its provenance-certified status and the dates its assessments were made. Establishing provenance and maintaining it as consent and currency change are recorded acts with owners — not assurances produced when a regulator asks.

Crucially, provenance attaches to the existing enforcement spine rather than sitting beside it as documentation nobody reads. The gate that clears a system for deployment is extended to require complete, current provenance for every material dataset. The register that records a system's authority carries whether its data is certified and resets that status the moment a retirement trigger fires. The people who supervise systems in operation are briefed to treat a lapsed consent basis or a stale dataset as a fault that suspends the system, not a housekeeping item. The requirement becomes a precondition that the system must satisfy and maintain, not a folder of forms.

The effect is that provenance becomes an artefact rather than an assurance. When a regulator asks where a system's training data came from and on what basis, the answer is the documented lineage and the consent-basis determination — not a search of people's memories. When asked whether the data is still fit, the answer is the current quality and distribution-shift assessments, as well as the subsequent retirement decisions. A provenance that can be shown, dataset by dataset, is one that was actually established; one that can only be asserted was the gap all along, waiting to be found.

None of this makes data governance glamorous — it is the least visible of the responsible-AI obligations, and the easiest to defer because nothing appears to break when it is skipped. What the architecture provides is a refusal to build on unverified ground: a requirement, enforced at the gate and maintained in the register, that the organisation know the data before it trusts the

system, and keep knowing it for as long as the system runs. That is what it means to govern from the foundation up, rather than to decorate a structure whose base was never checked.

SECTION EIGHT

Five Questions *for Your Board.*

The measure of a data-governance regime is not the sophistication of its quality tooling. It is whether the organisation can account for every system that decides about people, for the data on which the system was built. Each of the following is a question a regulator, a court, or an affected individual could put to the board — and each should be answerable today, with an artefact rather than an assurance.

For each system, can we say where its data came from? For every AI system that makes or materially contributes to decisions about individuals, can you produce the documented provenance of its material training datasets — origin, custody chain, and transformations — or are there datasets whose source no one can now establish?

Do we have a consent basis for the use, not just the collection? For training data containing personal information, can you show a consent basis that specifically covers AI training in the relevant domain — or are you relying on generic collection consent that was never given for this purpose?

Is a provenance gap a barrier, or a risk someone can accept? When a dataset's provenance cannot be established, does the system stop until the gap is closed — or can a manager, under a deadline, rate and accept it? Can you show a case where a system did not ship because its data could not be traced?

Does our accountability extend to the data we didn't collect? For vendor data and foundation models, do you obtain and document the same provenance assurances you would for your own data — or does "the data is the vendor's" stand in for the governance you are still accountable for?

Do we know the foundation on which the rest of our governance rests? Can you show that your fairness, explanation, and contestation controls are operating on data whose provenance is verified — or might they be running, and passing, on systems whose data was never established? A framework that cannot answer these questions is not governing its AI. It is operating it, and hoping the ground holds.

ABOUT

42 Consulting.AI.

42 Consulting.AI builds constitutional AI governance for regulated enterprises. Its core work, the MANDATE Suite, is a comprehensive governance architecture comprising 64 normative instruments across five frameworks — governing the delegation of machine decision authority, the ethical and rights boundaries on its use, the management system that maintains it, and the regulatory extensions for EU and group structures. The Suite is designed for organisations in financial services, insurance, superannuation, healthcare, government, and other regulated sectors operating autonomous AI systems at scale.

The architecture is supported by an academic preprint series published on Zenodo, which establishes the theoretical basis for each governance mechanism, and by the forthcoming book *Governance-First AI*. The Suite's distinguishing conviction is that the AI governance that regulated enterprises require cannot be delivered by a compliance framework alone — it requires a constitutional architecture, adopted by the board, enforced by the infrastructure, and continuously tested against current conditions.

Dr M Maruf Hossain, PhD, GAICD · Chief AI Strategist
enquire@42consulting.ai · www.42consulting.ai

The preprint series is available at: zenodo.org/communities/mandate/records

REFERENCES

References *and Sources.*

The regulatory instruments referenced in this paper were verified at the primary source. The governance mechanisms are set out in full in the MANDATE preprint series, published through Zenodo, and in the forthcoming book Governance-First AI. The full MANDATE preprint series is available at: zenodo.org/communities/mandate/records

REGULATORY AND INSTITUTIONAL SOURCES

European Union (2024). Regulation (EU) 2024/1689 (Artificial Intelligence Act), Article 10 — data and data-governance obligations for high-risk AI, including relevance, representativeness, and examination of training data for biases.

International Organisation for Standardization (2023). ISO/IEC 42001:2023, Annex A.7 — data for AI systems: data governance, quality, provenance, and preparation requirements.

Office of the Australian Information Commissioner. Privacy Act 1988 (Cth), Australian Privacy Principles 1, 3, 6 and 11 — transparent management, collection, use, and security of personal information used to train and operate AI systems.

Australian Prudential Regulation Authority. CPS 234 (information security) and CPS 230 (operational risk management) — data governance and data quality as prudential obligations for regulated entities.

United Nations (2011). Guiding Principles on Business and Human Rights — human rights due diligence, applicable to harms arising from the data on which automated decisions are built.

BOOK

Hossain, M Maruf and Ponnusamy, Ahilan (2026). Governance-First AI. Apress (forthcoming) — the primary public articulation of the responsible AI governance doctrine.

MANDATE PREPRINT SERIES

Hossain, M Maruf (2026). Foundation Model Dependency Governance: A Framework for Governing Enterprise AI Systems Whose Behaviour Is Materially Determined by Foundation Models. Preprint. Zenodo. https://doi.org/10.5281/zenodo.20680495.

Consulting.AI

Scaling Intelligence

enquire@42consulting.ai · www.42consulting.ai