Architecture for Interoperable Agricultural Evidence

First-Time Reader Brief, Seven-Paper Scientific Series, and Reference Appendix

The Australia-First Thin-Waist Strategy

AgEvidence, its go-to-market architecture, scientific program, and interoperability thesis

Consolidated scientific working draft and IC briefing edition. Named manuscript authors have been removed. Scientific authorship is represented only by role-based personas pending written consent, qualifying contribution, affiliation, conflict disclosure, CRediT allocation, and final-manuscript approval. Bibliographic author names inside scholarly citations and reference lists are retained because they identify the underlying literature, not proposed authors of this series.

Version: Working Draft 0.7 (format-hardened first-time-reader IC edition)
Date: August 2026
Status: Internal scientific publication and go-to-market strategy document
Revision: format-hardened reader flow; inline abstracts; compact portfolio; literature-review appendix retained at end

First-Time Reader Brief: What AgEvidence Is and Why This Scientific Series Exists

Reader orientation. This section is written for a reader encountering AgEvidence for the first time. AgEvidence should not be understood first as a carbon calculator, a registry, a new agricultural methodology, a verifier, or a generic farm-data platform. It is an evidence-infrastructure company built around a narrower proposition: agricultural evidence should be captured with stable provenance once, remain inspectable as methods and institutions change, and be reusable for multiple bounded decisions without silently changing what the historical record says.

The problem in plain language

Agricultural climate and environmental decisions are increasingly assembled from heterogeneous evidence. A single downstream claim or assurance process may depend on device telemetry, laboratory assays, intervention records, animal or cohort information, model runs, spatial observations, producer records, methodology versions, reviewer judgments, verification outcomes, rights declarations, and buyer-specific requirements. Those inputs are usually created by different systems for different purposes.

The recurring structural problem is not simply that data are missing. It is that evidence and interpretation are often fused too early. A record may be created for one buyer questionnaire, one project method, one calculator, one verification engagement, or one jurisdiction. When the methodology changes, the buyer changes, a new verifier asks a different question, a correction is made, or a company needs to support a second use of the same underlying activity, the evidence package is frequently reconstructed rather than reinterpreted. The same facts are recopied, transformed, renamed, or re-keyed into another workflow.

AgEvidence starts from a different design assumption:

Capture the historical evidence once → preserve provenance and limitations → apply versioned interpretations later → generate bounded artifacts for specific relying parties.

This is the basis of the "thin-waist" architecture developed throughout the scientific series. The shared waist is intentionally small. It preserves source identity, time, actors, relationships, commitments, parentage, authority, rights, limitations, method/model identity where relevant, and the distinction between observation, assertion, calculation, review, determination, and reliance. Country rules, scientific methods, eligibility tests, institutional policies, reporting formats, and buyer requirements remain above that common layer.

What AgEvidence is building

AgEvidence is best understood as two deliberately connected product layers around one evidence grammar.

Layer	What it does	Why it exists
Open/free developer layer	Gives agricultural technology teams common evidence primitives, manifests, schemas, fixtures, local verification, and a disciplined way to emit source-linked evidence from their existing products.	Lower integration friction. The first useful action should be technical adoption, not a sales process or a migration into an AgEvidence-owned system of record.

Layer	What it does	Why it exists
Commercial Evidence Plane	Continuously evaluates evidence against maintained profiles; surfaces gaps; supports human review and determinations; generates statements and reliance artifacts; manages sharing, API access, integrity, program versions, and operational support.	Monetise the operational work around evidence reuse without making AgEvidence the scientific, regulatory, legal, or verification authority.

The distinction matters. The open layer is intended to make the grammar easy to adopt. The paid layer becomes useful when evidence has to survive organizational boundaries, changing requirements, repeated evaluation, review, statement generation, external sharing, program maintenance, or institutional reliance. In that model, AgEvidence does not need to own the upstream farm, sensor, laboratory, model, or intervention application. It needs those systems to be able to emit evidence into a stable, inspectable contract.

The operating chain from source system to reliance

A simplified deployment path is:

Agricultural technology or source system

↓

AgEvidence SDK / evidence contract — source records, observations, interventions, model runs, operational events

↓

Append-only evidence graph — provenance, relationships, commitments, limitations, authority, rights

↓

Versioned evidence profiles / adapters — scientific method, program, buyer, assurance, jurisdiction

↓

Evidence Plane — evaluation, gaps, review, determination, statement generation, sharing

↓

Bounded reliance artifact — what can be supported, under which version, with which limitations, for which purpose

The source system remains the source system. AgEvidence is not supposed to become a shadow farm-management, animal-management, laboratory, sensor, registry, or buyer system merely because it can receive evidence from those systems. The architecture is valuable only if it preserves that boundary.

Why Australia is first

The scientific series uses Australia as the first reference context, not because Australian rules are meant to become universal, but because the architecture needs a concrete place in which to be falsified. Australian livestock and agricultural systems force the design to confront meaningful heterogeneity: pasture-based and intensive production, enteric-methane interventions, spatial and remote-sensing evidence, measurement and modelling, producer workflow, co-claim and rights questions, assurance pathways, and downstream value-chain use.

Australia therefore functions as the first pressure vessel for a country-neutral architecture. The success condition is not that one Australian rule becomes the global rule. It is that the common evidence layer remains stable while Australian scientific and institutional authority stays explicit and local.

The commercial entry strategy: developers first, institutions later

The initial adoption motion is intentionally different from selling a large enterprise system directly to a processor, retailer, bank, registry, or government program. Funded agricultural technology companies already produce the data that downstream institutions need. Their immediate problem is making that data legible, supportable, reusable, and defensible across counterparties.

The developer-first path is therefore:

1. integrate a narrow evidence primitive into an existing product;
2. emit a real source-linked event or record without replacing the upstream workflow;
3. verify the resulting object and preserve its provenance;
4. map it against a bounded evidence profile and expose missing or ambiguous evidence;
5. generate an artifact another internal or external actor can inspect;
6. expand to the paid Evidence Plane when the evidence needs continuous evaluation, review, sharing, maintenance, or reliance support.

This is why the scientific program cannot be separated from the go-to-market program. The product needs real source diversity before the common evidence waist can be considered stable. Conversely, developer adoption becomes easier if the primitives are supported by a coherent scientific explanation of why those fields, boundaries, and validity states exist.

How the Scientific Program Supports the Go-to-Market Strategy

The seven papers are an operating sequence, not a parallel academic project

The seven manuscripts are designed as a progressive falsification program for the AgEvidence architecture. Each paper removes one class of objection that would otherwise prevent the company from scaling a shared evidence layer.

Paper	Question it must answer	GTM dependency	What failure would mean
1	Can a sufficiently small, country-neutral evidence contract represent materially different agricultural sources?	Early Wave 1 integrations provide heterogeneous source contracts and reconstruction cases.	The shared waist is too narrow, too abstract, or already contaminated by application-specific assumptions.
2	Can the system distinguish integrity, method compatibility, scientific judgment, verification, and institutional reliance without laundering one into another?	Real interfaces and artifacts reveal where a single "verified" status would create misleading product behavior.	The product cannot communicate trust states safely enough for high-assurance use.
3	What evidence is actually necessary for bounded enteric-methane decisions?	Late Wave 1 source reality constrains the shared schema and prevents speculative field accumulation.	The architecture either omits scientifically decisive context or burdens developers with unnecessary fields.
4	Can the same primitives survive biological and source-record-first deployment from trial evidence through farm-gate operations?	Wave 2 is the decisive pressure test beyond clean event telemetry.	The initial developer wedge does not generalise to messy operational biology.
5	Can farm, spatial, grazing, remote-sensing, and producer evidence compose without creating a second core architecture?	Whole-farm primitives harvested during earlier integrations are composed after the first four papers stabilise.	The architecture fragments when evidence becomes spatially and temporally compositional.
6	Can the same evidence grammar extend to manure, waste, digestion, energy balances, and continuous engineering systems?	Post-Wave-2 domain extension deliberately tests a different physical system.	The supposed thin waist is actually livestock-intervention-specific.

Paper	Question it must answer	GTM dependency	What failure would mean
7	Can one frozen evidence history be reinterpreted under multiple jurisdictions without rewriting the past or pretending policy is universal?	Only tested after the Australian root is stable.	International expansion requires country-specific cores rather than adapters and profiles.

The sequencing is therefore commercial as well as scientific:

discover the waist → separate trust states → define sufficiency → prove biological deployment → prove composition → prove domain extension → prove jurisdictional portability

Wave 1 and Wave 2 are architecture-discovery waves

Wave 1 and Wave 2 should not be read as arbitrary customer cohorts or as a promise that a particular number of companies equals product-market fit. Their purpose is to create enough heterogeneity to learn what belongs in the shared evidence contract and what belongs at the edge.

Wave 1 emphasizes event-first sources: systems that already generate operational or measurement events that can be expressed cleanly. Wave 2 deliberately introduces biological, assay-heavy, model-heavy, and source-record-first evidence where identity, dose, persistence, cohorts, calibration domain, exceptions, and rights are harder to infer from telemetry alone.

The freeze doctrine is central. After sufficient source diversity, new companies should map into existing primitives, manifests, evidence profiles, connectors, and adapters. Repeated requests for company-specific core objects would not be a sign of healthy customization; they would be evidence against the scalability thesis.

How GTM work becomes scientific material

Every serious integration should produce more than integration code. It should generate reusable research and product artifacts: source manifests, event contracts, canonical fixtures, gap cases, reconstruction tests, ambiguity cases, reviewer questions, method mappings, and reliance artifacts. Those materials provide the empirical substrate for Papers 1–4.

The direction also runs the other way. Scientific writing forces the company to state exactly what a primitive means, which inferences are prohibited, what evidence is sufficient for a bounded decision, where model-domain limits belong, and what the system must not imply. That discipline becomes SDK documentation, conformance tests, interface rules, product copy, and sales guardrails.

The GTM-science flywheel

Free SDK adoption → real source diversity → shared primitives + failure cases → scientific manuscripts + external review → clearer category vocabulary + institutional legibility → paid Evidence Plane demand → more source diversity

This loop works only if both halves remain operationally coupled. Publications that do not incorporate what integrations reveal become a detached academic program. Integrations that do not improve the shared architecture become consulting. The strategy requires each to discipline the other.

How the Scientific Program Contributes to Moat Building

Moat interpretation. The papers do not create a moat merely by being published, and peer review does not confer product endorsement, regulatory approval, or commercial adoption. The moat-building thesis is cumulative: open technical adoption, a stable evidence grammar, executable conformance, source diversity, scientific vocabulary, external participation, maintained profiles, and repeated institutional reliance can reinforce one another over time.

1. Standard gravity around the thin waist

The free SDK lowers the cost of emitting evidence in the AgEvidence grammar. If multiple independent agricultural technology products adopt the same small set of primitives because doing so is easier than repeatedly building bespoke evidence exports, the shared contract can acquire standard-like gravity before any formal standards body is involved. Papers 1 and 2 are especially important here: they give the grammar a citable technical rationale and a bounded trust vocabulary rather than leaving it as unexplained vendor schema design.

The defensibility mechanism is not secrecy. It is coordination. Once developers, reviewers, buyers, and assurance teams can discuss the same evidence objects and trust states, replacing the grammar requires more than copying a schema; it requires recreating the accumulated coordination around that schema.

2. A growing corpus of conformance cases and edge mappings

Each integration should add manifests, fixtures, source mappings, gap cases, exception patterns, and conformance tests. These are reusable assets. A new customer is cheaper to integrate when its evidence can be mapped against prior patterns rather than modeled from first principles. The moat grows if source diversity increases while the core remains stable.

This is also a falsifiable moat. If every new company requires a new core object, there is no compounding architecture advantage. There is only bespoke integration work.

3. Scientific vocabulary as category infrastructure

The papers introduce and sharpen concepts such as thin-waist agricultural evidence, validity separation, decision-linked sufficiency, bounded reliance, evidence continuity, independent reconstruction, and jurisdictional portability. If those concepts become useful to researchers, assurance practitioners, developers, buyers, and institutions independent of the AgEvidence product, they can define the category in a way that naturally favors a neutral evidence layer.

This is category-building rather than advertising. The scientific contribution must remain useful even to a reader who never buys AgEvidence. That independence is precisely what can make the vocabulary durable.

4. External scientific participation as distributed scrutiny and distribution

Prospective coauthors, reviewers, laboratories, universities, and domain specialists are not sales endorsers. Their value is different: they expose weak claims, force domain-specific precision, expand the set of communities that understand the architecture, and create independent scholarly routes through which the concepts can be encountered.

A successful publication program therefore expands the architecture's intellectual surface area while preserving the boundary between scientific participation and commercial endorsement.

5. Maintained profiles as the commercial layer around an open grammar

An open evidence grammar does not eliminate the need for maintained interpretation. Scientific methods, buyer requirements, program rules, reporting expectations, evidence sufficiency thresholds, and institutional policies change. The paid Evidence Plane can maintain versioned profiles, continuously reevaluate evidence, expose gaps, manage review, generate updated statements, and preserve prior determinations.

That creates a commercial asymmetry: the base evidence remains portable, while the operational burden of keeping it decision-ready becomes a recurring service. The company can therefore be open at the interoperability layer without giving away the maintained operational system around it.

6. Historical continuity and switching cost without data lock-in

Traditional software switching cost often comes from trapping data inside an application. The AgEvidence thesis is stronger if the evidence itself remains portable. The switching cost should come from accumulated continuity: years of versioned source records, corrections, parent relationships, determinations, limitations, review events, and reliance history that are already coherent under a shared contract.

A competing system can ingest the evidence, but recreating the meaning, conformance history, maintained mappings, and institutional context becomes increasingly expensive. That is a healthier form of defensibility because it is based on accumulated structure and use rather than export friction.

7. Institutional legibility without claiming institutional authority

Buyers, processors, assurance teams, banks, programs, registries, and regulators do not need to delegate authority to AgEvidence. The architecture is designed to make their own authority easier to express and inspect. A buyer can rely on an artifact for one purpose while another institution rejects or conditions the same underlying evidence for another purpose. Those outcomes can coexist without rewriting the history.

If the system becomes a useful neutral handoff layer between organizations with different mandates, its value rises with the number of relying contexts it can support. This is the beginning of a coordination/network effect around evidence reuse rather than a simple user-count network effect.

8. Expansion by adapters instead of product forks

Papers 5–7 test whether new complexity can be absorbed at the edge: whole-farm composition, engineering and LCA evidence, and jurisdiction-specific rules. If that works, expansion adds profiles, adapters, conformance cases, and maintained domain knowledge to the same system. The company can enter new use cases without recreating its trust architecture.

The moat claim fails if each new domain becomes a separate product with incompatible evidence semantics. The seven-paper sequence is designed to discover that failure early.

The compounding moat model

Developer adoption
↓
More heterogeneous source mappings and fixtures
↓
More stable shared grammar + stronger conformance suite
↓
Citable scientific vocabulary + external scrutiny
↓
Greater institutional legibility and more maintained profiles
↓
More reliance events + recurring Evidence Plane demand
↻

The strategic objective is not to maximize any one layer in isolation. A large number of SDK installs with no reliance is weak. A collection of papers with no developer adoption is weak. A commercial console that requires bespoke schemas for every customer is weak. The defensible outcome is convergence: one small grammar repeatedly reused across different sources, methods, institutions, and jurisdictions, with the operational layer becoming more valuable as that reuse grows.

What the Scientific Series Does - and Does Not - Prove

What successful execution would support

If executed as designed, the research program can support evidence that:

1. a stable agricultural evidence contract can represent materially different source systems without repeated core redesign;
2. trust states can be separated so cryptographic integrity is not mistaken for scientific validity, accredited verification, or institutional acceptance;
3. evidence sufficiency can be tied to bounded decisions rather than inflated into a universal "minimum dataset";
4. the architecture can survive biological deployment, whole-farm composition, engineering domains, and jurisdictional reinterpretation;
5. independent reviewers can reconstruct determinations from disclosed evidence, versions, and rules; and
6. the commercial product can remain operationally useful without claiming to be the scientific or institutional authority.

What it would not establish

Even successful peer review would not by itself establish efficacy of a specific intervention, approval of an emissions-reduction methodology, accredited validation or verification, legal entitlement to environmental attributes, government acceptance, registry eligibility, accounting treatment, buyer acceptance, product-market fit, or investment return. Those decisions remain with the relevant scientific, legal, assurance, governmental, accounting, commercial, or investment authorities.

That boundary is not a weakness in the product thesis. It is a design requirement. AgEvidence is intended to preserve evidence and make authority legible, not absorb every authority into software.

What an investment committee should watch

Question	Evidence of progress
Is the SDK creating real adoption?	Independent teams can emit and verify useful evidence quickly from existing products without rewriting their application around AgEvidence.
Is the thin waist stabilising?	New source systems map to existing primitives and profiles; core-schema changes become less frequent as diversity increases.
Are the papers grounded in operations?	Manuscripts use integration-derived fixtures, gaps, reconstruction cases, and reviewer problems rather than a parallel synthetic research universe.
Does science improve the product?	Manuscript work produces clearer schemas, conformance tests, UI trust states, evidence sufficiency rules, and non-assertion boundaries.
Does Paper 4 survive the hard transition?	Biological/source-record-first evidence reuses the same primitives used for cleaner event-first sources.
Is the architecture expanding rather than forking?	Papers 5–7 add profiles, adapters, composition rules, and conformance cases rather than new incompatible cores.
Is the open/paid boundary working?	Free adoption lowers friction while customers pay for continuous evaluation, review, statements, sharing, APIs, program maintenance, integrity, and operational support.
Is reliance appearing?	Artifacts are inspected or used by actors beyond the originating developer team, creating concrete demand for maintained evidence operations.

Scientific Author Persona Convention

This consolidated working draft uses role-based scientific personas in place of proposed named manuscript authors. The personas are recruitment and contribution placeholders only; they do not imply consent, affiliation, authorship, endorsement, or approval.

Persona	Scientific / operating domain
Persona A	Australian livestock greenhouse-gas systems, production context, methodology and local-authority boundaries.
Persona B	Enteric methane, ruminant nutrition, intervention efficacy, modelling and calibration-domain science.
Persona C	Agricultural MRV governance, Scope 3 translation, institutional use and bounded reliance.
Persona D	Methane measurement, experimental design, uncertainty and evidence sufficiency.
Persona E	Manure, anaerobic digestion, bioenergy, mass/energy balance and LCA engineering.
Persona F	Grazing systems, geospatial evidence, producer workflow and operational burden.
Persona G	Scientific data provenance, interoperability, conformance, schemas, verification tooling and reproducibility.
Persona H	Research governance, CRediT, disclosures, institutional use and scientific communications.
Persona I	Local method, standards, verification, rights and data-governance review.

Authorship rule. A real person becomes an author only after written consent, a qualifying contribution, affiliation confirmation, conflict disclosure, CRediT allocation, and final-manuscript approval. Bibliographic author names remain intact because they identify the underlying scholarly literature.

Seven-Paper Scientific Portfolio: Titles and Abstracts

How to use this section. The following are the manuscript titles and full working-draft abstracts reproduced before the papers so a first-time reader can understand the entire research sequence before entering the detailed manuscripts. The abstracts retain the scientific non-assertion boundaries of the papers themselves.

Paper 1: A Thin-Waist Architecture for Interoperable Agricultural Evidence

Role in the sequence: Constitutional architecture. Defines the country-neutral thin waist and the test for whether one frozen evidence root can support multiple later interpretations.

Abstract

Agricultural greenhouse-gas measurement, reporting, and verification systems are commonly implemented as country-specific methods, proprietary data pipelines, registry workflows, calculator exports, or project-level document collections. These arrangements can be effective for an initial submission, but they couple historical evidence to one interpretation and make later reconstruction expensive when methods, models, buyers, verifiers, claim rules, or jurisdictions change. This paper proposes an Australia-first reference architecture in which agricultural evidence is captured once as a signed, append-only graph and interpreted many times through immutable, versioned country, method, and institutional policy layers.

The architecture is organised around a deliberately narrow “thin waist”: a country-neutral receipt envelope; stable event and decision families; source commitments; immutable parent links; identity, role, authority, rights, and consent declarations; explicit limitations and non-assertions; cryptographic signatures; a bounded verification vocabulary; and a common bundle manifest. National methods, eligibility rules, model profiles, uncertainty treatments, claim policies, data requirements, verification expectations, and reporting artifacts remain above this layer as declarative adapters. The design therefore separates what happened and who asserted it from how a later decision-maker interprets it.

Australia is proposed as the first implementation context because the program must accommodate pasture-based and intensive livestock systems, local scientific participation, value-chain actors, co-funding and co-claiming questions, and a standards-alignment pathway without embedding Australian policy into the generic trust kernel. The paper specifies dependency direction, core semantics, receipt families, graph invariants, adapter manifests, reliance artifacts, conformance tests, and an independent reconstruction protocol. The central evaluation asks whether one frozen evidence root can support multiple method determinations and institution-specific reliance artifacts without rewriting historical evidence or changing its cryptographic commitments.

The contribution is not a new emissions-reduction methodology and does not establish scientific efficacy, government approval, verification, legal entitlement, or claim ownership. It is a testable interoperability contract intended to make agricultural evidence portable while preserving local authority over applicability, scientific judgment, verification, rights, and institutional acceptance.

Keywords: agricultural MRV; evidence architecture; interoperability; thin waist; digital provenance; livestock decarbonisation; policy adapters; evidence graph; Australia.

Paper 2: Beyond "Verified"

Role in the sequence: Trust semantics. Separates integrity, method compatibility, scientific judgment, accredited verification, and institutional reliance.

Abstract

Digital agricultural measurement, reporting, and verification systems routinely use the word "verified" to describe materially different propositions: that a digital record is intact; that an activity appears compatible with a scientific method; that a scientific reviewer regards an interpretation as credible; that an accredited validation or verification body has reached a determination; or that a buyer, registry, lender, processor, government-facing program, or other institution has relied on an artifact. Collapsing these meanings creates validity laundering, in which confidence generated at one layer is transferred rhetorically or visually to another layer without the required evidence, authority, or decision process.

This paper proposes a three-layer validity model for agricultural evidence systems. Cryptographic integrity is limited to whether canonical commitments, signatures, parent relationships, trust material, and bundle consistency can be evaluated as valid, invalid, or indeterminate. Method compatibility is produced only by a versioned policy evaluator and records whether declared evidence is eligible, eligible with conditions, outside the current method, requires method extension, lacks sufficient evidence, or remains unassigned. Institutional reliance records whether a named institution accepted, relied on, rejected, or requested more evidence from a specific artifact for a bounded purpose. Scientific judgment and accredited verification remain explicit events linked to these layers rather than being silently absorbed into them.

The paper formalises the three state machines, defines prohibited inferences and transition rules, specifies API and interface requirements, and proposes a mixed-method evaluation using concept analysis, structured expert elicitation, inter-rater classification, interface and report audit, and conformance cases. Australia-first livestock scenarios are used to test whether systems preserve ambiguity, missing evidence, model-domain limits, and institution-specific decisions. The contribution is a governance vocabulary and conformance framework for transparent decision support. It does not replace scientific peer review, accredited validation or verification, legal analysis, or Australian government and standards authority.

Keywords: validity laundering; verification; agricultural MRV; cryptographic integrity; method compatibility; institutional reliance; governance; conformance; interface design.

Paper 3: Minimum Evidence Requirements for Enteric Methane Interventions in Australian Beef and Dairy Systems

Role in the sequence: Evidence sufficiency. Defines decision-linked minimum evidence for enteric-methane interventions without confusing sufficiency with efficacy.

Abstract

Rapid development of enteric-methane interventions has created a recurring translation problem: evidence sufficient to demonstrate efficacy in a controlled study may not be sufficient to support repeatable implementation, quantification, verification, or value-chain claims. This paper proposes a minimum-evidence and conformance framework for enteric-methane interventions in Australian beef and dairy systems. The study combines a preregistered systematic or scoping review, structured expert consensus, evidence-to-decision mapping, production-context testing, and machine-readable conformance cases.

The framework covers intervention identity and product characterisation; animal and cohort context; production system and diet; baseline definition; prescribed and delivered dose; observed exposure; duration and interruptions; methane measurement; model execution; animal performance and health; adverse effects; persistence; operational delivery; authority; rights; limitations; and intended decision use. Evidence is classified by source, temporal relationship, authority, transformation status, uncertainty, and decision purpose. The framework separates observations from calculations, assumptions, scientific judgments, method-compatibility assessments, accredited validation or verification determinations, and institution-specific reliance.

Measurement requirements are grounded in the distinct error structures and operating domains of respiration chambers, sulfur hexafluoride tracer methods, automated spot-sampling systems, and model-based estimates. Intervention-specific requirements preserve formulation, active component, batch, stability, diet interaction, dose-response, adaptation, and persistence. Australian pasture-based beef, feedlot beef, and dairy systems are treated separately where delivery fidelity, animal tracking, intake measurement, recording burden, and calibration-domain applicability differ materially.

The planned outputs are a minimum dataset, controlled vocabulary, gap taxonomy, decision-linked requirement matrix, production-context profiles, conformance checklist, and machine-readable schema. These outputs are intended to support transparent evidence readiness and identify when additional evidence, conditions, independent review, or method extension is required. They do not establish official Australian requirements, rank commercial interventions, or certify emissions reductions.

Paper 4: From Trial to Farm Gate

Role in the sequence: Applied deployment. Tests whether trial evidence can remain continuous through real farm-gate implementation, models, exceptions, rights, and claims.

Abstract

Controlled feed-additive trials can produce credible evidence of enteric methane reduction, but commercial deployment introduces additional evidentiary dependencies distributed across manufacturers, logistics providers, producers, feed systems, animal cohorts, model operators, verifiers, registries, funders, and downstream claimants. Product batch variation, ration-inclusion error, missed feeding intervals, cohort turnover, changing diets, fragmented data custody, and retrospective spreadsheet assembly can weaken the link between trial efficacy and a reconstructable value-chain outcome.

This paper proposes an append-only directed acyclic evidence graph for feed-additive deployment in Australian beef and dairy systems. The graph implements the thin-waist receipt contract of Paper 1, the three-layer validity model of Paper 2, and the minimum-evidence vocabulary of Paper 3. It binds scientific trial evidence to intervention identity, batch and purchase records, delivery and storage, administration and ration events, animal-cohort membership, baseline performance, methane measurement or model execution, scientific review, verification-readiness artifacts, outcome recognition, funding contributions, rights evidence, co-claim allocation, producer payment, and institution-specific reliance. Every node declares issuer, authority, event time, source commitments, parent relationships, limitations, and supersession status.

The architecture is evaluated through a synthetic mirror, permissioned pilot subset, or approved post-pilot case and an independent reconstruction protocol. Reviewers receive a bounded bundle rather than access to a hosted workflow and attempt to reconstruct implementation, quantitative calculations, exception handling, and claim allocation. The scenario suite includes product substitution, incomplete feeding logs, missed dose, cohort change, animal movement, model-version update, missing baseline, conflicting claim, pending verification scope, and corrected source records.

A central contribution is mathematical separation of the recognised physical greenhouse-gas outcome Q from commercial or institutional claim allocations a_i. Under a declared policy profile, the graph enforces the aggregate-cap invariant $\sum_i a_i \leq Q$. This prevents recorded quantitative allocations from silently exceeding the recognised outcome while leaving entitlement, claim language, and legal effect to relevant authorities. The study evaluates traceability, reconstruction effort, quantitative discrepancy, graph completeness, exception visibility, and aggregate-cap behaviour rather than product efficacy or formal method, registry, VVB, or government approval.

Keywords: feed additive; enteric methane; evidence graph; chain of custody; Scope 3 accounting; co-claims; livestock supply chain; verification readiness; data provenance.

Paper 5: Whole-Farm Grazing Systems

Role in the sequence: Whole-farm composition. Tests spatial, temporal, grazing, producer, and remote-sensing composition without creating a second architecture.

Abstract

Extensive grazing systems present distinct verification and data-governance challenges compared with contained agricultural operations. Ruminant herds move dynamically across variable paddock topographies, exposure to methane-mitigating dietary interventions may be intermittent, environmental conditions fluctuate, and direct emissions measurement is sparse. Consequently, greenhouse-gas monitoring, reporting, and verification systems must rely on heterogeneous data streams spanning producer-generated operational logs, remote-sensing products, spatial property records, weather observations, livestock records, and empirical or process-based emissions models.

This paper formalises a whole-farm evidence model tailored to Australian grazing beef systems. The framework structures producer and organisational identity, land parcels and paddocks, herd and cohort movement, management and supplementation events, environmental proxy observations, model executions, scientific reviews, rights, consent, contractual allocations, and institutional decisions into a temporal

graph schema. Rights, consent, data-use authority, land control, herd authority, and claim permissions are encoded as primary evidence primitives rather than unstructured metadata.

The paper establishes explicit temporal-alignment rules for combining asynchronous and multi-frequency evidence without silent interpolation. Each transformation declares source interval, target interval, subject mapping, spatial support, aggregation rule, imputation or interpolation method, calibration domain, uncertainty, and conditional assumptions. Remote-sensing observations remain distinct from ground observations, model estimates, and producer assertions. A successful model execution is separated from scientific applicability and from institutional acceptance.

The proposed evaluation combines participatory workflow mapping, source-system inventory, graph construction, temporal-alignment tests, scenario testing, independent reconstruction, and producer burden assessment. Burden is treated as a direct performance metric through setup time, recurring manual effort, duplicate entry, missing-data frequency, reconstruction latency, clarification requests, and artifact utility for producers, scientists, verifiers, buyers, and financial entities.

The intended outputs are a grazing evidence ontology, whole-farm graph profile, temporal-alignment protocol, rights and consent profile, minimum producer dataset, burden measurement framework, uncertainty disclosure model, and conformance cases. The framework is designed to distinguish directly observed operational facts from proxy, calculated, modelled, assumed, scientifically judged, verified, and institutionally relied-on propositions. It does not assert that one profile is sufficient for every Australian region, production system, method, or decision.

Paper 6: Manure, Waste, and Farm Energy

Role in the sequence: Engineering-domain extension. Extends the grammar to manure, waste, digestion, energy, mass balance, and LCA-oriented evidence.

Abstract

Livestock decarbonisation extends beyond enteric methane to manure storage, waste treatment, anaerobic digestion, biogas upgrading, combined heat and power, renewable electricity, recovered nutrients, bedding, digestate, thermal energy, and other farm-scale outputs. These interventions combine biological conversion, engineering operation, environmental accounting, and contractual attribute allocation. Evidence is commonly distributed across laboratory assays, weighbridges, flow meters, gas analysers, engine controllers, invoices, maintenance systems, grid meters, environmental reports, and policy calculators. Retrospective aggregation into one project spreadsheet can obscure whether a result changed because the physical system changed, a sensor drifted, a boundary moved, a default factor was updated, or an environmental attribute was reallocated.

This paper extends the Paper 1 thin-waist evidence architecture to manure, waste, and farm-energy systems while preserving the Paper 2 separation of cryptographic integrity, method compatibility, scientific or engineering judgment, accredited verification, and institutional reliance. It reuses the Paper 3 decision-linked evidence-sufficiency vocabulary and the Paper 4 separation of physical outcomes, contributions, rights, allocations, payments, and reliance.

The proposed architecture represents material inputs, storage, pre-treatment, digester configuration, operational state, laboratory and sensor provenance, gas production and composition, leakage, flare and engine operation, parasitic loads, electricity and heat output, digestate disposition, avoided counterfactuals, mass and energy balances, life-cycle calculations, co-product allocation, environmental attributes, and institution-specific reliance as an append-only directed evidence graph. Mass and energy residuals are preserved rather than silently forced to zero:

$$\varepsilon_M = \sum M_{\text{in}} - \sum M_{\text{out}} - \Delta M_{\text{stock}}, \qquad \varepsilon_E = \sum E_{\text{in}} - \sum E_{\text{out}} - \Delta E_{\text{stock}}.$$

The study evaluates the architecture using a synthetic mirror, permissioned pilot subset, or approved post-pilot case. Independent reviewers receive a bounded offline bundle and attempt to reconstruct material flows, gas and energy calculations, leakage treatment, life-cycle boundaries, co-product allocation, and attribute claims under versioned Australian and synthetic comparison adapters. Planned outcomes include parent and source resolution, sensor-quality states, residual magnitude, quantitative discrepancy, exception visibility, adapter differences, double-attribution detection, reconstruction time, and reviewer disagreement. The framework is not a new emissions-reduction methodology and does not establish product efficacy, asset ownership, renewable-energy entitlement, Scope 3 acceptance, or government approval.

Keywords: manure management; anaerobic digestion; farm energy; evidence architecture; mass balance; energy balance; circular economy; environmental attributes; Australian livestock; verification readiness.

Paper 7: Portable Evidence, Local Authority

Role in the sequence: Jurisdictional portability. Tests whether local authority can reinterpret one historical evidence root without rewriting the past or universalising policy.

Abstract

International expansion in agricultural greenhouse-gas and Scope 3 programs often produces country-specific software branches, duplicated evidence pipelines, and inconsistent interpretations of the same underlying activity. This paper evaluates a different proposition: agricultural evidence can remain stable while method compatibility, verification expectations, rights, reporting artifacts, and institutional reliance remain explicitly local and versioned. The study freezes one Australian livestock evidence graph, identified by a cryptographic root digest, and evaluates it against an immutable Australian adapter and at least two synthetic or research-only foreign policy adapters. The underlying receipts, source commitments, event semantics, parent links, signatures, limitations, and offline integrity-verification process remain unchanged.

The architecture follows the thin-waist model developed in Paper 1, the validity separation formalised in Paper 2, the minimum-evidence vocabulary of Paper 3, and the applied evidence profiles of Papers 4–6. Each adapter declares its purpose, authority status, source documents, methodology and policy versions, production-context requirements, eligibility conditions, parameter and model bindings, evidence profile, verification expectations, claim policy, data-governance rules, artifact profile, limitations, and supersession relationships. The experiment compares which evidence fields are reusable, which require semantic mapping or local bindings, which produce missing, `conflicting`, out-of-domain, or conditionally sufficient states, and which require method extension. It also tests methodology supersession: a later adapter version generates a new determination linked to the same historical evidence rather than rewriting the original record.

The paper integrates FAIR data stewardship, agricultural knowledge graphs, international greenhouse-gas inventory tiering, Australian inventory methodology, Scope 3 accounting, digital provenance, greenhouse-gas verification, and data-sovereignty literature. It proposes a weighted and unweighted evidence-reuse profile, an adapter-difference taxonomy, a semantic-portability test, a conformance suite, a migration and supersession protocol, and an independent reconstruction study. The intended contribution is not a universal agricultural methodology. It is a testable contract for preserving evidentiary identity while making policy translation explicit, inspectable, and locally governed. Non-Australian adapters remain research artifacts unless reviewed and adopted by competent authorities.

Keywords: evidence portability; policy adapters; agricultural MRV; conformance; interoperability; semantic portability; methodology versioning; data sovereignty; local authority; Australia.

Contents

Paper 1: A Thin-Waist Architecture for Interoperable Agricultural Evidence

Reader transition. The first part of this document has explained the company and the strategic function of the publication program. The remainder now changes register: each paper must stand on its own as a scientific working manuscript, with narrower claims, explicit methods and non-assertions, and planned results distinguished from completed findings. The papers should be read in order from architecture and trust semantics through evidence sufficiency, deployment, composition, domain extension, and jurisdictional portability. Appendix A contains the literature-review and reference-architecture substrate used to harden the series.

Working-draft note. This manuscript expands Paper 1 from the *AgEvidence Australia Scientific Publication Series: Seven Manuscript First Drafts*. It is a concept manuscript, not an empirical report. Every title, author order, affiliation, method, dataset, result, journal target, and ownership designation remains proposed. No named person should be represented as an author until written consent, contribution, affiliation, conflict, and final-manuscript approval are complete.

Non-negotiable guardrail. Peer-reviewed publication can create a citable reference architecture and influence future guidance, procurement expectations, and standards discussions. It does not itself create an Australian Standard, confer government acceptance, approve a methodology, certify an emissions reduction, determine claim ownership, or replace an accredited validation or verification body.

Abstract

Agricultural greenhouse-gas measurement, reporting, and verification systems are commonly implemented as country-specific methods, proprietary data pipelines, registry workflows, calculator exports, or project-level document collections. These arrangements can be effective for an initial submission, but they couple historical evidence to one interpretation and make later reconstruction expensive when methods, models, buyers, verifiers, claim rules, or jurisdictions change. This paper proposes an Australia-first reference architecture in which agricultural evidence is captured once as a signed, append-only graph and interpreted many times through immutable, versioned country, method, and institutional policy layers.

The architecture is organised around a deliberately narrow "thin waist": a country-neutral receipt envelope; stable event and decision families; source commitments; immutable parent links; identity, role, authority, rights, and consent declarations; explicit limitations and non-assertions; cryptographic signatures; a bounded verification vocabulary; and a common bundle manifest. National methods, eligibility rules, model profiles, uncertainty treatments, claim policies, data requirements, verification expectations, and reporting artifacts remain above this layer as declarative adapters. The design therefore separates what happened and who asserted it from how a later decision-maker interprets it.

Australia is proposed as the first implementation context because the program must accommodate pasture-based and intensive livestock systems, local scientific participation, value-chain actors, co-funding and co-claiming questions, and a standards-alignment pathway without embedding Australian policy into the generic trust kernel. The paper specifies dependency direction, core semantics, receipt families, graph invariants, adapter manifests, reliance artifacts, conformance tests, and an independent reconstruction protocol. The central evaluation asks whether one frozen evidence root can support multiple method determinations and institution-specific reliance artifacts without rewriting historical evidence or changing its cryptographic commitments.

The contribution is not a new emissions-reduction methodology and does not establish scientific efficacy, government approval, verification, legal entitlement, or claim ownership. It is a testable interoperability contract intended to make agricultural evidence portable while preserving local authority over applicability, scientific judgment, verification, rights, and institutional acceptance.

Keywords: agricultural MRV; evidence architecture; interoperability; thin waist; digital provenance; livestock decarbonisation; policy adapters; evidence graph; Australia.

1 Manuscript status, scope, and contribution boundary

This paper is an architecture and methods manuscript. It proposes a reference model, a formal vocabulary, a conformance approach, and a planned evaluation. It does not report completed architecture testing, field results, emissions reductions, verifier determinations, or government acceptance. Any tables labelled "planned results" are reporting templates rather than findings.

The proposed architecture addresses a recurring structural problem in agricultural evidence systems: evidence is often assembled only after a decision context is known. A project may gather documents specifically for one methodology, one registry form, one buyer questionnaire, or one verification engagement. The resulting package can be internally coherent yet difficult to reuse because the evidence and the interpretation have been fused. When the decision context changes, the package must be rebuilt.

The paper's contribution is a separation of concerns. It defines:

1. a stable, country-neutral evidence contract;
2. a set of event and decision receipt families;
3. an append-only evidence graph with explicit source and parent relationships;
4. immutable, versioned country, method, and institutional adapters;
5. bounded reliance artifacts for specific decision contexts;
6. conformance tests that distinguish a genuine implementation from a proprietary export format using similar terminology; and
7. an independent reconstruction protocol that treats missing information, ambiguity, and unsupported inference as measurable outcomes.

The architecture deliberately does *not* define a livestock emissions-reduction method, a national eligibility rule, a registry status, a legal allocation of environmental attributes, or a universal meaning of "verified." Those matters are represented as versioned interpretations or decisions made by authorised actors.

2 Provisional scientific author personas

Persona	Provisional role	Contribution domain / deliverable
Persona A	Lead/corresponding-author persona	Scientific scope; Australian applicability and local-authority boundaries; discussion and conclusion.
Persona B	Scientific coauthor persona	Model identity, applicability, intervention science and global comparability.
Persona C	Scientific coauthor persona	Institutional use cases, Scope 3 translation and reliance-artifact language.
Persona D	Scientific coauthor persona	Measurement provenance, uncertainty boundaries and epistemic distinctions.
Persona E	Scientific coauthor persona	Engineering generality beyond enteric interventions.
Persona F	Scientific coauthor persona	Producer workflow, capture burden, exceptions and farm-operational realism.
Persona G	Technical scientific contributor persona	Formal architecture, schemas, verifier behaviour, conformance fixtures and reproducibility.
Persona H	Conditional contributor persona	Institutional framing, public-interest boundary, disclosures and CRediT governance.

Authorship control. Persona IDs are recruitment placeholders only. A real person is not an author until consent, qualifying contribution, affiliation, conflict disclosure, CRediT allocation, and final-manuscript approval are complete.

3 Introduction

3.1 The reconstruction problem

Agricultural greenhouse-gas evidence is generated across heterogeneous biological, technical, contractual, and institutional systems. A single intervention may involve laboratory characterisation, product manufacturing, purchase and delivery records, feed or grazing operations, animal-cohort definitions, sensor observations, laboratory assays, model executions, scientific review, validation or verification activity, registry processes, contractual allocations, and producer payment. Each component has a different issuer, authority basis, time scale, error structure, and evidentiary meaning.

In many implementations, these components are assembled late as a project dossier. The dossier is designed around the immediate question: does the project satisfy a specified method, can a verifier complete an engagement, or can a buyer rely on a report? This creates a tightly coupled architecture. Evidence fields, transformations, labels, and outputs are shaped by one methodology, one calculator, one registry schema, or one institutional workflow.

Tight coupling can reduce initial implementation time, but it also creates a hidden future liability. A change in baseline, model version, eligibility rule, uncertainty treatment, data policy, verification profile, reporting format, or claim-allocation convention can force reconstruction of the complete package. Producers and technology providers may need to recover source records, re-establish authority, reconcile identities, repeat calculations, and explain transformations that were never explicitly recorded. The historical evidence may still exist, yet its relationship to the previous interpretation may be unclear.

The central architectural question is therefore not merely how to digitise an existing form. It is how to preserve evidence so that later decision-makers can ask new questions without rewriting the past.

3.2 Evidence and interpretation are different objects

This paper advances a simple but consequential design principle: evidence should be preserved independently of the policy interpretation later applied to it. The same agricultural event should not require a different historical record merely because a buyer, method, registry, verifier, or country asks a different question.

An intervention delivery, for example, may be relevant to several later determinations. One method may require proof of delivery and proof of feeding. Another may require batch identity, dose, animal class, and a model-specific calibration domain. A buyer may require a contractual allocation statement. A verifier may require source access and uncertainty documentation. These are not reasons to rewrite the delivery event. They are reasons to link the event to additional evidence and evaluate the resulting graph under different declared rules.

The proposed architecture therefore adopts a “capture once, interpret many times” model:

1. the event or decision is represented as a bounded receipt;
2. the receipt is canonicalised, linked to sources and parents, and signed;
3. receipts accumulate as an append-only graph;
4. a versioned adapter evaluates the graph for a declared method or policy context;
5. a reliance artifact packages the relevant evidence, determination, exclusions, and limitations for a specific institution or decision.

3.3 Why a thin waist

The term *thin waist* is borrowed from protocol architecture. A narrow, stable interface permits diversity above and below it. Below the waist, evidence may originate from farm software, devices, laboratories, remote sensing, intervention providers, contracts, spreadsheets, or human review. Above the waist, methods, countries, verifiers, registries, buyers, lenders, insurers, processors, and government programs apply different criteria.

The thin waist contains only those semantics that must remain common for evidence to be inspectable and portable. It should be narrow enough to remain stable, but expressive enough to prevent materially different propositions from collapsing into one generic “record.” The design challenge is therefore one of disciplined minimalism. Every field added to the waist increases commonality but also increases the risk that local policy is accidentally embedded in the core. Every field omitted increases flexibility but may weaken reproducibility or permit validity laundering.

3.4 Australia-first, not Australia-inside-the-kernel

Australia is proposed as the first implementation context rather than the universal source of policy rules. The originating program identifies Australia as a useful test environment because it combines pasture-based and intensive livestock systems, local scientific capability, real value-chain actors, co-funding and co-claiming questions, and a standards-alignment pathway. The architecture must represent those characteristics faithfully while preserving a country-neutral kernel.

This distinction is essential. An Australia-first implementation may include an Australian adapter, Australian terminology mappings, local production-system categories, and local authority declarations. It must not make the generic receipt kernel depend on Australian eligibility rules, national registry states, or country-specific claim conventions. Otherwise, a purportedly interoperable architecture becomes a national data model with foreign extensions.

3.5 Research questions

The manuscript addresses four primary questions:

1. What is the minimum stable evidence contract reusable across livestock interventions, methods, institutions, and countries?
2. Which semantics must remain country-neutral, and which requirements belong in a versioned Australian policy adapter?
3. Can one evidence root generate multiple compatibility determinations and reliance artifacts without changing historical evidence or cryptographic commitments?
4. What conformance tests distinguish a genuine implementation from a proprietary export format that uses similar terminology but cannot be independently reconstructed?

A fifth, cross-cutting question is whether the architecture preserves uncertainty honestly. A successful implementation must not convert missing evidence into inferred completeness, cryptographic validity into scientific truth, method compatibility into government approval, or one institution's acceptance into universal reliance.

TODO: Replace the literature-review placeholder with a structured review of agricultural MRV architectures, digital provenance, event sourcing, verifiable data structures, interoperability standards, model governance, environmental validation and verification, and protocol-design literature. The final review should distinguish prior technical mechanisms from the specific agricultural governance problem addressed here.

4 Design principles and non-assertions

4.1 Design principles

The architecture is governed by nine principles.

1. Historical evidence is immutable; interpretation is versioned. A receipt is not edited to match a later policy. Corrections, disputes, reversals, and re-evaluations are appended as new receipts linked to the original.

2. Every receipt is a bounded proposition. A receipt states what an identified issuer asserts, under what authority, from which sources, at what time, and with which limitations. It does not inherit broader meaning merely because it appears in a trusted system.

3. Country and institutional policy remains outside the kernel. Eligibility, registry status, claim policy, and local reporting requirements are adapter concerns.

4. Missing information remains visible. Unavailable parents, absent sources, unresolved identity, or incomplete trust material must produce an indeterminate or insufficient-evidence state rather than a silent pass.

5. Observed, measured, calculated, modelled, reviewed, and verified propositions remain distinct. The architecture must preserve the epistemic route by which a value or conclusion was produced.

6. Authority is explicit and scoped. An issuer may be authorised to report delivery but not feeding, to run a model but not approve a method, or to verify a project but not allocate contractual claims.

7. Independent reconstruction is a first-class requirement. A bounded bundle must be inspectable without privileged access to the originating platform.

8. Open interfaces do not require uniform policy. Countries and institutions may disagree while still relying on a common evidence contract.

9. Conformance is behavioural. An implementation conforms because it produces deterministic, inspectable, independently verifiable behaviour, not because it uses the same labels or exposes a similar JSON export.

4.2 Non-assertions

The architecture embeds explicit non-assertions. These are not legal disclaimers appended after the fact; they are semantic boundaries that determine what a receipt or determination can mean.

- A valid signature does not prove that the physical event occurred.
- A source commitment does not prove that the source was accurate or complete.
- A delivery receipt does not prove administration, consumption, efficacy, or method eligibility.
- A model execution receipt does not convert a model output into an observation.
- A scientific review receipt does not constitute an accredited validation or verification determination.
- A compatibility determination does not establish government approval, registry acceptance, or legal entitlement.
- An institutional reliance event is bounded to the named institution, artifact, purpose, time, and conditions.
- An asset issuance receipt does not by itself prove ownership of an environmental attribute or prevent double claiming.

5 Reference architecture

5.1 Dependency direction

The architecture imposes a one-way dependency:

$$\begin{aligned} \text{institutional application} &\rightarrow \text{versioned adapter and artifact profile} \\ &\rightarrow \text{global agricultural evidence contract} \\ &\rightarrow \text{generic receipt kernel.} \end{aligned}$$

The generic receipt kernel must never import national eligibility rules, registry states, buyer claim conventions, or method-specific outcomes. Country and institutional layers may depend on the kernel; the kernel must not depend on them.

This direction prevents national exceptions from becoming cryptographic branches. It also allows a common verifier to inspect bundles from multiple jurisdictions because canonicalisation, signatures, parent links, source commitments, and bundle rules remain stable even when policy interpretations differ.

5.2 The stable evidence contract

The proposed thin waist contains the following elements:

1. a canonical envelope and schema digest;
2. stable event and decision semantics;
3. source commitments and source-access metadata;

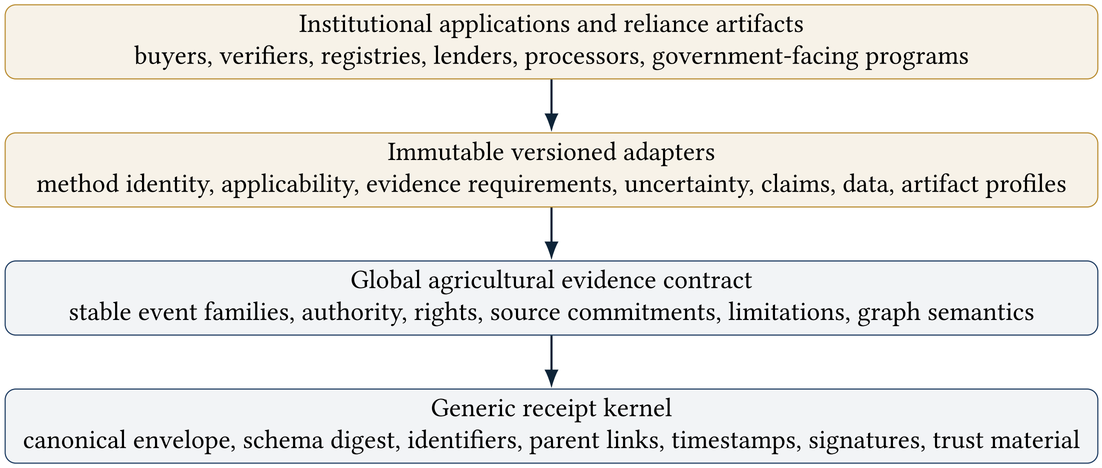

Figure 1: Dependency direction in the proposed thin-waist architecture. Higher layers may depend on lower layers; the generic kernel does not import national or institutional policy.

4. immutable parent links;
5. issuer, subject, identity, role, authority, rights, and consent declarations;
6. event time, recording time, and effective time where relevant;
7. policy references without embedded policy outcomes;
8. explicit limitations, assumptions, and non-assertions;
9. signatures and trust material;
10. a common integrity-verification vocabulary; and
11. a common bundle manifest.

A receipt records a bounded proposition made by an identified issuer within a declared authority scope. In compact form, let a receipt be

$$R = (h, \tau, i, s, a, t_e, t_r, P, C, L, \sigma)\,,$$

where:

- h is the canonical receipt identifier or digest;
- τ is the stable receipt family and subtype;
- i is the issuer identity and role;
- s is the subject or subjects;
- a is the declared authority and rights scope;
- t_e is event time and t_r is recording time;
- P is the ordered set of immutable parent references;
- C is the set of source commitments and content fields;
- L is the set of limitations, assumptions, and non-assertions; and
- σ is the signature and associated trust material.

The receipt identifier is computed over a canonical representation:

$$h = H(\mathrm{Canon}(R \setminus \{h, \sigma\}))\,,$$

and the issuer signs the canonical commitment:

$$\sigma = \mathrm{Sign}_{k_i}(h).$$

The exact cryptographic suite is an implementation profile rather than an agricultural semantic. Conformance requires deterministic canonicalisation, explicit algorithm identifiers, and preservation of the bytes or canonical data necessary for independent verification.

5.3 Stable receipt families

Receipt families describe event or decision type rather than national recognition. The initial proposed families are shown in Table 2.

Table 2: Proposed stable receipt families and their semantic boundaries.

Family	Bounded proposition	Must not be inferred
Authority	An identified actor holds a declared role or delegated authority for a specified scope and period.	Universal authority, authority outside scope, or continuing authority after expiry or revocation.
Rights	A party declares specified rights, licences, restrictions, or allocation terms over data, products, or environmental attributes.	Legal validity in every jurisdiction or absence of competing rights.
Identity binding	An identifier is linked to a person, organisation, property, device, herd, cohort, product, batch, or asset.	Biological, legal, or operational equivalence beyond the binding scope.
Consent	A subject or authorised party grants a specified permission for a stated purpose and period.	Consent for unrelated purposes or future policy versions.
Intervention	A product or practice was manufactured, delivered, installed, administered, or otherwise implemented as specified.	Efficacy, persistence, eligibility, or emissions reduction.
Observation	A phenomenon or state was directly observed under declared conditions.	Measurement accuracy beyond the observation process.
Measurement	A quantity was measured using a declared instrument, method, calibration, unit, and quality process.	Applicability to an undeclared population or decision.
Model execution	A declared model version was executed with identified inputs, parameters, and environment.	That outputs are observations, scientifically adequate, or method-approved.
Evidence candidate	Material is nominated as potentially relevant to a specified decision.	Sufficiency, acceptance, or verification.
Evidence gap	Required or expected material is missing, inaccessible, inconsistent, or unresolved.	Failure of the underlying event or automatic ineligibility unless an adapter says so.
Human review	An identified reviewer reached a bounded judgment using declared evidence and criteria.	Accreditation, institutional acceptance, or universal scientific consensus.
Verification determination	An authorised verification actor reached a declared determination under a specified profile.	Government approval, asset ownership, or acceptance by other institutions.
Asset issuance	A named system issued an identified asset or record under declared rules.	Scientific validity, legal title, or absence of duplicate claims elsewhere.
Contribution	A party's quantified or qualitative contribution to an intervention or outcome is recorded.	Ownership, payment entitlement, or claim priority unless separately established.

Family	Bounded proposition	Must not be inferred
Claim	A party made a bounded environmental, product, financial, or reporting claim.	Truth, exclusivity, acceptance, or legal enforceability.
Retirement	An identified asset was retired or cancelled in a named system for a stated purpose.	Elimination of all possible duplicate assertions outside that system.
Payment	A payment obligation or transfer occurred under declared terms.	Environmental validity or ownership of the underlying outcome.
Methodology delta	A difference between method or adapter versions is declared.	That the newer version is universally superior or retroactively applicable.
Artifact assembly	A bounded package was assembled from identified receipts, sources, determinations, and profiles.	Completeness beyond the manifest or suitability for every decision.
Reliance event	An institution accepted, rejected, relied on, or requested more evidence from a specific artifact for a bounded purpose.	Universal acceptance or binding effect on future institutions.

The family list is expected to evolve through controlled versioning. A new family should be added only when the proposition cannot be represented without semantic loss using an existing family and subtype. Local terminology may be mapped to a stable family, but the mapping must remain explicit.

5.4 Evidence graph

Let the evidence graph be a directed acyclic graph $G = (V, E)$, where each vertex $v \in V$ is a receipt and each edge $(u, v) \in E$ means that receipt v references receipt u as a parent. The graph is append-only: new vertices and edges may be added, but an existing receipt's canonical content is not changed.

The graph supports several relationship types:

- **derivation**: a calculation or model output depends on declared inputs;
- **authority**: an event receipt depends on a role, delegation, or consent;
- **identity**: a record depends on bindings among identifiers and real-world subjects;
- **correction**: a later receipt corrects or supersedes a prior proposition without deleting it;
- **review**: a human or institutional judgment depends on an identified evidence set;
- **allocation**: a claim, contribution, payment, or right depends on prior contractual or evidentiary receipts; and
- **assembly**: an artifact depends on a bounded set of receipts, sources, adapters, and determinations.

A correction is therefore not an in-place edit. If R_1 contains an incorrect quantity, a correction receipt R_2 references R_1, states the corrected proposition, identifies the reason and authority for correction, and preserves both histories. A later adapter or artifact profile determines whether R_1, R_2, or both must be displayed.

5.5 Source commitments and disclosure

A receipt may commit to source material without embedding the full source. A source commitment should identify:

- source digest and algorithm;
- media type and encoding;
- source creator or custodian where known;
- creation, observation, and acquisition times;
- access conditions and retention expectations;

- redaction, transformation, aggregation, or de-identification steps;
- whether the source is embedded, externally resolvable, permissioned, or unavailable; and
- limitations on independent access.

A source digest proves only that a disclosed object matches the committed bytes. It does not prove the source's truth, completeness, calibration, chain of custody before commitment, or scientific adequacy. Those questions require additional receipts and later evaluation.

5.6 Identity, authority, rights, and consent

Agricultural evidence frequently crosses organisational boundaries. A producer may operate a property but use a contractor, reseller, laboratory, processor, model provider, project developer, verifier, and buyer. A single identifier field cannot represent the resulting authority structure.

The architecture separates:

1. **identity**: who or what is being referred to;
2. **role**: the function performed;
3. **authority**: the scope within which an actor may assert or decide;
4. **rights**: data, product, contractual, or environmental-attribute interests;
5. **consent**: permission for a defined use; and
6. **reliance**: a later institution's bounded decision.

These relationships are time-bounded and revocable. An authority receipt may be superseded or revoked, but the historical fact that it existed remains part of the graph. A later evaluator must apply the authority state that was effective for the event and decision being assessed.

5.7 Country, method, and institutional adapters

An adapter is an immutable, versioned policy pack. It declares the rules by which a graph is interpreted for a specific context. At minimum, an adapter should identify:

- adapter identity, jurisdiction, institution, and version;
- referenced method identity and version;
- applicability criteria and exclusions;
- required and optional evidence;
- accepted units, conversions, and temporal windows;
- model identity, calibration domain, and supersession rules;
- uncertainty treatment and missing-evidence behaviour;
- verification expectations and authorised roles;
- rights, consent, and claim requirements;
- data-access, retention, and disclosure policy;
- artifact profiles and mandatory limitation language;
- effective date, superseded versions, and authority; and
- explicit non-assertions.

Historical versions are never overwritten. If method M_v is replaced by M_{v+1}, the same evidence root can be evaluated under both versions:

$$D_v = \Psi(G, A_v), \qquad D_{v+1} = \Psi(G, A_{v+1}),$$

where A_v is the adapter at version v and D_v is the resulting determination. The evaluation may differ, but the historical receipts and their digests remain unchanged.

5.8 Deterministic and judgment-bearing components

Not every adapter rule is mechanically decidable. The architecture distinguishes:

- **deterministic checks**, such as schema validity, required-field presence, unit compatibility, date-window logic, signature verification, and parent resolution;
- **parameterised calculations**, such as conversions or model executions with explicit versions and inputs;
- **bounded human judgments**, such as scientific applicability, materiality, adequacy of explanation, or acceptance of an exception; and
- **institutional decisions**, such as verification determination, buyer reliance, or registry acceptance.

An adapter may specify that a human judgment is required, identify the evidence to be considered, and constrain the allowed output states. It must not disguise a judgment as a deterministic calculation.

5.9 Reliance artifacts

A reliance artifact is a bounded package for a specific decision context. It references:

- the evidence root and bundle manifest;
- adapter and method identities and versions;
- verifier or review profile where applicable;
- claim and rights policy;
- included and excluded evidence;
- unresolved gaps, exceptions, disputes, and assumptions;
- deterministic results and human judgments;
- institution, purpose, and validity period;
- display and disclosure requirements; and
- explicit non-assertions.

The artifact should be reproducible from the declared graph and versions. A later institution may assemble a different artifact from the same evidence root because its purpose, adapter, or risk tolerance differs.

6 Australia adapter manifest

The Australia adapter is proposed as the first concrete implementation. It should not be treated as a universal policy model or an assertion of formal government or standards status. Its purpose is to test whether Australian scientific and production context can be represented above a country-neutral core.

A draft manifest should include the categories in Table 3.

Table 3: Draft Australia adapter manifest categories.

Category	Required content
Identity and authority	Adapter identifier, version, accountable issuer, effective date, superseded versions, approval process, and explicit statement of non-governmental status unless formally changed.
Production context	Australian livestock production categories, regional and seasonal descriptors, pasture and intensive system distinctions, herd and cohort definitions, and local terminology mappings.
Method references	Method identity and version, relevant eligibility boundaries, required calculations, model references, and uncertainty treatment.
Evidence requirements	Required receipt families, source expectations, temporal windows, minimum identity and authority bindings, and missing-evidence outcomes.
Model profile	Model name, version, calibration domain, input requirements, parameter sources, execution environment, and supersession rules.
Verification profile	Required reviewer or verifier roles, accreditation or competence requirements where applicable, evidence-access expectations, and allowed determination states.
Rights and claims	Required rights, consent, contribution, allocation, claim, and retirement evidence, while avoiding automatic legal conclusions.
Data policy	Access, retention, confidentiality, de-identification, redaction, permitted disclosure, and independent reproduction conditions.
Artifact profiles	Technical bundle, scientific review package, verifier package, buyer package, and other bounded outputs, each with mandatory limitations.
Exception handling	Out-of-domain conditions, missing data, disputed records, corrections, method extensions, and escalation pathways.

The final adapter must be reviewed by Australian scientific and institutional participants. Dissent should be recorded rather than silently harmonised. Where participants disagree about whether a requirement is generic or local, the default should be to keep it outside the kernel until broader evidence supports promotion into the common contract.

7 Methods

7.1 Study design

The proposed study is a design-science and conformance evaluation. It combines requirements decomposition, ontology development, reference implementation, adapter construction, synthetic or permissioned case assembly, independent reconstruction, and structured review.

The evaluation unit is not a claimed emissions reduction. It is the behaviour of the evidence architecture under controlled scenarios. The study asks whether the architecture preserves provenance, separates evidence from interpretation, exposes gaps, supports multiple versioned evaluations, and can be independently reconstructed.

7.2 Requirements decomposition

Requirements will be classified into four layers:

1. **generic trust**: canonicalisation, identifiers, signatures, parent links, source commitments, and bundle structure;
2. **global agricultural evidence**: stable event families, authority, rights, biological and operational subjects, measurement and model distinctions, and correction semantics;
3. **country and method interpretation**: applicability, exclusions, model profiles, uncertainty, evidence sufficiency, and verification expectations; and
4. **institutional reliance**: artifact profiles, decision purpose, acceptance criteria, claim policy, disclosure, and reliance outcomes.

Each candidate field will be tested against a placement question:

> Does every country and intervention require this semantic to preserve the meaning and provenance of evidence, or does the requirement belong in a versioned local profile?

Eligibility, registry status, government acceptance, and claim ownership are excluded from the core by default.

The decomposition process will maintain a decision log containing the candidate field, proposed layer, alternatives considered, rationale, dissent, and downstream implications. Fields promoted into the thin waist must have a documented cross-context justification.

7.3 Domain ontology workshop

A structured workshop will involve proposed Scientific Advisory Board authors, Australian scientific participants, technical architects, producer or implementation representatives, and independent reviewers. Participants will classify:

- subjects and identifiers;
- event and decision types;
- evidence types and epistemic status;
- authority, rights, and consent roles;
- lifecycle and decision states;
- units, temporal semantics, and model identity;
- limitations and non-assertions; and
- correction, dispute, exception, and supersession behaviour.

The workshop will use scenario cards rather than abstract terminology alone. Example cards may include a product delivery without feeding evidence, a sensor file with an expired calibration certificate, a model execution outside its declared calibration domain, a producer-authorised correction, a verifier request for additional evidence, and a buyer rejection for contractual reasons despite intact scientific evidence.

Participants will independently classify each scenario before discussion. Agreement, disagreement, and changes after discussion will be recorded. The purpose is not to force consensus on policy but to identify which distinctions the architecture must preserve.

7.4 Reference evidence graph

One Australian beef evidence graph will be constructed using one of three approved data routes:

1. a synthetic mirror that reproduces the structure and difficulty of a real workflow without using project data;
2. a permissioned, de-identified pilot subset; or
3. a post-pilot dataset released only after productisation, rights, and confidentiality issues are resolved.

The graph should include, where available:

- identity, authority, rights, and consent;
- intervention identity and product characterisation;
- purchase, delivery, storage, administration, and operational evidence;
- herd, cohort, property, and production-context bindings;
- observations, measurements, laboratory results, and sensor provenance;
- calculation and model-execution receipts;
- scientific review and verification-readiness events;

- evidence gaps, exceptions, disputes, and corrections;
- claim, contribution, allocation, payment, and reliance events; and
- artifact assembly under at least two adapter profiles.

No synthetic event will be presented as an empirical finding. Synthetic content must be visibly labelled in the source receipts, bundle manifest, artifact, and manuscript.

7.5 Reference implementation

A reference implementation will include:

- machine-readable schemas and human-readable semantic documentation;
- canonicalisation and hashing rules;
- signing and verification utilities;
- bundle assembly and offline verification;
- adapter parsing and version commitment;
- graph traversal and parent resolution;
- deterministic conformance fixtures;
- display of limitations, gaps, and non-assertions; and
- reproducible environment manifests.

The implementation should be published under terms that permit independent inspection and implementation. The study should not require use of AgEvidence's hosted workflow to verify a bounded bundle.

7.6 Adapter implementation

The same evidence root will be evaluated under:

1. an Australia adapter;
2. at least one synthetic second-country adapter designed to introduce materially different requirements; and
3. at least one institution-specific artifact profile.

The synthetic second-country adapter is a software and architecture test only. It must not be represented as a valid foreign compliance pathway. Its purpose is to vary requirements such as eligibility windows, accepted model versions, evidence granularity, uncertainty treatment, claim policy, or artifact disclosure while preserving the same evidence root.

7.7 Conformance test suite

The conformance suite will cover at least the following test classes.

Canonical receipt validation. Equivalent semantic input must produce the same canonical commitment under the declared profile. Non-equivalent input must not collide under ordinary test conditions. Unknown fields, number formats, time zones, ordering, and character encodings must be tested.

Signature and trust validation. Valid, invalid, expired, revoked, unknown, and unavailable trust material must produce distinguishable outcomes. An unavailable certificate or key must not be silently treated as valid or invalid when the correct state is indeterminate.

Parent and source resolution. Missing parents, circular references, duplicate identifiers, unavailable sources, digest mismatch, and inconsistent media types must be detected. Missing material must remain visible.

Authority and rights scope. Tests must show that a valid identity does not imply authority, and that authority for one event or period does not extend automatically to another.

Observation, measurement, calculation, and model separation. A model output must retain links to inputs, model identity, version, parameters, and environment. It must not be relabelled as an observation.

Adapter version commitment. Every determination must identify the exact adapter and method versions. Re-running the same deterministic evaluation with the same graph and versions should produce the same machine result.

Method supersession. The same evidence root must be evaluated under at least two adapter versions. Determinations may differ, but historical receipt digests must remain unchanged.

Artifact assembly. A manifest must identify every included receipt and source, every exclusion, and every version. A verifier should detect omitted mandatory evidence, undeclared transformations, or mismatched adapter commitments.

Unsupported inference detection. The suite will include deliberately misleading cases, such as a valid delivery receipt presented as proof of efficacy, or a cryptographically valid model execution presented as government-approved. Conforming software and reviewers should flag the unsupported inference.

7.8 Independent reconstruction

An independent reviewer will receive a bounded bundle, public specifications, verifier instructions, and a task protocol rather than access to AgEvidence's hosted workflow. The reviewer will be asked to:

1. verify bundle and receipt integrity;
2. resolve parent and source relationships;
3. identify missing, inaccessible, or inconsistent dependencies;
4. reconstruct selected calculations or model executions where licensed and feasible;
5. run the declared adapters;
6. assemble or inspect specified reliance artifacts;
7. identify unsupported inferences and hidden judgments; and
8. record ambiguity, required clarifications, and time spent.

Primary outcomes will include reconstruction time, proportion of receipts and sources resolved, number and type of missing dependencies, deterministic agreement, reviewer disagreement, unsupported inference detection, and unresolved judgment.

7.9 Reviewer study

At least two independent reviewers should assess the same bounded bundle. Reviewers should have relevant but different expertise, such as agricultural science, verification, data architecture, or producer implementation.

Inter-reviewer agreement will be reported for categorical tasks. Where appropriate, Cohen's κ may be used for two reviewers:

$$\kappa = \frac{p_o - p_e}{1 - p_e},$$

where p_o is observed agreement and p_e is agreement expected by chance. The statistic will not be used to imply that judgment has become objective. Disagreement categories and reasons will be reported qualitatively.

7.10 Evaluation metrics

The study will report:

- canonicalisation pass, fail, and indeterminate rates;
- signature and trust-resolution outcomes;
- parent and source resolution rates;
- number of unresolved identity, authority, rights, or consent issues;
- adapter determinism for machine-decidable rules;
- divergence among adapter versions and explanation of each divergence;
- evidence-root stability across re-evaluations;
- independent reconstruction time and effort;
- reviewer agreement and disagreement;
- number of unsupported inferences detected;
- number and severity of evidence gaps;
- artifact completeness and disclosure compliance; and
- implementation burden for source systems and producers.

Because the architecture explicitly permits indeterminate outcomes, success is not defined as maximising pass rates. A correct indeterminate result is preferable to an unsupported pass.

8 Planned evaluation and results reporting

Working-draft note. No architecture evaluation has been completed. This section is a reporting template. The final paper must report pass, fail, and indeterminate outcomes, including failures and reviewer disagreement.

Table 4: Planned architecture-conformance results template.

Test	Result	Interpretation rule
Canonical receipt validation	TBD	Encoding and digest integrity only; no claim about physical truth.
Signature and trust validation	TBD	Distinguish valid, invalid, and indeterminate trust states.
Parent and source resolution	TBD	Missing material produces a visible gap or indeterminate state, not inferred completeness.
Identity and authority scope	TBD	Valid identity does not imply authority outside the declared scope.
Observation/model distinction	TBD	Model outputs retain model identity and are not relabelled as observations.
Adapter version commitment	TBD	Every determination identifies immutable adapter and method versions.
Method supersession	TBD	Re-evaluation may change determinations while historical receipt digests remain unchanged.
Independent reconstruction	TBD	Report time, gaps, disagreement, clarification requests, and unresolved judgment.
Evidence-root reuse	TBD	The same root supports multiple bounded determinations and artifacts without rewriting history.

Test	Result	Interpretation rule
Unsupported inference detection	TBD	Report which misleading propositions were detected, missed, or disputed.
Artifact assembly	TBD	Report included, excluded, missing, and mismatched materials.
Producer and source burden	TBD	Report capture steps, duplication, manual interventions, and exception-handling burden.

The final results section should include negative results. Examples include non-deterministic adapter behaviour, hidden dependencies on hosted services, ambiguous authority, source material that cannot be independently accessed, or a receipt family that fails to preserve a scientifically material distinction. Such outcomes are architecture findings rather than implementation embarrassments.

9 Worked architecture scenarios

The following scenarios are conceptual test cases and are not findings.

9.1 One delivery, two method versions

A producer receives an intervention batch. The delivery receipt identifies supplier, producer, property, batch, quantity, time, source documents, and signatures. A later feeding receipt references the delivery and identifies cohort, dose, period, and operational evidence. A model execution references both receipts and a declared model version.

Adapter A_1 requires delivery, feeding, and a model version that was current at the time. Adapter A_2 later requires an additional product-characterisation field and a newer model version. Evaluation under A_1 may produce "eligible with conditions," while A_2 may produce "insufficient evidence" or "method extension required." The historical delivery and feeding receipts do not change. A new product-characterisation receipt or model execution may be appended.

9.2 Cryptographically valid but scientifically out of domain

A sensor file, calibration record, model input bundle, and model execution are all cryptographically valid. However, the animal class or production conditions fall outside the declared calibration domain. The integrity layer passes. The method adapter returns "outside current method" or "method extension required." A scientific reviewer may record a bounded judgment, but the architecture does not allow the integrity result to be displayed as scientific adequacy.

9.3 Method-compatible but awaiting verification

A deterministic evaluator finds that required evidence is present and the declared context satisfies the adapter's machine-decidable rules. The system produces a compatibility state but no verification determination. A verifier later requests additional source access. Both states coexist: method-compatible according to the declared evaluator, and verification pending or additional evidence requested. Neither state erases the other.

9.4 Institutional rejection despite intact evidence

A buyer rejects a reliance artifact because its contract requires a different claim-allocation term or disclosure. The evidence root remains intact, and a scientific review may remain favourable. The rejection is represented as a bounded reliance event, not as proof that the underlying evidence is invalid.

9.5 Correction without erasure

A producer discovers that an administration quantity was recorded in the wrong unit. The original receipt remains in the graph. A correction receipt references it, identifies the error, provides the corrected quantity and unit, and includes authority and source evidence. Re-running the adapter produces a new determination linked to the same historical branch plus the correction. The artifact profile determines how both records are displayed.

10 Discussion

10.1 Interoperability without uniformity

The architecture separates interoperability from uniformity. A common evidence contract does not require countries to share methods, emission factors, eligibility criteria, legal rights, or claim rules. It requires evidence and interpretations to be linked transparently.

This distinction matters because agricultural applicability can vary materially within one country, let alone across jurisdictions. Production system, climate, feed base, animal class, intervention mode, model calibration, measurement method, and institutional purpose may all affect whether evidence is relevant. A thin waist should preserve those distinctions without forcing them into one universal policy.

10.2 The architecture as public infrastructure

Publishing the envelope, semantics, adapter interface, conformance behaviour, and limitations permits independent scrutiny and competing implementations. Commercial value can remain in source integrations, evidence operations, method mapping, verification readiness, artifact production, service commitments, migrations, and dispute handling.

The architecture should therefore be evaluated as public infrastructure rather than as a disguised product description. Independent implementation is a success criterion. A reference model that only the originating platform can interpret has not achieved interoperability.

10.3 Judgment is exposed, not automated away

The model makes model execution, scientific review, verification determination, and institutional reliance explicit events. It does not automate away judgment. Instead, it records where judgment occurs, who is authorised to make it, what evidence was available, which criteria were applied, and which limitations remain.

This is especially important in agricultural systems, where biological variability, incomplete records, model-domain questions, and operational exceptions are common. A system that hides judgment behind a single status may appear simple but is difficult to audit. A system that exposes judgment may appear more complex but is more honest and reconstructable.

10.4 Validity laundering and interface design

A technically correct backend can still mislead users if interfaces collapse distinct states. The conformance model should therefore extend to labels, summaries, exports, and generated reports. A green checkmark for signature validity must not be positioned or worded so that users reasonably infer scientific validity or government approval.

Interface conformance should test:

- whether integrity, compatibility, review, verification, and reliance states are separately displayed;
- whether indeterminate and insufficient-evidence states are visible;
- whether the exact adapter and method versions are shown;
- whether limitations survive export and print;
- whether excluded evidence is disclosed; and
- whether a reliance event names the institution, purpose, date, and conditions.

10.5 Method change as re-interpretation

A major benefit of the architecture is that method change becomes re-interpretation rather than historical rewriting. A new adapter can ask a new question of an old evidence root. It may determine that additional evidence is required, that an older model is no longer accepted, or that an event falls outside a new eligibility boundary. Those outcomes are linked to the new adapter version.

This approach creates a transparent record of policy evolution. It also allows researchers and institutions to compare why determinations changed: because evidence changed, because policy changed, because a model changed, or because a reviewer reached a different judgment.

10.6 Open core and commercial boundary

The originating portfolio distinguishes an open reference architecture from commercial implementation. The scientific manuscript should preserve that boundary. Core schemas, semantics, conformance fixtures, and limitations should be inspectable and independently implementable. Proprietary services may provide operational value above and below the core, but they should not be necessary to understand a receipt or verify a disclosed bundle.

The manuscript must avoid competitor-targeting language. The scientific claim is that a thin-waist design can improve portability and reconstruction if it passes the proposed tests. Market advantage, if any, should arise from implementation quality rather than exclusionary control of the reference model.

11 Limitations

This draft has several material limitations.

First, the literature review is incomplete. The final manuscript must position the architecture against existing work in agricultural MRV, provenance, event sourcing, verifiable data structures, interoperability, model governance, environmental validation and verification, and human-computer interaction.

Second, the ontology and receipt families have not completed structured expert review. The current vocabulary may omit material distinctions, contain unnecessary families, or place requirements at the wrong layer.

Third, no approved Project 4030 dataset, final Australian method, accredited verifier determination, or evidence of government acceptance is available in this draft. The Australia example must remain synthetic, permissioned, or post-pilot.

Fourth, synthetic adapters test software portability but not legal, scientific, or institutional adequacy in another jurisdiction. They must not be represented as foreign compliance pathways.

Fifth, cryptographic commitments protect integrity after capture. They do not prove that an event occurred, that an instrument was correctly installed, that a measurement was scientifically valid, that a source was complete, or that an issuer acted honestly.

Sixth, append-only graphs preserve history but may increase complexity, storage, and disclosure burden. Correction and dispute display must be carefully designed so that users see current status without losing historical context.

Seventh, deterministic adapter behaviour is only possible for rules that are themselves explicit and machine-decidable. Scientific judgment, materiality, exception acceptance, and legal interpretation may remain contested.

Eighth, independent reconstruction may be constrained by confidential sources, intellectual property, privacy, or licence restrictions. The study must report such constraints rather than treating them as implementation details.

Ninth, evidence portability does not imply claim portability. Rights, allocation, consumer claims, reporting claims, and legal consequences may remain local and contractual.

Finally, publication of the architecture does not create an Australian Standard, establish government acceptance, approve a methodology, certify an emissions reduction, determine claim ownership, or replace accredited validation or verification.

12 Conclusion

This paper proposes a narrow, testable contract for agricultural evidence. Its central move is to separate stable evidence semantics from versioned methods and institutional decisions. Receipts preserve bounded propositions, source commitments, authority, parent relationships, limitations, and signatures. Receipts accumulate as an append-only graph. Immutable adapters interpret that graph for declared methods, jurisdictions, and institutions. Reliance artifacts package the resulting evidence and decisions for bounded purposes.

The architecture succeeds only if it can be independently implemented and reconstructed; if missing information remains visible; if one evidence root can be evaluated under multiple versions without rewriting history; and if integrity, compatibility, scientific judgment, verification, and reliance are not collapsed into one status.

The next steps are to complete the literature review, freeze a versioned vocabulary and envelope, select an approved data route, implement the Australia and synthetic comparison adapters, run the conformance suite, commission independent reconstruction, disclose failures and disagreement, and revise the architecture through controlled change.

13 Declarations

Funding and conflicts of interest

TODO: Disclose AgEvidence support, employment, consulting, advisory, equity, intellectual-property, and other relevant relationships for every contributor. Distinguish scientific authorship from advisory status and operational sponsorship.

Author contributions

TODO: Maintain a live CRediT contribution record. Confirm author order and corresponding-author responsibility only after written consent and final contribution review.

Data availability

TODO: Select and document one data route: synthetic mirror, permissioned pilot subset, or post-pilot release. Identify access controls, de-identification, source permissions, retention, and archival location.

Code and specification availability

TODO: Archive schemas, canonicalisation rules, verifier, adapters, conformance fixtures, example bundles, environment manifests, version history, and release checksums.

Ethics, privacy, and legal review

TODO: Determine whether human-participant, producer, commercial, privacy, or contractual approvals are required. Complete legal review of data rights, trademarks, confidentiality, and prohibited claims.

Use of synthetic materials

Any synthetic receipt, adapter, source, result, or scenario must be clearly marked in machine-readable and human-readable forms. Synthetic materials may test architecture behaviour but do not establish legal, scientific, or institutional adequacy.

Standards and authority disclaimer

This manuscript does not claim to create or represent an Australian Standard, government-approved methodology, accredited verification outcome, registry decision, legal opinion, or ownership determination.

14 Executable TODO register

Table 5: Paper 1 executable TODO register.

Task	Owner	Exit evidence
Complete peer-reviewed literature review and claims map	Persona A with scientific technical editor	Search protocol, screened library, evidence table, and revised introduction.
Freeze vocabulary and receipt envelope v0.1	Persona G; approval by Scientific Chairs	Versioned specification, decision log, schemas, and examples.
Complete ontology workshop	Scientific Chairs and Program Manager	Participant list, materials, classifications, dissent log, and revisions.

Task	Owner	Exit evidence
Choose Australian data route	Data Steward; decision by Persona H	Written permission and release plan, or approved synthetic protocol.
Build reference evidence graph	Technical and data team	Versioned bundle with manifest, source commitments, and test labels.
Implement Australia adapter	Scientific and technical owners	Immutable adapter package, authority statement, tests, and limitations.
Implement synthetic comparison adapter	Technical team; scientific review	Clearly labelled synthetic adapter and difference rationale.
Publish offline verifier and conformance fixtures	Reproducibility engineer	Reproducible release, environment manifest, checksums, and documentation.
Run independent reconstruction	Independent reviewers	Time logs, issue logs, determinations, disagreement, and reproducibility report.
Evaluate producer and source-system burden	Persona F and implementation team	Workflow map, burden measures, exceptions, and mitigation options.
Resolve architecture deltas through change control	Scientific Chairs	Versioned issue decisions propagated across portfolio papers.
Complete CRediT, disclosures, and final author approval	Persona A	Signed authorship, affiliation, conflict, and approval records.

Source note

This expanded draft is derived from Paper 1 and the portfolio guardrails, ownership tables, sequencing rules, and disclosure language in the internal working document *AgEvidence Australia Scientific Publication Series: Seven Manuscript First Drafts*, Version 0.1, August 2026. No external literature has been silently added. The peer-reviewed literature review remains an explicit TODO.

Paper 2: Beyond "Verified"

Working-draft note. This manuscript expands Paper 2 from the *AgEvidence Australia Scientific Publication Series: Seven Manuscript First Drafts.* It is a conceptual governance and methods manuscript. The state labels, transition rules, worked cases, interface requirements, and conformance tests remain hypotheses to be tested through literature review, expert elicitation, case classification, interface audit, and independent reproduction.

Non-negotiable guardrail. A compatibility result indicates how declared evidence maps to a specified method and policy version. It is not government approval, an accredited validation or verification determination, legal advice, a determination of claim ownership, or universal institutional acceptance.

Abstract

Digital agricultural measurement, reporting, and verification systems routinely use the word "verified" to describe materially different propositions: that a digital record is intact; that an activity appears compatible with a scientific method; that a scientific reviewer regards an interpretation as credible; that an accredited validation or verification body has reached a determination; or that a buyer, registry, lender, processor, government-facing program, or other institution has relied on an artifact. Collapsing these meanings creates validity laundering, in which confidence generated at one layer is transferred rhetorically or visually to another layer without the required evidence, authority, or decision process.

This paper proposes a three-layer validity model for agricultural evidence systems. Cryptographic integrity is limited to whether canonical commitments, signatures, parent relationships, trust material, and bundle consistency can be evaluated as valid, invalid, or indeterminate. Method compatibility is produced only by a versioned policy evaluator and records whether declared evidence is eligible, eligible with conditions, outside the current method, requires method extension, lacks sufficient evidence, or remains unassigned. Institutional reliance records whether a named institution accepted, relied on, rejected, or requested more evidence from a specific artifact for a bounded purpose. Scientific judgment and accredited verification remain explicit events linked to these layers rather than being silently absorbed into them.

The paper formalises the three state machines, defines prohibited inferences and transition rules, specifies API and interface requirements, and proposes a mixed-method evaluation using concept analysis, structured expert elicitation, inter-rater classification, interface and report audit, and conformance cases. Australia-first livestock scenarios are used to test whether systems preserve ambiguity, missing evidence, model-domain limits, and institution-specific decisions. The contribution is a governance vocabulary and conformance framework for transparent decision support. It does not replace scientific peer review, accredited validation or verification, legal analysis, or Australian government and standards authority.

Keywords: validity laundering; verification; agricultural MRV; cryptographic integrity; method compatibility; institutional reliance; governance; conformance; interface design.

1 Manuscript status and contribution boundary

This paper defines a governance vocabulary and a planned evaluation. It does not report completed elicitation, inter-rater agreement, interface-audit findings, verifier decisions, government acceptance, or empirical proof that the proposed state model reduces user error. Tables labelled "planned results" are reporting templates, not findings.

The manuscript depends on Paper 1 for the receipt envelope, evidence graph, adapter versioning, and bounded reliance-artifact concepts. It contributes the validity model that later portfolio papers should reuse. Applied papers must not create independent meanings for integrity, method compatibility, scientific judgment, accredited verification, institutional reliance, indeterminate evidence, or prohibited inference without versioned portfolio change control.

2 Provisional scientific author personas

Persona	Provisional role	Contribution domain / deliverable
Persona C	Lead/corresponding-author persona	Validity framing, institutional language, reliance, bounded decision use and governance interpretation.
Persona D	Scientific coauthor persona	Measurement, uncertainty and distinctions among observation, calculation, model output, judgment and verification.
Persona A	Scientific coauthor persona	Australian method, standards and local-authority context.
Persona B	Scientific coauthor persona	Model compatibility, calibration-domain logic and evidence sufficiency.
Persona G	Technical scientific contributor persona	Formal state model, API/interface states, transitions and audit/conformance tests.
Persona H	Conditional contributor persona	Institutional scenarios, communications-risk review, disclosures and CRediT governance.

Authorship control. Persona IDs are recruitment placeholders only. A real person is not an author until consent, qualifying contribution, affiliation, conflict disclosure, CRediT allocation, and final-manuscript approval are complete.

3 Introduction

3.1 The overloaded status problem

The word "verified" compresses complexity into a reassuring status. In agricultural greenhouse-gas systems, however, it may refer to at least five propositions:

1. a digital record has not been altered;
2. an activity appears to satisfy a specified method;
3. a scientific reviewer regards an interpretation as credible;
4. an accredited validation or verification body has reached a determination; or
5. an institution has relied on an artifact for a bounded decision.

These propositions involve different evidence, accountable actors, standards, and consequences. A signature check may be performed automatically. Method compatibility may depend on a versioned policy evaluator. Scientific credibility may require domain expertise and judgment. Accredited verification requires an authorised process and scope. Institutional reliance may depend on contract terms, supply-shed boundaries, risk tolerance, reporting purpose, or legal review.

A single status creates a transfer-of-confidence problem. Users may see a cryptographic check and infer scientific validity. They may see a compatibility assessment and infer government approval. They may see one buyer accept an artifact and infer that every future buyer, registry, verifier, or lender must do the same. This paper calls that transfer *validity laundering*: confidence generated at one layer is represented as authority at another.

3.2 Why additional labels are not enough

The problem is not solved by replacing one word with several near-synonyms. A useful model must identify independent state machines, responsible actors, permitted transitions, missing-information states, and prohibited inferences. It must support real ambiguity.

A record may be cryptographically valid but scientifically out of domain. Evidence may be method-compatible but await a VVB determination. A buyer may reject an artifact for contractual reasons even when the evidence is intact and scientifically credible. A receipt may be well formed but impossible to verify because a parent, key, or source is unavailable.

The model must therefore preserve coexistence rather than force all questions into one linear maturity ladder.

3.3 Scientific and technical foundations

Agricultural MRV operates across biological variability, measurement limitations, models, accounting frameworks, and institutional use. Reviews of livestock mitigation and rumen methane show why observed effects, model outputs, production context, and implementation conditions must remain distinguishable (Eckard, Grainger, et al. 2010; Hristov et al. 2013; Beauchemin et al. 2020; Eckard and Clark 2018; Kebreab et al. 2016). Agricultural MRV reviews and inventory or target-setting guidance likewise show that evidence may be interpreted differently for national inventories, intervention assessment, value-chain reporting, and land-sector claims (IPCC 2019; DCCEEW 2023; Paustian et al. 2019; SBTi 2022; WRI 2023).

Data-provenance research provides a basis for separating source lineage and transformation from later judgments (Buneman et al. 2001; Moreau 2010; Lebo et al. 2013). Content addressing and distributed integrity systems motivate tamper-evident commitments, while also illustrating that integrity is narrower than truth or scientific adequacy (Benet 2014; Cachin et al. 2016). Agricultural interoperability literature demonstrates that sensors, farm platforms, models, and institutional systems remain heterogeneous (Wolfert et al. 2017; Villa-Henriksen et al. 2020). The validity model is therefore designed for a plural system in which different actors reach different bounded decisions over a shared evidence graph.

Value-chain and Scope 3 contexts add a further reason to distinguish scientific evidence from reliance and claims. Insetting, corporate target setting, and agricultural data-pipeline work involve contractual allocation, reporting boundaries, and institution-specific use (Tricarico 2021; Liao et al. 2022; SBTi 2022; WRI 2023). A scientifically credible artifact may therefore be rejected for reasons unrelated to the truth of the underlying evidence.

Working-draft note. The references in this working draft were supplied for integration. Bibliographic metadata, especially supplementary entries, must be independently checked against publisher or official records before submission.

3.4 Research questions

1. Can cryptographic integrity, method compatibility, scientific judgment, accredited verification, and institutional reliance be represented without validity laundering?
2. Which states are necessary to preserve missing information, disagreement, out-of-domain conditions, and pending decisions?
3. What transitions are permitted within each layer, and which cross-layer inferences must be prohibited?
4. Can independent reviewers classify realistic agricultural cases reliably using the proposed model?
5. Do prototype dashboards, API responses, evidence packets, and reports display the states without misleading transfers of confidence?

4 Conceptual model

Let E denote an evidence envelope and resolved parent graph, S the associated signature and trust material, M_v a scientific method at version v, P_v an institutional policy adapter at version v, and A an assembled artifact.

4.1 Layer 1: cryptographic integrity

$$\Phi(E, S) \in \{\texttt{valid}, \texttt{invalid}, \texttt{indeterminate}\}.$$

A `valid` result means that the receipt is well formed, canonical commitments match, signatures verify, required parents and trust material are available, and the disclosed bundle is internally consistent under the declared verification profile.

It does *not* mean:

- the physical event occurred;
- the issuer was honest;
- the source was complete;
- the measurement was scientifically accurate;
- the model was appropriate; or

- a later institution should accept the artifact.

`Invalid` means a required integrity condition demonstrably failed. `Indeterminate` preserves uncertainty when trust material, parents, sources, algorithms, or required verification inputs are unavailable or unsupported. The distinction matters because "not proven valid" is not always equivalent to "proven invalid."

4.2 Layer 2: method compatibility

$$\Psi(E, M_v) \in \left\{ \begin{array}{l} \texttt{eligible}, \\ \texttt{eligible\_with\_conditions}, \\ \texttt{outside\_current\_method}, \\ \texttt{method\_extension\_required}, \\ \texttt{insufficient\_evidence}, \\ \texttt{unassigned} \end{array} \right\}.$$

Compatibility maps declared evidence to a specified method and version. It answers a bounded question: given the evidence available and the evaluator encoded in M_v, how does the case map to the method's applicability, evidence, model, and uncertainty requirements?

`eligible` Declared evidence satisfies the specified requirements without unresolved conditions identified by the evaluator.

`eligible_with_conditions` The case may proceed only with named limitations, controls, additional review, or conditional requirements.

`outside_current_method` The declared context is outside the method's defined scope.

`method_extension_required` The case raises a scientifically or operationally plausible extension that is not currently covered.

`insufficient_evidence` The evaluator cannot complete the mapping because required evidence is missing, inaccessible, inconsistent, or inadequate.

`unassigned` No applicable evaluator or method mapping has been selected.

Compatibility is version-specific:

$$D_v = \Psi(E, M_v), \qquad D_{v+1} = \Psi(E, M_{v+1}).$$

The same evidence may produce different outcomes under different versions without contradiction. A change in result must be attributable to changed evidence, changed method, changed model, or changed judgment.

4.3 Scientific judgment and accredited verification

Scientific judgment and accredited verification are explicit evidence-graph events rather than hidden sub-states.

A scientific review event may be represented as

$$J = (r_j, Q, E_J, c, \ell),$$

where r_j is the reviewer and declared competence, Q is the question, E_J is the evidence considered, c is the conclusion, and ℓ contains limitations, conflicts, and dissent.

A verification determination may be represented as

$$V = (b, s_v, A, q_v, d_v, \ell_v),$$

where b is the authorised verification body or professional, s_v is scope and applicable standard or program, A is the artifact reviewed, q_v is assurance or verification profile, d_v is the determination, and ℓ_v contains qualifications and limitations.

Neither event is automatically equivalent to method compatibility or institutional reliance. A method evaluator may require one or both, but the resulting states must remain traceable to the responsible actor.

4.4 Layer 3: institutional reliance

For institution i, purpose p, and artifact A,

$$\Omega_i(A, P_v, p) \in \left\{ \begin{array}{l} \texttt{accepted}, \\ \texttt{relied\_upon}, \\ \texttt{rejected}, \\ \texttt{more\_evidence\_requested}, \\ \texttt{pending}, \\ \texttt{not\_assessed} \end{array} \right\}.$$

Reliance is institution-specific and purpose-specific. A buyer may rely on an artifact for a supplier-engagement decision but not for a public product claim. A lender may accept it for risk screening but require additional evidence for a financial covenant. A verifier may request additional source access. A registry may reject it because a program rule is unmet.

The reliance event must name:

- institution and responsible role;
- artifact and version;
- purpose and decision scope;
- applicable policy or contract;
- outcome and conditions;
- date and validity period;
- evidence requested or excluded; and
- limitations and non-assertions.

4.5 Composite state

The composite state is an ordered tuple rather than a single score:

$$\Gamma = (\Phi(E, S), \Psi(E, M_v), J, V, \Omega_i(A, P_v, p)) .$$

For example:

$$\Gamma = (\texttt{valid}, \texttt{outside\_current\_method}, \texttt{scientifically plausible}, \texttt{not assessed}, \texttt{rejected})$$

is coherent. The record may be intact, the intervention scientifically plausible, outside the current method, not yet verified, and rejected by a particular institution. A scalar "verified/not verified" status would erase these distinctions.

5 Prohibited inference and transition rules

5.1 Prohibited cross-layer inferences

A conforming system must prohibit or explicitly warn against the following inferences:

1. $\Phi = \texttt{valid} \not\Rightarrow$ physical truth.
2. $\Phi = \texttt{valid} \not\Rightarrow \Psi = \texttt{eligible}$.
3. $\Psi = \texttt{eligible} \not\Rightarrow$ scientific consensus.
4. $\Psi = \texttt{eligible} \not\Rightarrow$ accredited verification.
5. $\Psi = \texttt{eligible} \not\Rightarrow$ government approval.

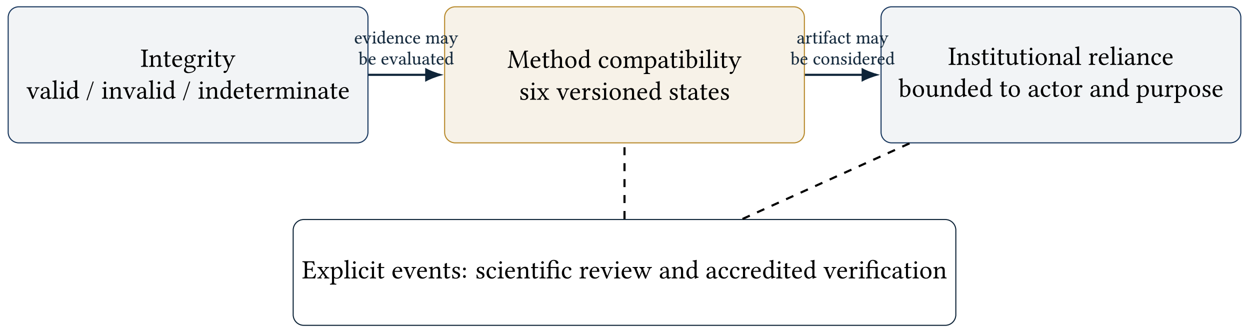

Figure 1: The three layers coexist but do not imply one another. Dashed links indicate that scientific review or accredited verification may inform a decision without being collapsed into it.

6. $J = \texttt{favourable} \not\Rightarrow V = \texttt{positive}$.
7. $V = \texttt{positive} \not\Rightarrow \Omega_i = \texttt{relied\_upon}$ for every institution.
8. $\Omega_i = \texttt{relied\_upon} \not\Rightarrow \Omega_k = \texttt{relied\_upon}$ for $i \neq k$.
9. An asset issuance or retirement event does not by itself establish scientific validity, legal ownership, or absence of duplicate claims elsewhere.

5.2 Within-layer transitions

Integrity transitions require a change in evidence availability, trust material, or demonstrated validity. A receipt may move from `indeterminate` to `valid` when a missing key or parent becomes available. It may move to `invalid` when a mismatch is demonstrated. The original verification event should remain in history.

Compatibility transitions require a new evidence graph, method version, adapter version, or bounded human judgment. A case does not silently become eligible because a UI state was edited.

Reliance transitions are new institutional events. An institution may move from `pending` to `more_evidence_requested`, then to `accepted` or `rejected`. Each event must preserve purpose, date, responsible actor, and artifact version.

5.3 Supersession

No state is overwritten. A later determination supersedes a prior determination by explicit reference. Let x_t be an earlier state event and x_{t+1} a later event. Then

$$x_{t+1} \xrightarrow{\texttt{supersedes}} x_t$$

does not erase x_t; it changes which event is current for a declared context. Historical reports remain reproducible.

6 Interface and API requirements

6.1 Minimum API representation

An API response should expose separate objects rather than one generic status. At minimum it should include:

- integrity state, verifier version, trust basis, and unresolved items;
- method identity, version, compatibility state, conditions, and evidence gaps;
- scientific-review events and declared reviewer roles;
- verification determinations and scope;
- reliance events by institution and purpose;

- timestamps and supersession links;
- limitations and non-assertions; and
- human-readable display labels that preserve the machine semantics.

Null, absent, `not assessed`, and `indeterminate` must not be treated as synonyms. Each represents a different condition.

6.2 Display requirements

Dashboards, reports, and exported artifacts should:

1. display the layers separately;
2. avoid a single green badge that visually dominates all states;
3. place limitations near the positive status they qualify;
4. show exact method, adapter, and artifact versions;
5. display missing evidence and unresolved judgment;
6. name the institution and purpose for reliance;
7. preserve state distinctions in PDF, print, CSV, and API exports; and
8. provide accessible text labels rather than relying only on colour or icons.

6.3 Validity-laundering risk patterns

The interface audit should search for:

- a checkmark labelled "verified" that represents only signature validation;
- scientific or policy claims in tooltips that exceed the underlying state;
- method-compatible evidence displayed beside government logos without authority;
- a buyer acceptance badge presented as universal acceptance;
- missing-evidence warnings hidden behind expandable panels;
- superseded or expired decisions displayed as current;
- colour hierarchies that make conditional or pending states appear final; and
- exported reports that omit limitations shown in the hosted interface.

7 Methods

7.1 Concept analysis

The study will review uses of *verified*, *validated*, *assured*, *approved*, *eligible*, *accepted*, *certified*, *registered*, *issued*, and *audited* across agricultural MRV, assurance, software integrity, environmental accounting, registry documentation, model-risk governance, and institutional reporting.

Each use will be classified by:

- proposition asserted;
- accountable actor;
- evidence required;
- decision consequence;

- duration and scope;
- qualification or limitation;
- possibility of an indeterminate state; and
- likely user inference.

The analysis should distinguish terminological disagreement from proposition disagreement. Two systems may use different labels for the same proposition, or the same label for different propositions.

7.2 Expert elicitation

A structured elicitation will involve scientists, VVB professionals, developers, buyers, processors, government-facing participants, data-governance experts, and producer or implementation representatives. Participants will review the model, identify missing decisions, classify worked cases, and evaluate prohibited inferences.

The panel should be balanced across institutional interests. Conflicts, employment, funding, and relevant commercial relationships will be disclosed. Advisory status alone does not establish authorship.

The elicitation will include:

1. independent pre-discussion classification;
2. facilitated review of disagreements;
3. revision of definitions and cases;
4. independent post-discussion reclassification; and
5. a dissent log that records unresolved objections.

7.3 Case set

At least twelve worked cases will be created. The set should vary:

- integrity state;
- evidence sufficiency;
- model-domain status;
- scientific-review outcome;
- verification status;
- institution and purpose;
- rights or claim conditions;
- missing sources or trust material;
- method and adapter versions; and
- interface presentation.

Cases may be synthetic, permissioned, or adapted from public materials. Synthetic cases must be clearly labelled and must not identify or imply criticism of named companies.

7.4 Inter-rater classification

Reviewers will classify each case by layer, state, responsible actor, missing evidence, permitted conclusion, and prohibited inference. Primary outcomes include percentage agreement, agreement by layer and state, confusion matrices, and qualitative disagreement categories.

For two reviewers, Cohen's κ may be reported:

$$\kappa = \frac{p_o - p_e}{1 - p_e},$$

where p_o is observed agreement and p_e is chance-expected agreement. For more than two reviewers or ordinal judgments, an appropriate alternative will be preregistered.

Agreement must not be interpreted as proof that the model is correct. Training may increase agreement without improving external validity. The study should report which distinctions remain difficult even after clarification.

7.5 Interface and report audit

Prototype dashboards, API responses, evidence packets, and verifier reports will be examined for misleading transfers of confidence. The audit will test:

- colour, icon, and badge semantics;
- labels and plain-language definitions;
- hierarchy and visual prominence;
- tooltip and disclosure proximity;
- export consistency;
- supersession and expiry display;
- accessibility for users who do not perceive colour; and
- user interpretation under task-based scenarios.

A structured rubric will score whether the interface states the bounded proposition, actor, version, purpose, limitations, and unresolved evidence.

7.6 Conformance cases

Machine-readable conformance cases will specify:

- input evidence and trust state;
- method and policy versions;
- expected layer states;
- permitted and prohibited display language;
- expected API fields;
- required warnings;
- supersession behaviour; and
- rationale.

A conforming implementation must produce the expected state separation even if its visual design differs.

8 Worked cases

8.1 Valid receipt, out-of-domain intervention

A signed feed-delivery record is intact and its parents resolve. The animal cohort, however, falls outside the calibration domain of the selected model. The record remains useful evidence. The appropriate composite state is valid integrity, outside-current-method or method-extension-required compatibility, no verification determination, and no final reliance.

8.2 Method-compatible record, pending verification

A complete bundle satisfies declared method inputs and deterministic rules but awaits VVB review. Compatibility and pending verification must be displayed separately. The system must not use "verified" merely because the method evaluator passed.

8.3 Institutional rejection despite scientific credibility

A buyer rejects an artifact because the project is outside its supply shed, the claim is allocated elsewhere, or a contract condition is unmet. The rejection does not make the scientific evidence false and does not invalidate the cryptographic bundle.

8.4 Cryptographic indeterminacy

A receipt is well formed but cannot be verified because a key, parent, or source is unavailable. The system reports `indeterminate`. It does not fabricate a binary result.

8.5 Valid signature, unauthorised issuer

A signature verifies, but the issuer's authority receipt does not cover the event type or time period. Integrity may be valid while authority is insufficient. A method evaluator may therefore return insufficient evidence or a condition requiring authorised confirmation.

8.6 Favourable scientific review, method extension required

A domain expert judges an intervention scientifically plausible and the evidence informative, but the current method does not cover the production context. The scientific review is recorded positively while compatibility remains method-extension-required.

8.7 Positive verification, buyer requests more evidence

A VVB reaches a positive bounded determination under its engagement. A buyer nevertheless requests additional chain-of-custody or claim-allocation evidence. The buyer's request does not negate the verification outcome.

8.8 Two institutions, different reliance outcomes

A processor relies on an artifact for internal supplier engagement. A lender rejects the same artifact for covenant reporting because its policy requires a different assurance level. Both reliance events are valid and institution-specific.

8.9 Method supersession

Evidence is compatible under method version v but becomes conditional under version $v + 1$ because uncertainty treatment changes. Both determinations remain linked to the same evidence root and exact method versions.

8.10 Expired reliance

An institution previously relied on an artifact for a twelve-month decision period. The reliance event has expired. The interface must not display it as current, although the historical event remains visible.

8.11 Conflicting scientific reviews

Two qualified reviewers reach different conclusions about model applicability. The system records both reviews, their evidence sets, competence declarations, conflicts, and limitations. It must not silently average them into a single "verified" state.

8.12 Artifact export drops limitations

The hosted dashboard displays conditional eligibility and limitations, but a PDF export shows only a green badge. The backend states may be correct, yet the export fails interface conformance because it launders a conditional result into an apparently final one.

9 Planned results reporting

Working-draft note. The draft reports no completed elicitation, inter-rater agreement, or interface audit. The final paper must include the case set, participant characteristics, agreement measures, dissent, interface findings, and model revisions.

Table 2: Paper 2 planned results template.

Analysis	Result	Required reporting
Concept inventory	TBD	Terms, sources, propositions, actors, consequences, and ambiguity categories.
Expert panel	TBD	Recruitment, expertise, institutional balance, conflicts, attrition, and dissent.
Inter-rater classification	TBD	Agreement by layer and state, confusion matrix, disagreement analysis, and changes after discussion.
Interface audit	TBD	Number and type of validity-laundering risks by medium and severity.
User interpretation	TBD	Task performance, misinterpretation patterns, confidence, and accessibility findings.
Expert revision	TBD	Accepted changes, rejected changes, unresolved dissent, and version history.
Conformance cases	TBD	Expected API and display behaviour, implementation deviations, and reasons.
Independent reproduction	TBD	Environment, time, missing dependencies, deterministic agreement, and unresolved judgment.

10 Discussion

A tripartite model changes the design objective from producing a reassuring status to exposing accountable decisions. It allows systems to be precise without claiming certainty. Institutions can ask whether a platform exposes integrity, compatibility, scientific review, verification, and reliance separately; whether states are versioned; whether unresolved evidence remains visible; and whether positive outcomes trace to responsible actors.

The framework should not become a semantic shield for weak science. A platform can display separate states and still rely on poor models, incomplete evidence, conflicted review, or inadequate capture controls. The state model improves transparency but does not replace substantive scientific evaluation. Cryptographic integrity is only as useful as the authority and capture controls surrounding the original assertion.

The model also resists a simplistic maturity ladder. Integrity is not a lower-quality version of verification, and institutional reliance is not a universally higher state than scientific review. The layers answer different questions. A buyer's rejection may be commercially decisive for that buyer but scientifically irrelevant. A scientific review may be favourable but insufficient for accredited verification. A record may remain valuable even when it is outside a current method.

Versioning is central. Method compatibility without method identity and version is not reproducible. Reliance without institution, purpose, and artifact version is not bounded. Verification without scope and applicable criteria is ambiguous. Interfaces should therefore make version metadata visible rather than treating it as technical detail.

The approach may also improve dispute handling. When actors disagree, they can identify whether the disagreement concerns integrity, evidence sufficiency, model domain, scientific judgment, verification procedure, contract terms, rights, or institutional risk appetite. This is more actionable than arguing over whether something is simply "verified."

11 Limitations

The vocabulary may not align with every registry, standards body, VVB, assurance provider, or jurisdiction. Mapping studies will be required, and some institutions may need additional layers such as legal opinion, registry issuance, accreditation level, or financial audit.

Expert elicitation may be biased by participant selection, facilitation, funding, or commercial interests. Inter-rater agreement may reflect training rather than intuitive clarity or real-world usability.

The proposed state sets may be too coarse for some applications and too complex for others. Additional sub-states may be needed, but excessive granularity could reproduce the ambiguity the model seeks to reduce.

Interface findings from prototypes may not generalise to production systems or users under time pressure. Colour, icon, language, and hierarchy effects may vary by culture, literacy, role, and accessibility needs.

Cryptographic integrity depends on capture controls, key management, identity, and authority. A valid signature can faithfully preserve a false or unauthorised assertion.

Method compatibility may include human judgment. Deterministic software cannot eliminate ambiguity where a method itself is open to interpretation.

Institutional reliance can be confidential, conditional, or legally sensitive. Public evidence graphs may not contain every reason for acceptance or rejection.

Finally, publication of the state model does not confer government acceptance, approve a methodology, create an Australian Standard, replace accredited validation or verification, determine legal rights, or establish claim ownership.

12 Conclusion

Agricultural MRV systems should stop using "verified" as an undifferentiated endpoint. Cryptographic integrity, method compatibility, scientific judgment, accredited verification, and institutional reliance are different propositions with different evidence, actors, consequences, and limitations.

The proposed model represents integrity, compatibility, and reliance as separate state machines while preserving scientific review and verification as explicit events. It defines indeterminate and insufficient-evidence states, requires exact version commitments, and makes prohibited inference a conformance issue rather than a marketing-style choice.

The next step is empirical evaluation: complete the terminology map, recruit a balanced expert panel, preregister the classification and interface studies, create at least twelve worked cases, publish the rubric and conformance fixtures, report disagreement and failure, and revise the model through versioned change control.

13 Declarations

Funding, conflicts, and ethics

TODO: Collect disclosures for AgEvidence support, employment, consulting, advisory roles, equity, intellectual property, industry relationships, and other interests. Determine ethics requirements for expert elicitation, interface studies, and any participant data.

Author contributions

TODO: Maintain a live CRediT contribution record. Confirm author order and corresponding-author responsibility only after written consent and final contribution review.

Cases and software

TODO: Publish synthetic cases, classification rubric, state definitions, API schema, UI test materials, conformance fixtures, and environment manifests where rights permit.

Data availability

TODO: Publish de-identified ratings, questionnaires, analysis code, dissent logs, and interface-audit outputs where ethics, privacy, and contractual conditions permit.

Synthetic-case disclosure

Synthetic cases are research artifacts designed to test classification and interface behaviour. They are not claims about named companies, projects, methods, or institutions.

14 Executable TODO register

Table 3: Paper 2 executable TODO register.

Task	Owner	Exit evidence
Complete assurance and terminology literature map	Persona C	Search protocol, concept inventory, source table, and ambiguity taxonomy.
Freeze states after Paper 1 terminology	Persona G; approval by authors	Versioned state specification, mapping, and decision log.
Preregister expert and interface studies	Methods team	Public or archived protocol, outcomes, exclusions, and analysis plan.
Recruit balanced experts	Persona C and Program Manager	Recruitment record, expertise matrix, conflicts, and consent.
Create at least twelve worked cases	Authors and technical team	Case set, answer key, rationale, and synthetic labels.
Run inter-rater study	Methods team	De-identified classifications, agreement analysis, and disagreement log.
Run interface and report audit	Governance and HCI reviewers	Audit rubric, screenshots, issue register, and severity ratings.
Publish API and UI conformance fixtures	Persona G and reproducibility engineer	Machine-readable fixtures, expected outputs, and environment manifest.
Run independent reproduction	Independent reviewers	Reproduction report, time, gaps, deviations, and unresolved issues.
Report dissent and version all changes	Persona C and scientific editor	Version history, accepted/rejected changes, and dissent appendix.
Complete CRediT, disclosures, and final approval	Persona C	Signed authorship, affiliation, conflict, and approval records.

References

Beauchemin, K. A., E. M. Ungerfeld, R. J. Eckard, and M. Wang (2020). “Review: Fifty years of research on rumen methanogenesis: What have we learned and where are we heading?” In: *Animal* 14.s1, s22–s36.

Benet, J. (2014). “IPFS - Content Addressed, Versioned, P2P File System”. In: *arXiv preprint arXiv:1407.3561*.

Buneman, P., S. Khanna, and W.-C. Tan (2001). "Why and Where: A Characterization of Data Provenance". In: *Database Theory – ICDT 2001*. Vol. 1973. Lecture Notes in Computer Science. Springer, pp. 316–330.

Cachin, C., M. Vukolić, and K. Kursawe (2016). "Blockchains and Consensus Protocols in the Wild". In: *IEEE Security & Privacy* 14.4, pp. 28–30.

Department of Climate Change, Energy, the Environment and Water (2023). *National Inventory Report 2021: Volume 1*. Australian Government.

Eckard, R. J. and H. Clark (2018). "Climate Change Mitigation in Livestock Production Systems". In: *Animal Production Science* 58.8, pp. 1379–1387.

Eckard, R. J., C. Grainger, and C. A. M. de Klein (2010). "Options for the Mitigation of Methane and Nitrous Oxide Emissions from Ruminant Production Systems". In: *Agriculture, Ecosystems & Environment* 136.3-4, pp. 209–216.

Hristov, A. N., J. Oh, C. Lee, R. Meinen, F. Montes, T. Ott, J. Firkins, A. Rotz, C. Dell, A. Adesogan, W. Yang, J. Tricarico, E. Kebreab, G. Waghorn, J. Dijkstra, and S. Oosting (2013). *Mitigation of Greenhouse Gas Emissions in Livestock Production – A Review of Technical Options for Non-CO2 Emissions*. FAO Animal Production and Health Paper 177. Rome: Food and Agriculture Organization of the United Nations.

Intergovernmental Panel on Climate Change (2019). *2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories*. Switzerland: IPCC.

Kebreab, E., A. Liedke, D. Caro, S. Deimling, C. E. van Middelaar, and J. Dijkstra (2016). "Environmental Impact of Livestock Production and Options for Mitigation". In: *Journal of Animal Science* 94.8, pp. 3121–3131.

Lebo, T., S. Sahoo, D. McGuinness, K. Belhajjame, J. Cheney, D. Corsar, D. Garijo, S. Soiland-Reyes, S. Zednik, and J. Zhao (2013). *PROV-DM: The PROV Data Model*. W3C Recommendation. World Wide Web Consortium.

Liao, W., K. Johnson, and L. Thompson (2022). "Integrated Data Pipelines for Agricultural Footprinting and Scope 3 Reporting". In: *Journal of Cleaner Production* 365, p. 132710.

Moreau, L. (2010). "The Foundations for Provenance on the Web". In: *Foundations and Trends in Web Science* 2.2-3, pp. 99–241.

Paustian, K., S. Collier, J. Baldock, S. Ogle, M. Eve, S. Del Grosso, and G. P. Robertson (2019). "Quantifying Greenhouse Gas Mitigation Potential in Agriculture: A Review of MRV Tools and Paradigms". In: *Environmental Research Letters* 14.12, p. 123001.

Science Based Targets initiative (2022). *Forest, Land and Agriculture (FLAG) Target Setting Guidance*. Science Based Targets initiative.

Tricarico, J. M. (2021). "Carbon Insetting: Opportunities for the Dairy Sector to Reduce Emissions Along the Supply Chain". In: *Journal of Dairy Science* 104.E-Suppl. 1, pp. 45–52.

Villa-Henriksen, A., G. T. Edwards, L. A. Pesonen, O. Green, and C. A. G. Sørensen (2020). "Internet of Things in Agriculture: A Survey on Smart Farming and Challenges for Technical Interoperability". In: *Biosystems Engineering* 191, pp. 62–80.

Wolfert, S., L. Ge, C. Verdouw, and M.-J. Bogaardt (2017). "Big Data in Smart Farming – A Review". In: *Agricultural Systems* 153, pp. 69–80.

World Resources Institute (2023). *Land Sector and Removals Guidance: Draft for Pilot Testing*. GHG Protocol, WRI & WBCSD.

Source note

This expanded draft is derived from Paper 2 and the portfolio guardrails, ownership tables, sequencing rules, and disclosure language in the internal working document *AgEvidence Australia Scientific Publication Series: Seven Manuscript First Drafts*, Version 0.1, August 2026. The scientific and technical literature listed in the bibliography was supplied for integration and has been cited where relevant. Bibliographic accuracy and source suitability must be independently verified before submission.

Paper 3: Minimum Evidence Requirements for Enteric Methane Interventions in Australian Beef and Dairy Systems

Working-draft note. This manuscript expands Paper 3 from the *AgEvidence Australia Scientific Publication Series: Seven Manuscript First Drafts*. It is a review, consensus, and conformance-framework manuscript. It does not report a completed systematic review, Delphi process, field trial, conformance study, verifier determination, emissions reduction, product ranking, or institutional acceptance. All study yields, ratings, thresholds, case outcomes, and burden estimates remain planned outputs.

Non-negotiable guardrail. A minimum-evidence framework defines what must be available for a bounded scientific, method, verification, or institutional decision. It does not create an Australian Standard, establish an approved methodology, certify a commercial product, determine legal rights or claim ownership, replace an accredited validation or verification body, or predetermine acceptance by a government agency, method authority, registry, buyer, lender, processor, or other institution.

Abstract

Rapid development of enteric-methane interventions has created a recurring translation problem: evidence sufficient to demonstrate efficacy in a controlled study may not be sufficient to support repeatable implementation, quantification, verification, or value-chain claims. This paper proposes a minimum-evidence and conformance framework for enteric-methane interventions in Australian beef and dairy systems. The study combines a preregistered systematic or scoping review, structured expert consensus, evidence-to-decision mapping, production-context testing, and machine-readable conformance cases.

The framework covers intervention identity and product characterisation; animal and cohort context; production system and diet; baseline definition; prescribed and delivered dose; observed exposure; duration and interruptions; methane measurement; model execution; animal performance and health; adverse effects; persistence; operational delivery; authority; rights; limitations; and intended decision use. Evidence is classified by source, temporal relationship, authority, transformation status, uncertainty, and decision purpose. The framework separates observations from calculations, assumptions, scientific judgments, method-compatibility assessments, accredited validation or verification determinations, and institution-specific reliance.

Measurement requirements are grounded in the distinct error structures and operating domains of respiration chambers, sulfur hexafluoride tracer methods, automated spot-sampling systems, and model-based estimates. Intervention-specific requirements preserve formulation, active component, batch, stability, diet interaction, dose-response, adaptation, and persistence. Australian pasture-based beef, feedlot beef, and dairy systems are treated separately where delivery fidelity, animal tracking, intake measurement, recording burden, and calibration-domain applicability differ materially.

The planned outputs are a minimum dataset, controlled vocabulary, gap taxonomy, decision-linked requirement matrix, production-context profiles, conformance checklist, and machine-readable schema. These outputs are intended to support transparent evidence readiness and identify when additional evidence, conditions, independent review, or method extension is required. They do not establish official Australian requirements, rank commercial interventions, or certify emissions reductions.

Keywords: enteric methane; evidence sufficiency; feed additives; beef; dairy; Australia; measurement provenance; model provenance; implementation fidelity; conformance; uncertainty.

1 Manuscript status, scope, and contribution boundary

This manuscript defines a proposed evidence framework and a planned process for evaluating it. It does not assert that the proposed fields are already agreed minimum requirements. It does not report review search yields, study-quality ratings, consensus scores, stakeholder agreement, inter-rater reliability, producer burden, or conformance performance. Tables labelled as planned results are reporting templates rather than findings.

The paper depends on Paper 1 for the stable receipt envelope, append-only evidence graph, source commitments, parent relationships, versioned adapters, and bounded reliance-artifact concepts. It depends on Paper 2 for the separation of cryptographic integrity, method compatibility, scientific judgment, accredited verification, and institutional reliance. Paper 3 contributes the scientific and operational evidence vocabulary for enteric-methane interventions. Later applied papers should reuse this vocabulary rather than creating independent meanings for intervention identity, dose, exposure, measurement provenance, model execution, persistence, evidence sufficiency, or decision readiness.

The unit of analysis is not a product in the abstract. It is a declared intervention implementation involving a defined product or practice, animal cohort, production context, period, delivery pathway, measurement or model pathway, evidence bundle, and decision purpose. The same commercial product may therefore produce different sufficiency outcomes across diets, animal classes, systems, durations, methods, and institutional uses.

The framework is intended to answer a bounded question:

> Given a declared intervention implementation and decision purpose, what evidence must be present, attributable, temporally aligned, scientifically applicable, and reproducible for a competent reviewer to reach a transparent decision without silently reconstructing material facts?

It is not intended to define one universal dataset for every purpose. Scientific efficacy, implementation fidelity, model quantification, protocol compatibility, verification readiness, payment, and Scope 3 reliance require overlapping but non-identical evidence. Proportionality therefore requires each field to be linked to a decision use.

2 Provisional scientific author personas

Persona	Provisional role	Contribution domain / deliverable
Persona D	Lead/corresponding-author persona	Review protocol; measurement quality; uncertainty; evidence sufficiency; consensus methodology and scientific terminology.
Persona B	Scientific coauthor persona	Intervention efficacy, dose-response, diet/intake interactions, persistence and model calibration domains.
Persona A	Scientific coauthor persona	Australian pasture, feedlot and dairy production context; applicability and local-method boundaries.
Persona F	Scientific coauthor persona	Producer implementation, delivery mechanisms, cohort continuity, recording burden and proportionality.
Persona C	Scientific coauthor persona	Buyer, processor, program, protocol and Scope 3 evidence-readiness translation.
Persona G	Technical scientific contributor persona	Machine-readable schema, gap states, conformance tooling, synthetic cases and reproducibility fixtures.
Persona I	Invited domain-review persona	Measurement, method, verification, producer or Australian domain expertise where qualifying.

Authorship control. Persona IDs are recruitment placeholders only. A real person is not an author until consent, qualifying contribution, affiliation, conflict disclosure, CRediT allocation, and final-manuscript approval are complete.

3 Introduction

3.1 The efficacy-to-evidence-readiness gap

The scientific evaluation of an enteric-methane intervention and the institutional use of that intervention are related but different problems. Controlled trials may establish a treatment effect under defined diets, animal classes, doses, durations, measurement systems, and environmental conditions. Commercial implementation adds product identity, batch variation, storage, purchase and delivery, administration, animal-cohort continuity, operational deviations, model execution, record retention, verification scope, rights, and claim allocation.

Evidence gaps frequently emerge during translation rather than efficacy research. A trial may report a treatment effect without preserving identifiers needed to connect the research formulation to a later commercial batch. A producer may document purchase but not administration. A feedlot may record ration inclusion but not pen-level refusals or substitutions.

A grazing enterprise may document supplement delivery without credible animal-level or cohort-level exposure. A model may estimate reductions without preserving the software version, input transformations, calibration domain, or uncertainty treatment. A buyer may seek a value-chain claim without evidence of contractual allocation, overlap, or the boundary of the relied-on outcome.

The resulting problem resembles a bridge whose laboratory-tested material strength is well documented while its field assembly records are missing. The strength test is still scientifically valuable, but it cannot by itself demonstrate that the installed bridge used the same material, in the same configuration, under the same loads. Enteric-methane evidence requires the equivalent of both material testing and construction records: efficacy evidence must remain distinguishable from evidence that the declared intervention was actually delivered to the declared animals, for the declared duration, under conditions in which the quantified effect is scientifically applicable.

Rigorous evidence synthesis must therefore draw upon established measurement methodologies [7], Australian production-system parameters [3, 5], and empirical model calibration-domain boundaries [4]. While additives such as 3-nitrooxypropanol and *Asparagopsis taxiformis* demonstrate material efficacy under controlled trials [9, 4], operational delivery and cohort-binding integrity remain major friction points in extensive grazing systems [12].

3.2 Why minimum evidence is not a minimum efficacy threshold

This paper proposes a minimum-evidence framework rather than a minimum efficacy threshold. An efficacy threshold asks how large, consistent, or durable an effect must be. A minimum-evidence framework asks what information must be available for a competent reviewer to evaluate an implementation and determine what conclusions are supported.

The distinction matters because a small, well-characterised effect may be more decision-ready than a large but poorly documented effect. Conversely, extensive evidence does not guarantee efficacy. Evidence sufficiency is a property of the decision pathway: whether identity, context, delivery, measurement, model, uncertainty, authority, and limitations are adequately documented for the stated purpose.

The framework therefore avoids a single binary label. It represents evidence at field and domain level using states such as `present`, `missing`, `conflicting`, `out_of_domain`, `conditionally_sufficient`, and `not_applicable`.A final decision remains the responsibility of the named scientific, method, verification, or institutional actor.

3.3 Australian production systems require context-specific sufficiency

Australian pasture-based beef, feedlot beef, and dairy systems differ in diet control, animal movement, supplementation, records, performance data, measurement opportunities, and model applicability. A field that can be captured automatically in a total mixed ration system may require proxy evidence or remain unavailable in extensive grazing. A delivery method that produces relatively uniform exposure in a controlled feed bunk may produce highly variable exposure through blocks, lick supplements, water medication, or intermittent mustering.

The framework therefore combines stable evidence domains with production-context profiles. Stable domains preserve the questions that must be considered; profiles define which sources, frequencies, tolerances, proxies, and exceptions may be credible in a given system. Context specificity is not a relaxation of scientific rigor. It is an explicit statement that rigor must be achieved through methods appropriate to the operating environment rather than through fictional uniformity.

3.4 Evidence continuity from source to decision

The framework treats evidence continuity as a chain. Product characterisation must connect to a batch. The batch must connect to purchase and delivery. Delivery must connect to administration. Administration must connect to animals or cohorts. Cohorts must connect to production context and time. Measurements or model inputs must connect to those same animals, periods, and conditions. Calculations must connect to declared methods and versions. Scientific judgments, verification determinations, and institutional reliance must identify the artifact and scope on which they depend.

A broken link does not always invalidate every downstream use. It changes what can be concluded. Purchase without administration may support a commercial transaction but not implementation fidelity. Model execution without calibration-domain evidence may support a reproducibility check but not scientific applicability. A positive scientific review may support credibility but not legal claim allocation. The framework is designed to preserve these distinctions.

3.5 Research questions

The manuscript addresses six primary questions:

1. Which evidence domains are necessary to evaluate enteric-methane interventions across efficacy, implementation fidelity, quantification, method compatibility, verification readiness, and institutional use?
2. Which fields must be captured at source because later reconstruction would be unreliable, unverifiable, or vulnerable to bias?
3. How should minimum evidence differ among Australian pasture-based beef, feedlot beef, and dairy systems?
4. How should requirements differ among intervention classes, including 3-nitrooxypropanol, *Asparagopsis* and other halogenated macroalgae, nitrate, lipids, and management or dietary interventions?
5. Can experts classify realistic cases consistently using the proposed vocabulary, gap states, and decision-linked requirement matrix?
6. Can a machine-readable schema identify missing, conflicting, out-of-domain, conditional, and not-applicable evidence without converting those states into unsupported approval?

3.6 Novelty relative to efficacy-only and checklist approaches

Table 2: Novelty analysis for the proposed minimum-evidence framework.

Dimension	Common efficacy-only or checklist approach	Proposed framework
Evidence object	Treats the published intervention effect as the primary transferable object.	Treats a declared implementation, evidence bundle, and decision purpose as the unit of analysis.
Identity	Uses intervention class or product name without preserving formulation, active component, batch, stability, and substitution.	Requires identity continuity from research formulation to commercial product, batch, storage, and administration.
Dose	Conflates offered, formulated, delivered, and consumed dose.	Separates prescribed, manufactured, delivered, available, and observed exposure.
Context	Records broad animal category or diet description.	Binds cohort, animal class, production stage, diet, management, location, baseline, and material changes.
Measurement	Reports a methane value without complete provenance.	Preserves method, device, calibration, sampling design, operator, exclusions, raw/processed status, and uncertainty.
Models	Treats successful execution as evidence of applicability.	Requires version, code, inputs, transformations, factors, calibration domain, assumptions, sensitivity, and review state.
Persistence	Extrapolates short trials without explicit temporal qualification.	Preserves adaptation, discontinuation, rebound, duration, carryover, and persistence uncertainty.
Decision linkage	Requires fields "because they may be useful."	Links every required field to a named decision use and challenges fields without one.
Gap handling	Uses pass/fail or complete/incomplete.	Preserves `missing`, `conflicting`, `out_of_domain`, `conditionally_sufficient`, and `not_applicable`.
Authority	Treats evidence possession as authority to assert or allocate outcomes.	Records issuer role, authority basis, rights, consent, claim allocation, and institutional purpose separately.

4 Scientific and methodological foundations

4.1 Measurement methodologies and provenance

Enteric-methane measurement methods do not produce interchangeable observations merely because they share a unit. Respiration chambers provide high control and continuous capture but can alter normal behaviour, diet access, and environment. Sulfur hexafluoride tracer techniques can support measurements under less constrained conditions but depend on tracer release, sampling, background correction, equipment integrity, and operator practice. Automated head-chamber or spot-sampling systems can collect repeated observations at voluntary visits but require adequate visit frequency, temporal coverage, calibration, and attention to selection effects. Reviews comparing these methods emphasise method-specific uncertainty, calibration, sampling, and applicability [7].

Automated measurement can expose substantial between-animal and within-animal variation. Repeated observations are therefore not simply redundant copies; they are samples from a temporally varying biological process. Work measuring dairy cows during milking illustrates the importance of individual variation and sampling design [6]. Minimum evidence must consequently include not only the device name but the sampling frame, number and timing of observations, exclusions, missingness, calibration history, operator interventions, and the relationship between observed animals and the declared cohort.

Measurement provenance should permit a reviewer to answer five questions:

1. What physical quantity was observed, by what method and device?
2. Under what calibration and quality-control regime was it observed?
3. Which animals, periods, diets, and intervention conditions did the observation represent?
4. What processing, filtering, exclusion, aggregation, and unit conversion occurred?
5. What uncertainty and limitations apply to the resulting estimate?

4.2 Intervention identity, dose-response, and persistence

Intervention classes differ in mechanism, formulation sensitivity, delivery requirements, interaction with diet, safety profile, and persistence. The framework must preserve those differences rather than reducing all interventions to “feed additive present.”

For 3-nitrooxypropanol, meta-analytic evidence indicates that antimethanogenic effects vary with dietary composition [4]. The minimum evidence profile should therefore preserve diet composition, dry-matter intake, inclusion rate, feeding frequency, production stage, and the domain in which the selected effect relationship was derived. A claimed dose cannot be interpreted independently of diet and intake.

For *Asparagopsis taxiformis*, intervention identity includes species, cultivation and processing conditions, formulation, active-component concentration, batch stability, storage, and delivery. Dose-response evidence and the volatility or stability of active components make product and batch continuity particularly important [9]. Evidence associated with one preparation cannot silently transfer to another preparation whose active concentration or storage history differs.

For lipid supplementation, the response may be non-linear and constrained by intake, diet, digestibility, animal performance, and practical inclusion limits. Meta-analysis in beef cattle supports the need to preserve supplement type, fatty-acid profile where material, inclusion, basal diet, intake, duration, and co-effects [1].

For nitrate and alternative electron acceptors, evidence must include dose escalation, adaptation, diet, administration, safety monitoring, and adverse-event handling. Meta-analysis provides a basis for effect parameterisation but does not eliminate implementation-specific safety and exposure requirements [2].

Persistence is not a single field. It includes duration of exposure, duration of observed effect, adaptation during treatment, effect after interruption, rebound after discontinuation, carryover, and uncertainty when trial duration is shorter than the intended claim period.

4.3 Model calibration and reproducible execution

Model-based quantification is a sequence of declared transformations, not a black box. Tiered inventory and agricultural emissions work demonstrates that model outputs depend on input quality, parameterisation, system boundaries, emission factors, and assumptions [8]. Australian forage-fed cattle equations further illustrate the importance of dry-matter intake and the empirical population from which a relationship was derived [3].

Minimum model provenance includes:

- model name, owner or source, version, release date, code or executable commitment;
- declared purpose and output quantity;
- all inputs, units, sources, timestamps, and missing-data treatments;
- transformations, lookup factors, coefficients, and rounding rules;
- calibration population and domain;
- applicability assessment for animal, diet, production system, climate, and intervention;
- uncertainty method, parameter distributions, sensitivity analysis, and covariance assumptions where material;
- exclusions, overrides, reviewer interventions, and software environment;
- output artifact, execution timestamp, and reproducibility status.

A successful execution establishes that declared software processed declared inputs. It does not by itself establish that the selected model is scientifically applicable. Compatibility therefore requires both execution provenance and domain evidence.

4.4 Australian production-system feasibility

Australian mitigation reviews emphasise that technical potential and operational feasibility differ across grazing and intensive systems [5]. Forage-fed cattle models reflect Australian production data but remain conditioned on the populations, intake estimates, and circumstances used in their development [3]. Grazing systems introduce particular challenges for daily supplementation, exposure uniformity, intake estimation, animal movement, and cohort continuity [12].

The framework therefore treats feasibility evidence as substantive rather than administrative. A scientifically effective compound that cannot be delivered consistently under the declared management system may have strong efficacy evidence and weak implementation evidence. The correct response is not to discard the science or to assume delivery. It is to report the gap and bound the supported conclusion.

4.5 Evidence synthesis and decision translation

Systematic reviews and meta-analyses provide structured ways to evaluate heterogeneous mitigation evidence, identify moderators, and separate broad patterns from individual studies [10]. Their role in this paper is procedural and substantive: the review identifies candidate evidence fields, intervention-specific moderators, measurement limitations, model domains, and unresolved uncertainty.

Evidence-to-decision translation also requires attention to boundaries and leakage. Land-use mitigation analysis illustrates why apparently favourable changes can be misinterpreted when system consequences and displaced effects are omitted [11]. For enteric interventions, analogous boundary questions include reduced intake, altered productivity, herd-size effects, upstream intervention production, manure consequences, substitution, and claim overlap. Paper 3 does not perform a full life-cycle assessment, but it requires declared decision boundaries and material co-effects where they influence interpretation.

> **Working-draft note.** The academic references integrated into this draft were supplied for incorporation. Bibliographic metadata and all interpretive claims must be independently checked against publisher records before submission. The final review should add current primary studies, regulatory and method documents where relevant, safety literature, and Australian system-specific evidence.

5 Conceptual framework

5.1 Evidence object and decision context

Let X denote a declared enteric-methane intervention implementation:

$$X = \langle P, B, C, S, T, D, Q, Y, U, R \rangle,$$

where P is intervention and product identity, B is batch and supply-chain identity, C is animal or cohort context, S is production system and diet, T is the intervention period, D is delivery and administration, Q is the measurement or model pathway, Y is performance and health evidence, U is uncertainty and limitations, and R is authority, rights, and intended decision use.

Let d be a bounded decision purpose, such as scientific efficacy assessment, implementation-fidelity assessment, model quantification, method compatibility, verification readiness, institutional reliance, claim allocation, or payment.

For evidence field f, define:

$$\begin{aligned}\sigma(f, X, d) \in \{&\texttt{present}, \texttt{missing}, \texttt{conflicting}, \\ &\texttt{out\_of\_domain}, \texttt{conditionally\_sufficient}, \texttt{not\_applicable}\}.\end{aligned}$$

The state does not represent universal truth. It records the result of evaluating a declared field against a versioned requirement for a declared implementation and purpose.

The domain-level result is:

$$\Sigma_k(X, d, v) = \langle \sigma(f_1, X, d), \ldots, \sigma(f_n, X, d) \rangle,$$

where k identifies the evidence domain and v identifies the framework version. No scalar confidence score is assumed.

5.2 Evidence classes

Each item should identify its epistemic class:

1. **Observation:** a direct or instrument-mediated record of an event or condition.
2. **Commercial or administrative record:** a transaction, custody, delivery, authorisation, or contractual record.
3. **Calculation:** a deterministic transformation of declared inputs.
4. **Model output:** an output produced under a specified model, calibration domain, and assumptions.
5. **Assumption or proxy:** a declared substitute used because direct evidence is unavailable.
6. **Scientific judgment:** an expert conclusion about applicability, sufficiency, uncertainty, or interpretation.
7. **Method-compatibility determination:** an evaluation against a specified method and version.
8. **Validation or verification determination:** a conclusion reached by an authorised body under a declared scope.
9. **Institutional reliance event:** a named institution's bounded decision to accept, rely, reject, or request more evidence.

These classes should not be collapsed. A model output is not an observation. A purchase record is not an administration record. A scientific judgment is not an accredited verification determination. A buyer's reliance is not universal acceptance.

5.3 Source-capture priority

A field has high source-capture priority when retrospective reconstruction would create material uncertainty, bias, or unverifiable inference. Candidate high-priority fields include product and batch identity; storage and expiry; delivery timestamp and quantity; ration formulation; administration events; interruption and substitution; cohort membership; device identifiers; calibration records; raw observations; model version; input values; transformations; and contemporaneous limitations.

The consensus process should rate each field on:

$$\text{Priority}(f) = g(N_f, R_f, T_f, A_f, F_f, B_f),$$

where N_f is decision necessity, R_f is reconstruction risk, T_f is temporal sensitivity, A_f is attribution importance, F_f is capture feasibility, and B_f is burden. The function g will be specified prospectively; this draft does not assign weights.

6 Minimum evidence domains

6.1 Intervention identity and product characterisation

The evidence chain should identify intervention class, product name, manufacturer or responsible supplier, formulation, active component where applicable, concentration, physical form, batch or lot, manufacture date, authorisation or legal status where relevant to the decision, storage conditions, expiry, certificate of analysis, substitution history, and the relationship between the commercial product and the research evidence cited.

Identity continuity prevents evidence associated with one product, species, strain, formulation, active concentration, or processing method from transferring silently to another. For biologically or chemically variable products, a generic class label is insufficient. Batch-level characterisation may be required when efficacy or safety depends on active concentration, stability, contamination, or degradation.

Minimum evidence should distinguish:

- **class identity**, such as inhibitor, macroalgae, nitrate, lipid, vaccine, breeding strategy, or management intervention;
- **commercial identity**, including product, supplier, formulation, and version;
- **batch identity**, including lot, manufacture, analysis, storage, and expiry;
- **research linkage**, identifying which trials or evidence profiles are considered applicable and why;
- **substitution**, including any change in formulation, supplier, active concentration, carrier, or delivery form.

6.2 Animal, herd, and cohort context

Fields include species, breed or relevant genetic grouping, sex, animal class, production stage, age, live weight or relevant size, physiological status, herd identity, cohort identity, location, entry and exit dates, mortality, sale, transfer, mixing, and material cohort changes.

Cohort identity is a temporal relationship, not a static label. The framework should permit a reviewer to determine which animals were exposed, measured, modelled, included in performance outcomes, and represented in the claimed period. Where individual identification is infeasible or unnecessary, group-level identity must still preserve the rules by which animals entered, remained in, and left the group.

6.3 Production system, diet, and baseline

Production context includes pasture-based beef, feedlot beef, pasture-based dairy, partial mixed ration dairy, total mixed ration dairy, or another explicitly defined system. Fields include location, climate or season where material, grazing management, confinement, feed delivery, water access, diet ingredients, forage quality, ration composition, supplements, dry-matter intake or intake proxy, baseline performance, and material management changes.

Baseline evidence should identify:

- the baseline purpose and boundary;
- whether baseline is contemporaneous, historical, modelled, control-based, or counterfactual;
- baseline period and cohort;
- diet, performance, management, and methane conditions;
- exclusions, missing data, and changes;
- comparability between baseline and intervention periods;
- the method used to adjust for production, intake, season, or cohort differences.

6.4 Dose, delivery, exposure, and duration

The framework separates five propositions:

$$\text{prescribed dose} \neq \text{formulated dose} \neq \text{delivered dose} \neq \text{available dose} \neq \text{consumed dose.}$$

Evidence may include procurement, delivery, storage, feed-mill incorporation, ration formulation, mixing verification, feeding records, supplement placement, water-system records, refusal, wastage, residuals, interruptions, missed events, product substitutions, and corrective actions.

For controlled feeding, evidence may support pen- or group-level exposure through ration batches, delivery logs, bunk management, and intake estimates. For extensive grazing, evidence may rely on supplement disappearance, access patterns, electronic feeders, marker systems, mustering, water-point controls, or validated proxies. The proxy must be named, justified, and uncertainty-bearing.

Duration includes start, stop, ramp-up, adaptation, interruption, withdrawal, and follow-up. A claim period should not exceed the period supported by exposure and persistence evidence without an explicit model or assumption.

6.5 Methane measurement provenance

Measurement evidence identifies method, device, serial number, configuration, calibration, reference gases or tracers, sampling design, operator, animal binding, time, duration, unit, environmental conditions, raw or processed status, quality-control results, exclusions, missingness, aggregation, and uncertainty.

A minimum measurement record should include:

1. measurement objective and estimand;
2. measurement method and rationale;
3. device identity, configuration, software, and calibration;
4. sampling population and animal or cohort binding;
5. sampling frequency, duration, temporal coverage, and missingness;
6. raw-data commitment and processing pipeline;
7. exclusion and quality-control rules;
8. unit conversions and aggregation;
9. uncertainty components and final uncertainty expression;
10. known limitations and applicability to the declared decision.

Method-specific profiles should define additional requirements for respiration chambers, SF_6, GreenFeed or comparable automated systems, portable accumulation chambers, micrometeorological methods, breath analysers, proxies, and emerging sensing systems.

6.6 Model provenance and reproducible execution

Model evidence identifies model name, version, code or binary, environment, input schema, input sources, transformations, factors, coefficients, assumptions, calibration domain, uncertainty, sensitivity, reviewer state, and output artifact.

The reproducibility test is whether an independent reviewer with authorised access can reproduce the declared result from the frozen inputs and environment, or can explain precisely why reproduction is not possible. Reproducibility is necessary but not sufficient: the model may run exactly and still be outside its scientific domain.

6.7 Performance, health, adverse effects, and co-effects

Where interpretation depends on intake, liveweight gain, milk yield, milk composition, feed efficiency, reproduction, morbidity, mortality, or health, those outcomes must be temporally and cohort linked to the intervention. The framework should identify adverse events, diagnostic criteria, monitoring frequency, treatment, discontinuation, escalation, and reporting authority.

Performance evidence serves at least three purposes: detecting adverse co-effects; interpreting emissions intensity; and assessing whether observed or modelled reductions are associated with changes in intake or production. It must not be used to imply safety or productivity benefits beyond the observed scope.

6.8 Persistence, adaptation, interruption, and rebound

Persistence evidence distinguishes:

- persistence during continuous exposure;
- biological adaptation during treatment;
- effect under variable or intermittent exposure;
- effect after temporary interruption;
- rebound or return to baseline after discontinuation;
- carryover into subsequent periods;
- uncertainty when long-term evidence is absent.

Short-duration efficacy evidence should not be stretched across a longer implementation period without an explicit temporal model, supporting evidence, and uncertainty qualification.

6.9 Authority, rights, consent, and decision use

The framework identifies who is authorised to assert implementation, issue product records, control farm data, execute models, review science, provide verification determinations, allocate claims, and authorise institutional use. Possession of data does not automatically confer authority to assert, disclose, sell, or allocate outcomes.

Fields include issuer identity, role, authority basis, consent, data rights, confidentiality, intellectual-property restrictions, contractual claim rights, co-funding, co-claiming, exclusivity, supply-shed boundary, reporting purpose, and decision expiration.

6.10 Limitations, exceptions, and non-assertions

Every domain should support explicit limitations. A useful evidence packet states not only what is present but what is not asserted. Examples include:

- purchase is documented; administration is not fully observed;
- exposure is estimated at group level; individual intake is unknown;
- measurement covers a sample; extrapolation to the full cohort is model-based;
- the intervention is within the product evidence domain but outside the selected accounting method;
- model execution is reproducible; scientific applicability remains under review;
- a scientific reviewer found the evidence credible; accredited verification has not occurred;
- a buyer relied on the artifact for one purpose; no universal acceptance is asserted.

7 Decision-linked requirement profiles

Table 3: Illustrative decision-linked evidence profiles for consensus testing.

Decision purpose	Core evidence emphasis	Prohibited shortcut
Scientific efficacy	Intervention identity, study design, animals, diet, dose, measurement, statistics, performance, adverse effects, persistence.	Transferring an effect to a different formulation, diet, animal class, or duration without applicability evidence.
Implementation fidelity	Product and batch, delivery, administration, exposure, cohort continuity, interruptions, substitutions, records.	Treating purchase or delivery as proof of consumption.

Quantification	Baseline, activity data, measurement or model provenance, transformations, calibration domain, uncertainty, period and cohort.	Treating successful model execution as scientific applicability.
Method compatibility	Versioned method requirements, eligibility, model domain, exclusions, missing evidence, conditions.	Treating compatibility as approval, verification, or registry acceptance.
Verification readiness	Complete evidence graph, source access, controls, uncertainty, exceptions, authority, reproducibility.	Treating a complete checklist as a positive VVB determination.
Institutional reliance	Named artifact, institution, purpose, policy version, contractual conditions, rights, limitations, expiration.	Treating one institution's reliance as universal acceptance.
Claim allocation	Recognised outcome, contributors, rights, allocation rule, overlap, exclusivity, temporal and organisational boundaries.	Treating physical contribution or data possession as legal entitlement.
Payment	Contract, performance trigger, evidence source, calculation, approval, dispute and correction process.	Treating payment as scientific proof or emissions certification.

8 Methods

8.1 Study design

The proposed study has five linked components:

1. preregistered systematic or scoping review;
2. candidate-field extraction and evidence-domain synthesis;
3. modified Delphi or structured consensus process;
4. production-context and intervention-class case testing;
5. machine-readable schema and conformance evaluation.

The protocol should be frozen before screening and consensus results are known. Any amendment should record date, reason, decision-maker, and effect on interpretation.

8.2 Review protocol

The review will identify evidence domains used in controlled efficacy research, field implementation, methane quantification, model application, method assessment, verification, and value-chain claims.

The final protocol should specify databases, search strings, date range, language restrictions, grey-literature treatment, deduplication, inclusion and exclusion criteria, title and abstract screening, full-text screening, reviewer independence, conflict resolution, quality appraisal, extraction forms, and synthesis method.

Candidate databases include Web of Science, Scopus, CAB Abstracts, AGRICOLA, PubMed where relevant, and discipline-specific sources. Standards, official methods, and verifier guidance should be searched separately and clearly distinguished from peer-reviewed scientific evidence.

The review should include intervention classes broad enough to test framework generality while preserving class-specific requirements. Candidate classes include 3-NOP, macroalgae and halogenated compounds, nitrate, lipids, tannins and other plant secondary compounds, dietary manipulation, vaccines, breeding, rumen microbiome interventions, and management strategies.

8.3 Data extraction

Each included source should be coded for:

- study and publication identity;
- country, production system, animal class, diet, and environment;
- intervention class, product, formulation, active component, batch information;
- dose definition, delivery, exposure evidence, duration, interruptions;
- baseline and comparator;
- methane method, device, calibration, sampling, processing, uncertainty;
- model identity, inputs, domain, transformations, uncertainty;
- performance, health, safety, adverse effects, persistence;
- source-capture fields and retrospective reconstruction;
- decision use and supported conclusion;
- limitations, conflicts, funding, and study-quality indicators.

Extraction should distinguish absence from non-reporting. A paper that does not report a batch identifier does not necessarily prove that the identifier never existed. The evidence map should therefore represent `not_reported` internally during review, even if the final conformance vocabulary uses `missing` for the assembled decision artifact.

8.4 Quality appraisal and evidence synthesis

The review should use appraisal tools appropriate to study design rather than one generic score. Controlled animal experiments, observational field studies, validation studies, model studies, systematic reviews, and institutional documents have different bias structures.

The synthesis should identify:

- fields consistently reported across credible studies;
- fields associated with measurement validity, model applicability, safety, or persistence;
- intervention-specific moderators;
- production-context differences;
- recurring omissions;
- fields needed for implementation or verification but rarely reported in efficacy literature;
- unresolved scientific disagreements.

Meta-analysis may be used for selected questions if populations, interventions, comparators, outcomes, and designs are sufficiently comparable. The framework itself should not be justified solely by vote-counting study reports.

8.5 Consensus process

A modified Delphi or comparable structured process will involve researchers, Australian production experts, measurement specialists, model developers, producers, veterinarians or safety specialists where relevant, technology developers, VVB professionals, institutional users, and data-governance experts.

Participants should rate each candidate field on:

1. necessity for each decision purpose;
2. source-capture priority;
3. timing and frequency;
4. responsible issuer or authority;
5. acceptable evidence sources;

6. feasibility by production system;
7. uncertainty consequence if absent;
8. whether a proxy may be acceptable;
9. conditions under which the field is not applicable.

Panel composition, recruitment, expertise criteria, sector representation, conflicts, attrition, response distributions, comments, changes between rounds, and dissent should be reported. Consensus thresholds must be set prospectively. Minority disagreement should not be erased when it reveals material scientific or operational uncertainty.

8.6 Evidence-to-decision mapping

Each field will map to one or more decision uses. Fields without a defined use should be challenged to prevent unnecessary burden. Conversely, each decision should identify the minimum set of fields and permissible alternatives.

The mapping will produce a versioned matrix:

$$M_v(f, d, c, i) \to \{\text{required, conditional, recommended, optional, not applicable}\},$$

where f is the field, d the decision purpose, c the production context, i the intervention profile, and v the framework version.

8.7 Production-context matrix

Pasture-based beef, feedlot beef, and dairy systems will be evaluated separately. The matrix should compare:

- animal identification and cohort continuity;
- diet control and intake measurement;
- intervention delivery mechanisms;
- exposure variability;
- record systems and automation;
- measurement feasibility;
- model availability and calibration;
- performance data;
- exception frequency;
- producer burden and cost.

8.8 Case construction and independent classification

At least twelve synthetic or permissioned cases should cover:

- three production contexts;
- multiple intervention classes;
- measured and modelled quantification;
- complete, missing, conflicting, and out-of-domain evidence;
- product substitution;
- interrupted delivery;
- cohort change;
- short-duration persistence evidence;

- reproducible but scientifically inapplicable model execution;
- positive scientific review with pending verification;
- institution-specific acceptance or rejection.

Independent reviewers will classify each field and domain, identify the responsible actor, state the supported conclusion, and record required escalation. Agreement should be reported by field, domain, decision purpose, production context, and gap state.

8.9 Machine-readable schema and conformance tests

The approved vocabulary will be represented as a versioned schema with controlled terms, cardinality, data types, provenance, authority, temporal bounds, uncertainty, limitations, and decision links. The schema should not contain hidden approval logic.

Conformance tests should include:

1. valid complete case;
2. missing batch identity;
3. purchase without administration;
4. delivered dose without exposure evidence;
5. cohort discontinuity;
6. device calibration expired;
7. raw data unavailable;
8. model version unspecified;
9. calibration-domain mismatch;
10. conflicting input records;
11. persistence extrapolation beyond evidence;
12. unauthorised issuer;
13. claim-rights ambiguity;
14. method update and re-evaluation;
15. export that drops limitations.

A conforming system should preserve the gap state and diagnostic basis. It should not transform a structurally valid record into a positive scientific or institutional determination.

8.10 Producer burden and proportionality

Burden should be measured rather than assumed. Candidate metrics include setup time, recurring manual minutes, automated capture share, number of source systems, exception frequency, training burden, data-access cost, equipment cost, record-retention period, and audit retrieval time.

The study should identify fields that can be reused from existing farm, feed, herd, dairy, laboratory, and commercial systems. Specialist measurement should be required only where the scientific or method question demands it. A minimum framework that maximises fields rather than decision value would fail its purpose.

8.11 Analysis plan

Planned quantitative analyses include review agreement, consensus distributions, attrition, field-level necessity ratings, production-context differences, case-classification agreement, and burden distributions. Appropriate statistics may include

Cohen's or Fleiss' kappa, Gwet's agreement coefficient, Krippendorff's alpha, intraclass correlation, ordinal models, and descriptive uncertainty intervals, selected according to data structure and preregistered assumptions.

Qualitative comments and dissent should be coded using a transparent thematic process. Any change to field definitions resulting from comments should be traceable to the version and rationale.

9 Illustrative production-context profiles

Table 4: Illustrative context differences for consensus testing; not final requirements.

Dimension	Pasture-based beef	Feedlot beef	Dairy
Diet control	Variable pasture, seasonal forage, supplements, lower direct intake observability.	Formulated rations, feed delivery and bunk records, pen-level intake.	Pasture, partial mixed ration, or TMR; milk and feeding systems may provide repeated records.
Delivery	Blocks, lick, loose supplement, water, strategic feeding, bolus, other intermittent routes.	Feed-mill or bunk incorporation, water, direct dosing.	TMR/PMR, concentrate at milking, automated feeders, direct dosing.
Exposure fidelity	Often group-level and variable; proxy validation may be central.	Pen-level delivery can be strong; individual consumption may remain uncertain.	Individual visits or milk-system data may support finer binding, subject to sampling bias.
Cohort continuity	Movement across paddocks, properties, sales, musters, and extensive conditions.	Pen entry, transfer, hospital pen, exit, and slaughter records.	Herd-management systems, lactation stage, dry-off, culling, and treatment groups.
Measurement	Field methods and models may dominate; repeated direct measurement can be difficult.	Research or commercial measurement may be more feasible in controlled settings.	Automated spot sampling and repeated visits may be feasible in some systems.
Burden risk	Manual records, connectivity, sparse infrastructure, and long intervals.	Integration across feed, yard, animal, and commercial systems.	Integration across herd, milk, feed, sensor, and laboratory systems.

10 Worked cases

10.1 Commercial product name matches, batch identity missing

A dairy operation reports use of a commercial inhibitor whose product name matches published research, but the batch identifier and active concentration are unavailable. Purchase and administration records are present.

The evidence supports product-class and commercial-identity continuity but not batch-level characterisation. The appropriate state may be `conditionally_sufficient` for an implementation record and `missing` for a decision requiring batch confirmation. The system must not infer that every batch has the same active concentration or stability.

10.2 Purchase and delivery without observed administration

A grazing enterprise purchases enough supplement for the declared herd and period. Delivery records are complete, but administration events, supplement disappearance, animal access, and interruptions are not recorded.

Commercial supply is `present`; delivered exposure is `missing`. The evidence may support procurement but not implementation fidelity or quantified abatement. A proxy may be proposed, but it must be declared and evaluated rather than silently substituted.

10.3 Feedlot ration inclusion with cohort transfer

A feedlot documents ration inclusion and pen-level feed delivery. During the intervention, animals are transferred among pens and a hospital group. Quantification uses the original pen population.

Dose and delivery evidence are present, but cohort continuity is conflicting. The reviewer must reconcile entry, transfer, treatment, and exit dates. An aggregate estimate may remain possible if the method defines a defensible allocation rule; otherwise the cohort result is insufficient.

10.4 Valid measurements with inadequate temporal coverage

An automated system has current calibration and valid spot measurements, but only a small subset of animals visited during a short interval after feeding.

Device provenance is present. Sampling representativeness is conditionally sufficient or missing, depending on the estimand and protocol. The system must not convert valid instrument readings into a representative cohort mean without supporting sampling evidence.

10.5 Reproducible model outside calibration domain

A model execution is fully reproducible, but the intervention, diet, animal class, or production system lies outside the empirical calibration domain.

Execution provenance is present; scientific applicability is `out_of_domain`. A method extension, alternative model, or additional study may be required. Reproducibility does not cure domain mismatch.

10.6 Short-term trial extrapolated across an annual claim

A controlled trial demonstrates an effect for eight weeks. A project applies the effect factor to a twelve-month period without persistence evidence or an explicit temporal model.

Initial efficacy evidence may be present. Annual persistence is missing or out of domain. The supported period must be bounded, or the extrapolation must be separately modelled, justified, and uncertainty-adjusted.

10.7 Scientifically credible evidence, unresolved claim rights

The intervention, delivery, quantification, and scientific review are complete. Two value-chain participants assert overlapping rights to the same recognised outcome.

Scientific and implementation evidence may be sufficient. Claim allocation is conflicting. The conflict must not be represented as a scientific failure, and scientific credibility must not be represented as legal entitlement.

10.8 Australian extensive-system proxy accepted conditionally

An extensive beef operation cannot measure individual intake. It uses a validated group-level exposure proxy supported by supplement disappearance, animal access controls, periodic observation, and uncertainty bounds.

The field may be `conditionally_sufficient` for a specific method and context. The condition, proxy validation, uncertainty, and decision scope must remain visible. The acceptance does not establish that the proxy is universally sufficient.

11 Planned results reporting

Planned results - not findings. No review, consensus, case-classification, conformance, or producer-burden findings are reported in this draft. The final manuscript must publish search yield, study characteristics, quality assessment, field extraction, consensus ratings, dissent, context differences, case-test results, schema revisions, and burden estimates.

Table 5: Paper 3 planned results template.

Output	Result	Required reporting

Review evidence map	TBD	Search dates, databases, records, screening flow, included studies, intervention classes, contexts, measurement methods, models, quality, and gaps.
Candidate field inventory	TBD	Field definitions, evidence source, intervention specificity, production context, and decision use.
Consensus outcome	TBD	Panel composition, rounds, thresholds, distributions, dissent, conflicts, attrition, and changes.
Decision matrix	TBD	Required, conditional, recommended, optional, and not-applicable fields by purpose, context, and intervention profile.
Context profiles	TBD	Pasture, feedlot, and dairy differences; acceptable sources and proxies; unresolved issues.
Inter-rater classification	TBD	Agreement by field, domain, state, context, and decision purpose; disagreement analysis.
Conformance cases	TBD	Inputs, expected states, diagnostics, prohibited outputs, and implementation results.
Producer burden	TBD	Setup and recurring effort, automation share, exceptions, costs, retrieval time, and distribution across systems.
Framework revision	TBD	Accepted changes, rejected changes, rationale, dissent, and version history.

12 Discussion

12.1 Minimum evidence as an interface between science and implementation

A minimum-evidence framework can reduce late reconstruction by making expectations visible before protocol, verification, or buyer review. Its value is not that every participant stores every possible datum. Its value is that the chain from scientific evidence to implementation and decision can be inspected without silently filling gaps.

The framework acts like a shipping manifest for scientific meaning. A physical shipment can arrive with intact packaging while lacking the manifest needed to determine what it contains, where it came from, or whether it meets import conditions. Likewise, a methane estimate can be numerically precise while lacking the identity, context, exposure, provenance, and authority needed to interpret it. The manifest does not prove the goods are valuable; it makes their identity and handling inspectable.

12.2 Measurement validity and operational validity must remain connected but distinct

Measurement literature supports strong method-specific provenance requirements [7, 6]. Yet a high-quality measurement does not prove that a commercial implementation matches the measured trial. Conversely, imperfect field evidence may still be informative when limitations are explicit.

The framework therefore avoids two extremes: laboratory exceptionalism, in which only tightly controlled measurements are treated as evidence; and operational permissiveness, in which any farm record is treated as sufficient. The correct standard depends on the decision and the inferential distance between the evidence and the conclusion.

12.3 Calibration domain is an evidence requirement

Diet-dependent effects for 3-NOP [4], formulation and active-component issues for *Asparagopsis* [9], and context-sensitive responses to lipids and nitrate [1, 2] demonstrate why calibration domain cannot be an optional model footnote. Domain evidence belongs in the primary decision artifact.

A model or effect factor is like a map drawn at a particular scale for a particular terrain. The map may be exact within its surveyed region. Carrying it into unsurveyed terrain does not make the map fraudulent, but it changes the confidence that a route exists where the paper shows one. Calibration-domain evidence marks the surveyed boundary.

12.4 Australian context changes the evidence pathway

Australian extensive systems do not justify lower transparency. They require different evidence strategies. Where daily individual intake is infeasible, the framework should specify acceptable group-level sources, validated proxies, uncertainty, and conditions. Where connectivity is limited, records may be captured offline and committed later, but temporal integrity and correction rules must remain visible.

The framework should also avoid embedding one current method or program into the scientific core. Australian profiles should be versioned and authority-specific. This preserves local relevance without turning a scientific publication into an unofficial regulatory instrument.

12.5 Proportionality is a scientific requirement

Excessive data requirements can degrade evidence by encouraging retrospective fabrication, low-quality box-ticking, or producer disengagement. Proportionality is therefore not merely a cost concern. It affects validity.

Every required field should answer: which decision becomes materially less reliable if this field is absent? If the answer is unclear, the field should be recommended, optional, or removed. Where a low-burden proxy can support the decision with explicit uncertainty, the framework should permit it. Where the decision cannot be supported without direct evidence, the framework should say so.

12.6 Machine readability does not create scientific authority

A schema can enforce structure, types, required relationships, and controlled states. It can reveal that calibration evidence is missing or that a model version is unspecified. It cannot determine by itself that a scientific inference is credible, that a method authority accepts an extension, that a VVB has reached a conclusion, or that a buyer owns a claim.

The schema is analogous to a well-designed laboratory notebook template. It can prompt the scientist to record batch, calibration, method, and exceptions. It cannot conduct the experiment or interpret every anomaly. Conformance means the notebook preserves the necessary distinctions, not that every entry is scientifically correct.

12.7 Value-chain use requires bounded reliance

Evidence-to-decision mapping should support buyers, processors, lenders, programs, and other institutions without implying universal reliance. An institution may accept a bounded artifact for supplier engagement while rejecting it for a public claim. A verifier may regard the bundle as ready for engagement while making no positive determination. A method authority may require an extension despite favourable scientific review.

Paper 2's three-layer validity model should therefore remain visible in every Paper 3 output. Evidence sufficiency supports a decision pathway; it does not collapse integrity, compatibility, verification, and reliance into one status.

13 Limitations

The proposed framework has several limitations.

First, the review may not capture unpublished operational knowledge, confidential product information, or emerging interventions. Grey literature may contain practical evidence but variable quality.

Second, consensus depends on panel composition. Researchers, producers, developers, verifiers, buyers, and government-facing participants may value different fields. Conflicts and sector imbalance may shape outcomes.

Third, Australian production systems are heterogeneous. Three broad profiles cannot represent every region, breed, climate, enterprise, delivery system, or data environment.

Fourth, intervention classes evolve. Product formulations, active concentrations, delivery systems, safety evidence, and model relationships may change. Requirements must be versioned.

Fifth, some fields may be infeasible or costly in extensive systems. The framework can expose the gap and evaluate proxies, but it cannot make unavailable evidence appear.

Sixth, machine-readable conformance tests assess representation and declared logic. They do not prove factual accuracy, scientific efficacy, regulatory compliance, legal entitlement, or institutional acceptance.

Seventh, the supplied citations are a high-impact starting set rather than a completed review. The final manuscript must verify metadata, add primary studies and recent reviews, and document search completeness.

14 Conclusion

The transition from controlled efficacy to repeatable implementation and institutional use requires more than a published treatment effect. It requires evidence of intervention and batch identity, animal and production context, dose and exposure, duration, measurement or model provenance, uncertainty, performance, safety, persistence, authority, rights, and limitations.

This paper proposes a transparent process for defining those minimums rather than asserting them by author preference. A preregistered evidence review, structured consensus process, production-context testing, independent case classification, and machine-readable conformance framework can identify what must be captured at source, what may be calculated or proxied, what requires scientific judgment, and what remains an institutional decision.

The intended result is not a universal approval checklist. It is a versioned evidence contract that makes gaps, conditions, disagreements, and out-of-domain cases visible. Such a framework can improve evidence readiness for Australian beef and dairy interventions while preserving the authority boundaries of scientists, method bodies, accredited verifiers, governments, registries, producers, and value-chain institutions.

15 Declarations

Funding and support. TODO: Disclose AgEvidence support, direct and in-kind funding, data access, technical assistance, and the role of any sponsor in protocol design, analysis, interpretation, writing, and submission.

Competing interests. TODO: Collect and publish financial, advisory, governance, intellectual-property, product-development, method-development, verification, and value-chain relationships for all contributors and panel participants.

Ethics. TODO: Determine whether expert elicitation, producer interviews, usability studies, or identifiable operational cases require human-research ethics review or exemption.

Data availability. TODO: Publish search strategies, screening decisions where permitted, extraction forms, field dictionary, de-identified ratings, dissent summaries, synthetic cases, schemas, validation rules, and conformance fixtures.

Code availability. TODO: Archive schema validators, transformation code, test corpus, environment manifest, and version history under a declared licence.

Author contributions. TODO: Complete a live CRediT record and obtain written consent, affiliation confirmation, conflict disclosure, and final-manuscript approval from every author.

Use of generative tools. TODO: Apply the target journal's disclosure policy and retain human responsibility for all claims, citations, analysis, and final text.

16 Executable TODO register

Table 6: Executable Paper 3 TODO register.

ID	Task	Accountable owner	Status
P3.01	Confirm lead, co-authors, affiliations, CRediT roles, and conflicts.	Persona D	Open
P3.02	Select primary journal and article type; lock reporting requirements.	Lead author	Open
P3.03	Approve review scope, intervention taxonomy, production contexts, and decision purposes.	Persona D; co-author approval	Open
P3.04	Verify every supplied citation against publisher records and add current primary literature.	Literature-review team	Open
P3.05	Register or archive the systematic/scoping review protocol.	Persona D	Open

P3.06	Complete search, deduplication, screening, appraisal, extraction, and synthesis.	Review team	Open
P3.07	Freeze candidate field dictionary, definitions, evidence classes, and gap states before consensus.	Scientific authors	Open
P3.08	Define panel sampling, expertise criteria, conflicts, thresholds, attrition rules, and dissent reporting.	Methods adviser	Open
P3.09	Obtain ethics determination and participant consent materials where required.	Program manager	Open
P3.10	Recruit balanced Australian and international panel participants.	Persona D; Persona A	Open
P3.11	Conduct consensus rounds; publish distributions, comments, attrition, and dissent.	Consensus team	Open
P3.12	Build decision-linked requirement matrix and versioned production-context profiles.	Scientific authors	Open
P3.13	Construct at least twelve contrasting synthetic or permissioned cases.	Persona G; Persona F	Open
P3.14	Run independent case classification and agreement analysis.	Independent methods team	Open
P3.15	Build machine-readable schema, validators, and negative conformance fixtures.	Persona G	Open
P3.16	Measure producer capture burden and test proportionality.	Persona F	Open
P3.17	Obtain Australian external scientific, producer, method, and verifier review.	Persona A	Open
P3.18	Resolve terminology dependencies with Papers 1 and 2 through versioned portfolio change control.	Scientific Chairs	Open
P3.19	Complete independent methods and claims review; log unresolved issues.	External reviewers	Open
P3.20	Complete data/code statement, CRediT, funding, conflicts, limitations, author approval, and submission archive.	Persona D	Open

Bibliography

[1] Almeida, A. K., Hegarty, R. S., and Cacho, O. (2021). Meta-analysis of methane emissions from beef cattle fed lipid supplements. *Animal Feed Science and Technology*, 271, 114717.

[2] Appuhamy, J. A. D. R. N., Strathe, A. B., Jayasundara, S., Wagner-Riddle, C., Dijkstra, J., France, J., and Kebreab, E. (2016). Anti-methanogenic effects of nitrate supplementation in cattle: A meta-analysis. *PLoS ONE*, 11(8), e0160563.

[3] Charmley, E., Williams, S. R. O., Moate, P. J., Hegarty, R. S., Herd, R. M., Oddy, V. H., and colleagues (2016). A universal equation to predict methane production of forage-fed cattle in Australia. *Animal Production Science*, 56(3), 169–180.

[4] Dijkstra, J., Bannink, A., France, J., Kebreab, E., and van Gastelen, S. (2018). Antimethanogenic effects of 3-nitrooxypropanol depend on dietary composition in dairy cattle: A meta-analysis. *Journal of Dairy Science*, 101(10), 9041–9057.

[5] Eckard, R. J., Grainger, C., and de Klein, C. A. M. (2010). Options for the abatement of methane and nitrous oxide from ruminant production: A review. *Livestock Science*, 130(1–3), 47–56.

[6] Garnsworthy, P. C., Craigon, J., Hernandez-Medrano, J. H., and Saunders, N. (2012). Variation among individual dairy cows in methane emissions measured during milking. *Journal of Dairy Science*, 95(6), 3181–3189.

[7] Hammond, K. J., Crompton, L. A., Bannink, A., Dijkstra, J., Yáñez-Ruiz, D. R., O'Kiely, P., and colleagues (2016). Review of methodologies for measuring enteric methane emissions from ruminants. *Animal Feed Science and Technology*, 216, 19–37.

[8] Kebreab, E., Clark, K., Wagner-Riddle, C., and France, J. (2006). Methane and nitrous oxide emissions from Canadian animal agriculture: A review under Intergovernmental Panel on Climate Change Tier 1 and Tier 2 methodologies. *Canadian Journal of Animal Science*, 86(2), 135–158.

[9] Kinley, R. D., Martinez-Fernandez, G., Matthews, M. K., de Nys, R., Magnusson, M., and Tomkins, N. W. (2020). Mitigating enteric methane emissions from cattle using *Asparagopsis taxiformis*. *Journal of Cleaner Production*, 259, 120836.

[10] Minx, J. C., Callaghan, M., Lamb, W. F., Bornmann, L., and Fuss, S. (2017). Learning about climate change mitigation solutions: The role of systematic reviews and meta-analyses. *Environmental Research Letters*, 12(1), 013001.

[11] Searchinger, T. D., Wirsenius, S., Beringer, T., and Dumas, P. (2018). Assessing the efficiency of changes in land use for mitigating climate change. *Nature*, 564(7735), 249–253.

[12] Thompson, L. R., and Rowntree, J. E. (2020). Invited review: Methane mitigation in grazing ruminants. *Applied Animal Science*, 36(5), 577–587.

Paper 4: From Trial to Farm Gate

Working-draft note. This manuscript expands Paper 4 from the *AgEvidence Australia Scientific Publication Series: Seven Manuscript First Drafts.* It is an applied-methods and case-study manuscript. It does not report a completed pilot, approved Project 4030 dataset, verified reduction, recognised Scope 3 claim, registry outcome, government acceptance, or product comparison. All graph-completeness values, reconstruction times, discrepancies, claim allocations, burden estimates, and conformance outcomes remain planned results.

Non-negotiable guardrail. Verification readiness is not an accredited validation or verification determination. Method compatibility is not government approval. The aggregate-cap invariant is a minimum accounting control, not a legal allocation rule. Product, funding, contribution, contract, registry, and value-chain claim rights remain subject to the governing method, policy, contract, accounting framework, and competent authority.

Abstract

Controlled feed-additive trials can produce credible evidence of enteric methane reduction, but commercial deployment introduces additional evidentiary dependencies distributed across manufacturers, logistics providers, producers, feed systems, animal cohorts, model operators, verifiers, registries, funders, and downstream claimants. Product batch variation, ration-inclusion error, missed feeding intervals, cohort turnover, changing diets, fragmented data custody, and retrospective spreadsheet assembly can weaken the link between trial efficacy and a reconstructable value-chain outcome.

This paper proposes an append-only directed acyclic evidence graph for feed-additive deployment in Australian beef and dairy systems. The graph implements the thin-waist receipt contract of Paper 1, the three-layer validity model of Paper 2, and the minimum-evidence vocabulary of Paper 3. It binds scientific trial evidence to intervention identity, batch and purchase records, delivery and storage, administration and ration events, animal-cohort membership, baseline performance, methane measurement or model execution, scientific review, verification-readiness artifacts, outcome recognition, funding contributions, rights evidence, co-claim allocation, producer payment, and institution-specific reliance. Every node declares issuer, authority, event time, source commitments, parent relationships, limitations, and supersession status.

The architecture is evaluated through a synthetic mirror, permissioned pilot subset, or approved post-pilot case and an independent reconstruction protocol. Reviewers receive a bounded bundle rather than access to a hosted workflow and attempt to reconstruct implementation, quantitative calculations, exception handling, and claim allocation. The scenario suite includes product substitution, incomplete feeding logs, missed dose, cohort change, animal movement, model-version update, missing baseline, conflicting claim, pending verification scope, and corrected source records.

A central contribution is mathematical separation of the recognised physical greenhouse-gas outcome Q from commercial or institutional claim allocations a_i. Under a declared policy profile, the graph enforces the aggregate-cap invariant $\sum_i a_i \leq Q$. This prevents recorded quantitative allocations from silently exceeding the recognised outcome while leaving entitlement, claim language, and legal effect to relevant authorities. The study evaluates traceability, reconstruction effort, quantitative discrepancy, graph completeness, exception visibility, and aggregate-cap behaviour rather than product efficacy or formal method, registry, VVB, or government approval.

Keywords: feed additive; enteric methane; evidence graph; chain of custody; Scope 3 accounting; co-claims; livestock supply chain; verification readiness; data provenance.

1 Manuscript status, scope, and contribution boundary

Paper 4 is the first applied graph paper in the portfolio. It depends on Paper 1 for the stable receipt envelope, append-only graph, source commitments, immutable parent relationships, versioned adapters, and bounded reliance artifacts. It depends on Paper 2 for the separation of cryptographic integrity Φ, method compatibility Ψ, scientific judgment, accredited verification, and institutional reliance Ω. It depends on Paper 3 for intervention identity, dose, exposure, measurement provenance, model provenance, persistence, evidence sufficiency, and decision-linked minimum requirements.

The manuscript must therefore reuse, rather than redefine, the following portfolio terms:

- receipt, source commitment, parent, evidence root, adapter, artifact, and reliance event;

- `valid`, `invalid`, and `indeterminate` integrity states;
- `eligible`, `eligible_with_conditions`, `outside_current_method`, `method_extension_required`, `insufficient_evidence`, and `unassigned` compatibility states;
- `accepted`, `relied_on`, `rejected`, `needs_more_evidence`, `pending`, and `not_assessed` reliance states; and
- `present`, `missing`, `conflicting`, `out_of_domain`, `conditionally_sufficient`, and `not_applicable` evidence states.

The contribution is an applied architecture and evaluation protocol. It does not establish product efficacy, an official Australian minimum dataset, a recognised reduction, an accepted Scope 3 claim, legal entitlement, or a verified project.

2 Provisional scientific author personas

Persona	Provisional role	Contribution domain / deliverable
Persona B	Lead/corresponding-author persona	Trial-to-deployment scientific mapping, intervention efficacy, model identity/applicability and quantitative pathway.
Persona A	Scientific coauthor persona	Australian beef/dairy production, baseline, accounting and local-method boundaries.
Persona D	Scientific coauthor persona	Measurement provenance, uncertainty, evidence sufficiency and conformance requirements.
Persona C	Scientific coauthor persona	Verification readiness, value-chain use, rights, co-claims and allocation profiles.
Persona F	Scientific coauthor persona	Feed delivery, cohort change, implementation fidelity, exception realism and producer burden.
Persona G	Technical scientific contributor persona	Evidence graph schema, verifier, exception fixtures and independent reconstruction package.
Persona H	Governance contributor persona	Data route, permissions, de-identification, source custody and disclosure controls.
Persona I	Independent reviewer persona	Reconstruction, quantitative discrepancy and claims review without privileged platform access.

Authorship control. Persona IDs are recruitment placeholders only. A real person is not an author until consent, qualifying contribution, affiliation, conflict disclosure, CRediT allocation, and final-manuscript approval are complete.

3 Introduction

3.1 The efficacy-to-deployment gap

Controlled trials answer a bounded efficacy question: under specified animals, diets, doses, durations, measurements, and analytical procedures, did an intervention alter methane production or methane yield? Commercial deployment asks a broader set of questions. Was the same product and active formulation supplied? Was it stored correctly? Was the target dose mixed and offered? Which animals were exposed, for how long, and under which diet? Were interruptions, substitutions, cohort turnover, and refusals recorded? Which baseline and model version were used? Who reviewed the interpretation? Which institution relied on the result, and for what claim?

Research on 3-nitrooxypropanol and *Asparagopsis* shows why efficacy cannot be represented as a generic product constant. Response depends on mechanism, diet, dose, formulation, and operating conditions (Duin et al. 2016; Dijkstra et al. 2018; Alvarez-Hess et al. 2019; Kinley et al. 2020; Romero-Perez et al. 2014). Translating trial results into farm deployment therefore requires an evidence chain that preserves both scientific applicability and operational fidelity.

The deployment problem is particularly acute across fragmented value chains. Manufacturers hold batch records; distributors hold purchase and delivery records; farms hold ration and cohort records; laboratories and devices hold measurement data;

model operators hold calculations; verifiers hold review evidence; and buyers or processors hold claim and reliance decisions. Static folders and ad hoc spreadsheets may reproduce an initial report while leaving undocumented transformations and authority assumptions.

3.2 Why an evidence graph

An evidence graph represents each event, source, calculation, judgment, and institutional decision as a bounded node linked to its dependencies. Graph-based lineage and enterprise data-lineage research support explicit dependency resolution across distributed workflows and organisational boundaries (He et al. 2020; Kruse et al. 2021). Agricultural supply-chain research also shows that integrity mechanisms are useful only when governance, interoperability, and source authority are addressed (Tönissen and Teuteberg 2020).

The graph is append-only. Product substitution, missed dose, corrected ration data, cohort change, and model updates are appended rather than used to erase the original record. This is aligned with Paper 1's rule that historical evidence is immutable while interpretation is versioned.

3.3 Outcome and claim are different objects

A physical or recognised greenhouse-gas outcome is not the same as a commercial claim. Let Q denote the recognised outcome for one bounded cohort, period, baseline, and method profile. Let a_i denote the quantitative claim allocated to participant i. The minimum aggregate-cap rule is

$$\sum_{i=1}^{n} a_i \leq Q. \tag{1}$$

This rule prevents the graph from recording allocated quantities that exceed the recognised outcome. It does not determine who is entitled to claim, whether two actors may make differently worded claims, whether an intervention counts toward a target, or whether an institution accepts the allocation. Scope 3 accounting and target-setting frameworks make value-chain boundaries and claim use consequential (World Resources Institute and World Business Council for Sustainable Development 2011; SBTi 2022). Carbon-accounting scholarship likewise warns that attribution and double counting are governance problems rather than arithmetic alone (Ascui and Cojoianu 2019; Gopalakrishnan et al. 2023; Vanderhaegen et al. 2021).

3.4 Research questions

1. Can the Paper 1 receipt contract represent the full trial-to-farm-to-claim pathway without losing scientifically material context?
2. Can Paper 3 minimum-evidence fields be populated, linked, and evaluated across real operational exceptions?
3. Can the same evidence root support Paper 2 integrity, compatibility, verification, and reliance states without validity laundering?
4. Can an independent reviewer reconstruct implementation and quantification without privileged access to the hosted platform?
5. Does the graph expose missing, conflicting, out-of-domain, conditional, and not-applicable evidence rather than infer completeness?
6. Can contribution, rights, payment, and claim allocation remain separate from the physical outcome Q?
7. Does the aggregate-cap invariant prevent over-allocation across co-claimants while preserving institution-specific policy?

3.5 Novelty relative to conventional deployment dossiers

Table 2: Novelty analysis for the proposed trial-to-farm evidence graph.

Dimension	Conventional approach	Paper 4 proposed graph
Data provenance	Static retrospective folder assembly and ad hoc spreadsheet calculations.	Append-only directed graph with immutable lineage, source commitments, authority, and supersession.
Evidence sufficiency	Treats document presence as completion.	Reuses Paper 3 field states and decision-linked requirement profiles.
Validity	Uses one "verified" or "approved" status.	Preserves Paper 2 integrity, compatibility, review, verification, and reliance separately.
Claim allocation	Couples outcome quantification to claim rights.	Separates Q, contribution, rights, payment, and allocations a_i; enforces $\sum a_i \leq Q$.
Operational dynamics	Assumes perfect compliance and static cohorts.	Represents missed doses, batch changes, substitutions, movements, refusals, and cohort turnover natively.
Auditability	Requires manual reassembly for each review.	Produces bounded bundles for offline verification and independent reconstruction.
Method change	Rewrites or replaces prior spreadsheets.	Re-evaluates one evidence root under immutable method and policy versions.

4 Scientific and methodological foundations

4.1 Efficacy-to-deployment translation

Controlled efficacy is evidence about a treatment under declared conditions, not a transferable constant. For 3-NOP, biochemical mechanism and trial evidence support a product-specific pathway, while meta-analysis and pasture studies show that response depends on diet and delivery context (Duin et al. 2016; Dijkstra et al. 2018; Alvarez-Hess et al. 2019). Beef studies further show why dose, duration, and performance outcomes must remain visible (Romero-Perez et al. 2014). For *Asparagopsis*, product characterisation, active component, dose, storage, and diet remain material to interpretation (Kinley et al. 2020).

Paper 4 therefore does not copy a literature efficacy percentage into a farm outcome. It links the implemented product and operating conditions to the scientific evidence and records the compatibility judgment under a declared method version.

4.2 Implementation fidelity and cohort tracking

Implementation fidelity includes the relationship among target dose, product received, product loaded, ration mixed, feed offered, feed refused, animal exposure, and interval continuity. Commercial adoption research identifies cost, delivery, supply, management, and record constraints (Black et al. 2021). Productivity and animal-performance evidence are also necessary because absolute emissions, intensity, and production response may diverge (Capper and Bauman 2013).

Individual-animal variation and changing cohort membership motivate explicit animal or cohort bindings (Garnsworthy et al. 2012). A pen-level ration receipt cannot be applied automatically to animals that entered after the feeding period or left before it ended. Cohort membership is therefore time-indexed.

4.3 Digital chain of custody and provenance

The graph adopts Paper 1's bounded receipt model. Every receipt records issuer, subject, authority, event and recording times, parents, source commitments, limitations, and signature. Provenance literature distinguishes source lineage from later interpretation (Buneman et al. 2001; Moreau 2010; Lebo et al. 2013). Tamper-evident structures and content-addressed systems support independent integrity checks but do not prove physical truth (Crosby and Wallach 2009; Benet 2014).

In supply chains, authority and custody are distributed. A manufacturer may assert batch composition; a carrier may assert delivery; a producer may assert administration; a model operator may assert execution; a verifier may issue a determination; and a buyer may record reliance. No actor inherits authority to make all propositions.

4.4 Scope 3 accounting and allocation

Corporate value-chain accounting requires explicit organisational and inventory boundaries (World Resources Institute and World Business Council for Sustainable Development 2011). FLAG guidance creates additional target-setting context for land and agricultural emissions (SBTi 2022). IPCC guidance supplies inventory methods and factors but does not itself allocate corporate claim rights (Intergovernmental Panel on Climate Change 2019).

The graph therefore distinguishes:

- physical intervention and production evidence;
- quantified or recognised outcome;
- funding and operational contribution;
- contractual and data rights;
- quantitative allocation;
- claim language and reporting boundary;
- retirement or cancellation where relevant; and
- institution-specific reliance.

4.5 Verification readiness and independent reconstruction

ISO 14064-3 and broader assurance principles require sufficient appropriate evidence, defined criteria, scope, materiality, and competent review (International Organization for Standardization 2019; International Auditing and Assurance Standards Board 2013). Agricultural market literature also emphasises uncertainty, data availability, and verification burden (Oldfield et al. 2022). A graph can improve readiness by exposing lineage and gaps, but it cannot substitute for verifier judgment.

Working-draft note. The user-supplied bibliography has been integrated. Several supplied records are marked "verification pending" because their metadata or existence has not been independently confirmed. Those entries must be checked against publisher or official records before submission.

5 Evidence graph model

5.1 Graph definition

Let $G = (V, E)$ be a directed acyclic graph. Each vertex $v \in V$ is a signed receipt, and each edge $(u, v) \in E$ means that v depends on or references u. The graph root h_G is a commitment to the bounded receipt set and manifest:

$$h_G = H(\mathrm{Canon}(\mathcal{M}, \{h_v : v \in V\})) ,$$

where $\mathcal{M}$ is the bundle manifest and h_v is the canonical digest of receipt v.

The graph is append-only. A correction v_c references the original receipt v_o, declares a correction reason and authority, and does not erase v_o. A superseding method or policy determination similarly appends a new decision receipt linked to the same evidence root.

5.2 Scientific and product evidence

The scientific branch contains:

- literature and trial identifiers;
- intervention mechanism and class;
- research formulation and active component;
- animal, diet, dose, duration, and measurement context;

- effect estimate, uncertainty, performance, and adverse effects;
- calibration-domain and extrapolation limitations; and
- scientific review events.

The product branch contains manufacturer identity, product name, formulation, active concentration, batch, certificate of analysis, manufacture and expiry, storage, authorisation status, substitution, and linkage to the scientific formulation.

A product-name match without batch or formulation continuity is represented as a Paper 3 evidence gap rather than inferred equivalence.

5.3 Commercial and operational implementation

Operational nodes include:

- purchase order, invoice, dispatch, and delivery;
- custody transfer and storage;
- feed-mixer load and ration formulation;
- administration or feeding event;
- offered and refused feed;
- dose interruption, substitution, or missed feeding;
- operator, equipment, and source-system identity; and
- correction and exception events.

The graph separates target, manufactured, delivered, mixed, offered, and estimated consumed dose. Let D_t, D_m, D_d, D_r, D_o, and D_c denote these quantities. The pathway is not assumed to be lossless:

$$D_c \leq D_o \leq D_r \leq D_d,$$

unless a declared measurement or reconciliation explains otherwise. The inequality is a diagnostic expectation, not a universal biological law.

5.4 Animal and cohort binding

Let $C(t)$ be the set of animals in a cohort at time t. A cohort interval $[t_0, t_1]$ is not represented by one static list when membership changes. Entry and exit receipts update the time-indexed cohort:

$$C(t^+) = \big(C(t^-) \cup A_t\big) \setminus R_t,$$

where A_t and R_t are animals added and removed at time t.

Implementation evidence is eligible for an animal or cohort only where temporal overlap, identity, production context, and exposure rules are satisfied. This prevents a ration record from being applied to animals outside the relevant pen or interval.

5.5 Measurement and model execution

Measurement receipts preserve method, device, calibration, operator, sampling design, raw data, exclusions, transformations, aggregation, and uncertainty, consistent with Paper 3.

Model-execution receipts preserve:

- model and version;
- code or executable digest;
- input receipt references;

- units and transformations;
- parameters and factor sources;
- calibration-domain determination;
- execution environment;
- output and uncertainty;
- limitations; and
- supersession links.

A reproducible execution may still be `out_of_domain`. Paper 2 requires that cryptographic validity $\Phi = \texttt{valid}$ not be displayed as method eligibility $\Psi = \texttt{eligible}$.

5.6 Scientific review and verification readiness

A scientific review receipt identifies reviewer, declared competence, question, evidence set, conclusion, limitations, conflicts, and dissent. A verification-readiness artifact identifies the evidence root, adapter and method versions, included and excluded evidence, gaps, deterministic checks, human judgments, and required source access.

Readiness is not verification. A later VVB determination is a separate receipt with actor, scope, criteria, assurance profile, outcome, qualifications, and date.

5.7 Contribution, rights, and co-claim allocation

Contribution receipts record financial, operational, scientific, or data contributions. Rights receipts record contractual or declared rights and restrictions. Neither automatically creates a quantitative claim.

For recognised outcome Q, allocation vector $\mathbf{a} = (a_1, \ldots, a_n)$ must satisfy:

$$a_i \geq 0, \qquad \sum_{i=1}^{n} a_i \leq Q.$$

Define unallocated quantity:

$$u = Q - \sum_{i=1}^{n} a_i, \qquad u \geq 0.$$

If the policy permits percentage shares w_i, then

$$a_i = w_i Q, \qquad w_i \geq 0, \qquad \sum_i w_i \leq 1.$$

The graph must preserve the policy version, allocation authority, contractual source, units, cohort, interval, claim purpose, and supersession. A mathematically valid allocation does not establish legal entitlement or institutional acceptance.

5.8 Reliance artifacts

For institution I, policy adapter P_v, and artifact A, Paper 2 defines:

$$\Omega(A, P_v, I) \in \{\texttt{accepted}, \texttt{relied\_on}, \texttt{rejected}, \texttt{needs\_more\_evidence}, \texttt{pending}, \texttt{not\_assessed}\}.$$

The same evidence root may generate different artifacts and reliance outcomes for a verifier, buyer, processor, lender, registry, or government-facing pilot. These outcomes coexist without changing historical receipts.

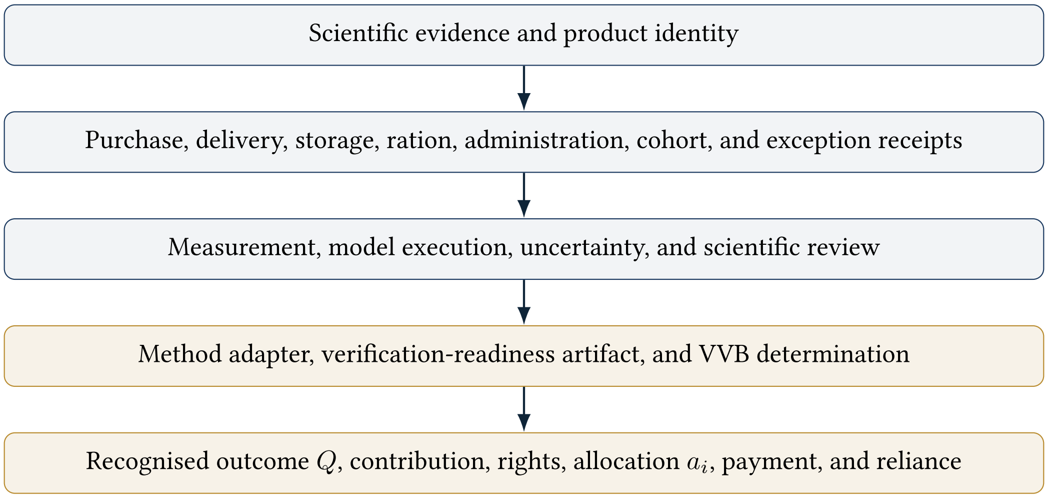

Figure 1: Aligned Paper 4 evidence pathway. Each layer is composed of Paper 1 receipts, evaluated with Paper 2 state separation and Paper 3 minimum-evidence requirements.

6 Methods

6.1 Study design

The study is an applied design-science evaluation using one bounded Australian beef or dairy deployment case. The approved data route will be one of:

1. a synthetic mirror reproducing the structural difficulty of a real case;
2. a permissioned and de-identified pilot subset; or
3. an approved post-pilot dataset released after rights and confidentiality review.

Synthetic data must be labelled in every receipt, manifest, artifact, figure, table, and result.

6.2 Case construction and graph assembly

The case will be assembled from source systems rather than reconstructed solely from a final spreadsheet. Each source object will receive a commitment, custodian, access condition, and transformation history. Receipt generation will occur as close to source as practicable.

Graph assembly will follow these steps:

1. bind producer, property, supplier, product, batch, devices, and cohorts;
2. capture authority, rights, and consent;
3. link scientific formulation to commercial product and batch;
4. ingest purchase, delivery, storage, mixing, feeding, and exception records;
5. construct time-indexed cohort membership;
6. ingest measurement or model inputs and provenance;
7. execute the declared model and preserve environment and outputs;
8. conduct scientific review and evidence-sufficiency classification;
9. evaluate the graph under an immutable method adapter;
10. assemble a verification-readiness bundle;
11. append contribution, rights, allocation, payment, and reliance events where available; and

12. freeze the bundle manifest and evidence root for independent reconstruction.

6.3 Paper 3 evidence-conformance mapping

Every required field will be classified as:

$$s_f \in \{\texttt{present}, \texttt{missing}, \texttt{conflicting}, \\ \texttt{out\_of\_domain}, \texttt{conditionally\_sufficient}, \texttt{not\_applicable}\}.$$

Domain sufficiency will not be calculated as a simple percentage unless the governing decision profile explicitly supports weighting. A critical missing field may be more consequential than many present optional fields.

6.4 Integrity and compatibility evaluation

The bundle will be evaluated using the Paper 2 topology:

$$\mathcal{V}(E, A) = \langle \Phi(E, S), \Psi(E, M_v), \Omega(A, P_v, I) \rangle .$$

Integrity tests include canonicalisation, signatures, trust material, parent resolution, source digests, and manifest consistency. Compatibility tests include applicability, required evidence, model domain, uncertainty, verification expectations, rights, and claim policy.

The conformance rules include:

$$\Phi = \texttt{valid} \nRightarrow \Psi = \texttt{eligible}, \qquad \Psi = \texttt{eligible} \nRightarrow \Omega = \texttt{accepted}.$$

6.5 Independent reconstruction

Independent reviewers will receive:

- the bounded bundle and manifest;
- public schemas and semantic documentation;
- adapter and method versions;
- verifier and reconstruction instructions;
- required software and environment manifest; and
- a task protocol.

They will not receive privileged access to the hosted workflow. Reviewers will attempt to reconstruct:

1. product and batch continuity;
2. delivery and administration;
3. cohort membership by interval;
4. dose and exposure pathway;
5. baseline and production context;
6. measurement or model inputs;
7. model execution and uncertainty;
8. method-compatibility result;
9. exceptions, corrections, and supersession;
10. recognised outcome Q; and
11. claim allocations and aggregate-cap conformance.

6.6 Exception and co-claim scenarios

At minimum, the test suite will include:

- missed feeding interval;
- product substitution or reformulation;
- missing certificate of analysis;
- dose outside target tolerance;
- cohort entry or exit during the claim period;
- unrecorded animal movement discovered later;
- missing or conflicting baseline;
- model execution outside calibration domain;
- model-version supersession;
- corrected source record;
- expired or revoked authority;
- overlapping allocation requests;
- aggregate allocation exceeding Q;
- buyer reliance with VVB review pending; and
- positive verification with buyer requesting additional rights evidence.

6.7 Evaluation metrics

Primary metrics are:

- proportion of required parent and source relationships resolved;
- number and type of Paper 3 evidence states;
- reconstruction time and clarification requests;
- deterministic agreement on integrity and machine-decidable compatibility checks;
- quantitative discrepancy in reconstructed inputs, outputs, and Q;
- number of hidden versus visible exceptions;
- number of unsupported inferences detected;
- allocation-cap pass, fail, and indeterminate outcomes;
- number of claim, rights, and reliance conflicts;
- producer and source-system capture burden; and
- reviewer disagreement and unresolved scientific judgment.

For reconstructed quantity $\widehat{Q}$, relative discrepancy may be reported as:

$$\delta_Q = \frac{|\widehat{Q} - Q|}{\max(|Q|, \epsilon)},$$

with declared $\epsilon > 0$ to avoid division by zero. A low discrepancy does not prove scientific validity if the same incorrect assumptions were reproduced.

7 Worked scenarios

7.1 Missed dosing interval

The target schedule specifies daily administration. Feed-mixer logs show no inclusion for two days. The graph appends an evidence-gap and exception receipt. The model must either quantify the interruption under the declared method or return conditional or insufficient evidence. The missed interval is not averaged away silently.

7.2 Product substitution

A substitute product is used after a supply interruption. A new product and batch branch is required. Scientific evidence for the original formulation cannot be inherited automatically. Compatibility may become `method_extension_required` or `insufficient_evidence`.

7.3 Cohort turnover

Animals enter and leave a feedlot pen during the reporting period. Time-indexed cohort membership determines exposure eligibility. A static end-of-period list would overstate or misassign exposure.

7.4 Model-version update

The same frozen evidence root is evaluated under M_v and M_{v+1}. Outputs and compatibility may differ, but the historical delivery, administration, cohort, and measurement receipts remain unchanged.

7.5 Corrected ration record

An operator discovers a unit error. A correction receipt references the original, records authority and reason, and supplies the corrected value. Both remain visible. The new model execution and determination link to the corrected branch.

7.6 Allocation exceeds outcome

A processor, retailer, and funder request quantitative allocations whose total exceeds Q. The allocation evaluator returns non-conformance:

$$\sum_i a_i > Q.$$

The system must not reduce or reassign claims automatically. It records the conflict and requires an authorised policy or contractual decision.

7.7 Scientifically credible, rights unresolved

The intervention, exposure, and model evidence are scientifically credible, but the contract does not establish who may make a Scope 3 claim. Integrity and compatibility may be positive while reliance remains `needs_more_evidence` or `pending`.

7.8 Positive VVB determination, different buyer policy

A VVB issues a bounded positive determination. A buyer rejects the artifact because its policy requires a different reporting boundary. The verification event remains valid; the buyer reliance event is separately `rejected`.

8 Planned results reporting

> **Working-draft note.** No graph evaluation or reconstruction has been completed. The following table is a reporting template, not findings.

Table 3: Paper 4 planned results template.

Analysis	Result	Required interpretation
Graph assembly	TBD	Report node and edge counts, source systems, missing parents, inaccessible sources, and synthetic content.
Paper 3 sufficiency mapping	TBD	Report states by domain and decision profile; do not collapse to one completion score without justification.
Integrity evaluation	TBD	Report valid, invalid, and indeterminate checks separately.
Method compatibility	TBD	Report exact method and adapter versions, conditions, gaps, and human judgments.
Independent reconstruction	TBD	Report time, clarifications, unresolved dependencies, and environment.
Quantitative reproduction	TBD	Report input, output, uncertainty, and Q discrepancies.
Exception visibility	TBD	Report detected, missed, disputed, and retrospectively discovered exceptions.
Method supersession	TBD	Report changed determinations and unchanged historical receipt digests.
Aggregate-cap tests	TBD	Report pass, fail, and indeterminate allocation scenarios.
Rights and claim review	TBD	Report unresolved authority, contract, overlap, or claim-language issues without making legal conclusions.
Reliance outcomes	TBD	Report institution, purpose, artifact, policy version, outcome, and conditions.
Producer burden	TBD	Report manual steps, duplicate entry, time, exception burden, and mitigation options.

9 Discussion

9.1 Trial efficacy is necessary but not sufficient

A trial establishes evidence under controlled conditions. Deployment requires identity continuity, operational records, cohort binding, model provenance, exception handling, and authority. The graph does not diminish trial evidence; it prevents trial evidence from being stretched beyond its domain.

9.2 Operational exceptions are evidence

Missed doses, substitutions, movements, and corrections are often treated as data-cleaning problems. In this architecture they are evidence. Recording them improves interpretability because a reviewer can determine whether the exception is immaterial, conditional, method-disqualifying, or unresolved.

9.3 One evidence root, multiple decisions

Paper 1's central claim is that historical evidence can be preserved while interpretations vary. Paper 4 operationalises that claim. The same root may support a scientific review, a method determination, a VVB artifact, a processor report, and a buyer reliance decision. Each artifact identifies versions, included evidence, exclusions, gaps, and limitations.

9.4 No validity laundering

Paper 2 requires that integrity, compatibility, verification, and reliance remain separate. A complete graph can still be scientifically out of domain. A method-compatible graph can still await verification. A positive VVB determination can still be rejected by a buyer for bounded policy reasons. The interface and exports must preserve those distinctions.

9.5 Minimum evidence as the applied contract

Paper 3 defines what must be available for a bounded decision. Paper 4 tests whether those requirements are implementable in real workflows. Fields that cannot be captured reliably should not be silently dropped. The applied study should determine whether a credible proxy, conditional state, method extension, or redesigned workflow is required.

9.6 Claim allocation is not outcome quantification

The aggregate-cap invariant is intentionally modest. It prevents quantitative over-allocation within the declared graph. It does not solve attribution, additionality, legal title, target accounting, public claim language, or interactions with registries. Those issues require versioned policy, contract, and institutional decisions.

9.7 Verification readiness and cost

A well-structured bundle may reduce reviewer time by exposing lineage, versions, and gaps. That is an empirical question. The study should report whether reconstruction effort actually falls, whether new burden is shifted upstream, and whether source access or confidentiality remains a bottleneck.

9.8 Open architecture and commercial implementation

The graph profile, receipt semantics, conformance fixtures, and limitations should remain independently inspectable. Commercial value may lie in source integrations, managed operations, exception handling, artifact assembly, and service commitments. Conformance should not require exclusive use of one hosted platform.

10 Limitations

This draft has no approved case dataset, completed field evaluation, verified outcome, recognised Scope 3 claim, or accredited determination.

The supplied bibliography includes entries whose metadata has not been independently verified. These must not be treated as submission-ready references until checked.

A synthetic mirror can test graph behaviour but not establish operational feasibility, legal adequacy, or institutional acceptance.

An append-only graph preserves history but does not prove the truth of source assertions. Capture controls, identity, authority, calibration, and fraud risk remain material.

The aggregate-cap invariant prevents one arithmetic form of over-allocation but does not prevent all forms of double claiming, such as differently worded claims over the same outcome, claims outside the graph, or inconsistent organisational boundaries.

Independent reconstruction may reproduce a shared incorrect assumption. Agreement is therefore not equivalent to scientific validity.

Producer burden may be uneven. A graph that improves downstream auditability could impose unacceptable upstream recording costs unless capture is integrated into existing workflows.

Rights, confidentiality, privacy, and commercial sensitivity may restrict source access. Such restrictions must be reported as limitations rather than hidden dependencies.

Finally, publication does not create an Australian Standard, approve a method, certify a product, establish Project 4030 endorsement, confer government acceptance, determine claim ownership, or replace an accredited validation or verification body.

11 Conclusion

Paper 4 translates the first three portfolio papers into one applied trial-to-farm-to-claim pathway. It uses the Paper 1 thin-waist receipt architecture, the Paper 2 validity topology, and the Paper 3 minimum-evidence vocabulary to represent feed-additive deployment as an append-only evidence graph.

The proposed graph preserves product identity, operational fidelity, cohort dynamics, measurement and model provenance, scientific review, verification readiness, outcome recognition, contribution, rights, allocation, payment, and reliance. It treats exceptions and corrections as evidence rather than defects to erase.

The aggregate-cap invariant separates a bounded recognised outcome Q from quantitative allocations a_i, preventing recorded over-allocation while preserving local authority over rights and claims. The next step is to select an approved data route, build the graph, preregister the reconstruction protocol, run independent review, publish conformance fixtures, disclose failures, and propagate architecture deltas through versioned portfolio change control.

12 Declarations

Funding and conflicts of interest

TODO: Disclose AgEvidence support, employment, consulting, advisory roles, equity, intellectual property, product relationships, pilot relationships, and other interests for every contributor.

Author contributions

TODO: Maintain a live CRediT contribution record. Confirm author order and corresponding-author responsibility only after written consent and final contribution review.

Data availability

TODO: Select synthetic mirror, permissioned pilot subset, or approved post-pilot route. Document permissions, de-identification, retention, confidentiality, and release conditions.

Code and specification availability

TODO: Archive schemas, graph profile, verifier, adapter versions, model environment, conformance fixtures, example bundles, checksums, and release history.

Rights and claims

No allocation, contribution, payment, or reliance receipt in the study should be represented as a legal determination. Legal and contractual review remains required.

Synthetic-case disclosure

Synthetic receipts, outcomes, allocations, and reliance events are test artifacts. They are not empirical findings or claims about named companies, products, farms, or institutions.

13 Executable TODO register

Table 4: Paper 4 executable TODO register.

Task	Owner	Exit evidence
Verify supplied bibliography and complete targeted review	Persona B and scientific editor	Publisher-checked library, evidence table, and revised claims.
Freeze Paper 4 graph profile against Papers 1–3	Persona G and portfolio authors	Versioned mapping, schemas, decision log, and dependency review.
Select approved case and data route	Data Steward and executive sponsor	Written permissions, de-identification plan, or synthetic protocol.

Task	Owner	Exit evidence
Build product-to-cohort graph	Technical and implementation teams	Versioned bundle, manifest, source commitments, and gap register.
Implement method and claim adapters	Technical team; scientific and legal review	Immutable adapter packages, tests, authority statements, and limitations.
Preregister reconstruction and analysis	Methods team	Archived protocol, outcomes, exclusions, discrepancy rules, and analysis code plan.
Create exception and allocation fixtures	Authors and reproducibility engineer	Machine-readable cases, expected states, and rationale.
Run independent reconstruction	Independent reviewers	Time logs, issue logs, outputs, discrepancies, and disagreement report.
Evaluate producer and source-system burden	Persona F and implementation team	Workflow map, time estimates, duplicate-entry count, and mitigation options.
Audit validity laundering in outputs	Persona C and interface reviewers	Dashboard, PDF, API, and export audit with issue register.
Resolve deltas through portfolio change control	Scientific Chairs	Versioned decisions propagated to Papers 1–3 and later papers.
Complete CRediT, disclosures, and final author approval	Persona B	Signed authorship, affiliation, conflict, and approval records.

References

Alvarez-Hess, P. S., P. J. Moate, S. R. O. Williams, J. L. Jacobs, K. A. Beauchemin, Z. Durmic, and R. J. Eckard (2019). "Efficacy of 3-Nitrooxypropanol in Reducing Enteric Methane Emissions from Dairy Cows Fed a Pasture-Based Diet". In: *Journal of Dairy Science* 102.6. Bibliographic details supplied for this working draft; verification pending, pp. 5141–5150.

Ascui, F. and T. F. Cojoianu (2019). "Redefining Carbon Accounting for the Climate-Resilient Economy". In: *Accounting, Auditing & Accountability Journal* 32.8, pp. 2186–2215.

Benet, J. (2014). "IPFS - Content Addressed, Versioned, P2P File System". In: *arXiv preprint arXiv:1407.3561*.

Black, J. L. et al. (2021). "Factors Affecting the Commercial Adoption of Enteric Methane Mitigation Technologies in Australian Livestock Industries". In: *Animal Production Science* 61.10. Bibliographic details supplied for this working draft; verification pending, pp. 1011–1025.

Buneman, P., S. Khanna, and W.-C. Tan (2001). "Why and Where: A Characterization of Data Provenance". In: *Database Theory – ICDT 2001*. Vol. 1973. Lecture Notes in Computer Science. Springer, pp. 316–330.

Capper, J. L. and D. E. Bauman (2013). "The Role of Productivity in Improving the Environmental Sustainability of Ruminant Production Systems". In: *Annual Review of Animal Biosciences* 1.1, pp. 469–489.

Crosby, S. A. and D. S. Wallach (2009). "Efficient Data Structures for Tamper-Evident Logging". In: *Proceedings of the 18th USENIX Security Symposium*, pp. 317–334.

Dijkstra, J., A. Bannink, J. France, E. Kebreab, and S. van Gastelen (2018). "Antimethanogenic Effects of 3-Nitrooxypropanol Depend on Dietary Composition in Dairy Cattle: A Meta-Analysis". In: *Journal of Dairy Science* 101.10, pp. 9041–9057.

Duin, E. C., T. Wagner, S. Shima, D. Prakash, B. Cronin, D. R. Yanez-Ruiz, S. Duval, R. Rumbeli, R. T. Stemmler, R. K. Thauer, and M. Kindermann (2016). "Mode of Action Uncovered for the Specific Reduction of Methane Emissions from Ruminants by the Small Molecule 3-Nitrooxypropanol". In: *Proceedings of the National Academy of Sciences* 113.22, pp. 6172–6177.

Garnsworthy, P. C., J. Craigon, J. H. Hernandez-Medrano, and N. Saunders (2012). "Variation among Individual Dairy Cows in Methane Measurements Made on Farm during Milking". In: *Journal of Dairy Science* 95.6. The supplied citation used a different title and page range; this working record should be verified against the publisher version, pp. 3181–3189.

Gopalakrishnan, S. et al. (2023). "Preventing Double Counting in Scope 3 Carbon Accounting across Interconnected Agricultural Supply Chains". In: *Environmental Science & Technology* 57.14. Bibliographic details supplied for this working draft; verification pending, pp. 5821–5831.

He, J. et al. (2020). "Graph-Based Data Lineage Tracking for Distributed Scientific Workflows". In: *Future Generation Computer Systems* 108. Bibliographic details supplied for this working draft; verification pending, pp. 381–393.

Intergovernmental Panel on Climate Change (2019). *2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories: Volume 4, Agriculture, Forestry and Other Land Use*. IPCC.

International Auditing and Assurance Standards Board (2013). *International Standard on Assurance Engagements 3000 (Revised): Assurance Engagements Other than Audits or Reviews of Historical Financial Information*. International Federation of Accountants.

International Organization for Standardization (2019). *ISO 14064-3:2019: Greenhouse Gases – Part 3: Specification with Guidance for the Verification and Validation of Greenhouse Gas Statements*. International Organization for Standardization.

Kinley, R. D., G. Martinez-Fernandez, M. K. Matthews, R. de Nys, M. Magnusson, and N. W. Tomkins (2020). "Mitigating Enteric Methane Emissions from Cattle Using Asparagopsis taxiformis". In: *Journal of Cleaner Production* 259, p. 120836.

Kruse, L. et al. (2021). "Design Principles for Enterprise Data Lineage in Complex Supply Chain Ecosystems". In: *Information Systems Frontiers* 23.4. Bibliographic details supplied for this working draft; verification pending, pp. 945–963.

Lebo, T., S. Sahoo, D. McGuinness, K. Belhajjame, J. Cheney, D. Corsar, D. Garijo, S. Soiland-Reyes, S. Zednik, and J. Zhao (2013). *PROV-DM: The PROV Data Model*. W3C Recommendation. World Wide Web Consortium.

Moreau, L. (2010). "The Foundations for Provenance on the Web". In: *Foundations and Trends in Web Science* 2.2-3, pp. 99–241.

Oldfield, E. E. et al. (2022). "Credit-Readiness and Verification Standards for Agricultural Soil Carbon and Greenhouse Gas Markets". In: *Environmental Research Letters* 17.12. Bibliographic details supplied for this working draft; verification pending, p. 124006.

Romero-Perez, A., E. K. Okine, S. M. McGinn, L. L. Guan, M. Oba, S. M. Duval, M. Kindermann, and K. A. Beauchemin (2014). "The Potential of 3-Nitrooxypropanol to Lower Enteric Methane Emissions from Beef Cattle". In: *Journal of Animal Science* 92.10, pp. 4682–4693.

Science Based Targets initiative (2022). *Forest, Land and Agriculture (FLAG) Target Setting Guidance*. Science Based Targets initiative.

Tönissen, S. and F. Teuteberg (2020). "Understanding the Impact of Blockchain Technology in Agricultural Supply Chain Management: A Systematic Literature Review". In: *Journal of Enterprise Information Management* 33.2, pp. 425–482.

Vanderhaegen, K. et al. (2021). "Accounting for Indirect Land-Use Change and Scope 3 Supply Chain Interventions: Methodological Gaps and Solutions". In: *Global Environmental Change* 68. Bibliographic details supplied for this working draft; verification pending, p. 102262.

World Resources Institute and World Business Council for Sustainable Development (2011). *Corporate Value Chain (Scope 3) Accounting and Reporting Standard*. WRI/WBCSD.

Source note

This expanded draft is aligned to the attached expanded first drafts of Paper 1, Paper 2, and Paper 3. It reuses their architecture, state vocabulary, evidence states, guardrails, authorship discipline, planned-results convention, and dependency model. The additional Paper 4 literature was supplied for integration. Bibliographic accuracy and source suitability must be independently verified before submission.

Paper 5: Whole-Farm Grazing Systems

Working-draft note. This manuscript expands Paper 5 from the *AgEvidence Australia Scientific Publication Series: Seven Manuscript First Drafts*. It is a whole-farm systems, evidence-architecture, and participatory-methods manuscript. No producer workflow study, burden assessment, remote-sensing integration, temporal-alignment evaluation, model validation, regional sufficiency determination, verifier decision, or institutional reliance outcome has been completed. All cases, measures, thresholds, results, and claims remain proposed.

Non-negotiable guardrail. The proposed whole-farm graph records evidence, transformations, rights, limitations, and bounded decisions. It does not itself establish an Australian Standard, approve a greenhouse-gas methodology, prove a physical emissions reduction, certify a farm or product, determine legal ownership of data or environmental attributes, replace accredited validation or verification, or compel reliance by a government agency, registry, buyer, lender, processor, insurer, or other institution.

Abstract

Extensive grazing systems present distinct verification and data-governance challenges compared with contained agricultural operations. Ruminant herds move dynamically across variable paddock topographies, exposure to methane-mitigating dietary interventions may be intermittent, environmental conditions fluctuate, and direct emissions measurement is sparse. Consequently, greenhouse-gas monitoring, reporting, and verification systems must rely on heterogeneous data streams spanning producer-generated operational logs, remote-sensing products, spatial property records, weather observations, livestock records, and empirical or process-based emissions models.

This paper formalises a whole-farm evidence model tailored to Australian grazing beef systems. The framework structures producer and organisational identity, land parcels and paddocks, herd and cohort movement, management and supplementation events, environmental proxy observations, model executions, scientific reviews, rights, consent, contractual allocations, and institutional decisions into a temporal graph schema. Rights, consent, data-use authority, land control, herd authority, and claim permissions are encoded as primary evidence primitives rather than unstructured metadata.

The paper establishes explicit temporal-alignment rules for combining asynchronous and multi-frequency evidence without silent interpolation. Each transformation declares source interval, target interval, subject mapping, spatial support, aggregation rule, imputation or interpolation method, calibration domain, uncertainty, and conditional assumptions. Remote-sensing observations remain distinct from ground observations, model estimates, and producer assertions. A successful model execution is separated from scientific applicability and from institutional acceptance.

The proposed evaluation combines participatory workflow mapping, source-system inventory, graph construction, temporal-alignment tests, scenario testing, independent reconstruction, and producer burden assessment. Burden is treated as a direct performance metric through setup time, recurring manual effort, duplicate entry, missing-data frequency, reconstruction latency, clarification requests, and artifact utility for producers, scientists, verifiers, buyers, and financial entities.

The intended outputs are a grazing evidence ontology, whole-farm graph profile, temporal-alignment protocol, rights and consent profile, minimum producer dataset, burden measurement framework, uncertainty disclosure model, and conformance cases. The framework is designed to distinguish directly observed operational facts from proxy, calculated, modelled, assumed, scientifically judged, verified, and institutionally relied-on propositions. It does not assert that one profile is sufficient for every Australian region, production system, method, or decision.

Keywords: grazing beef; whole-farm systems; evidence graph; data governance; producer burden; remote sensing; temporal alignment; greenhouse-gas MRV; land identity; herd movement.

1 Manuscript status, scope, and contribution boundary

This manuscript proposes a reference profile and an evaluation plan. It does not report participant recruitment, farm records, data completeness, burden measurements, remote-sensing accuracy, model-domain outcomes, reconstruction time, verifier conclusions, buyer acceptance, financing outcomes, or national representativeness. Tables labelled as planned results are reporting templates rather than findings.

The paper depends on Paper 1 for the country-neutral receipt envelope, append-only evidence graph, parent and source commitments, identity and authority declarations, versioned adapters, and bounded reliance artifacts. It depends on Paper 2 for the separation of cryptographic integrity, method compatibility, scientific judgment, accredited validation or verification, and institutional reliance. It depends on Paper 3 for minimum-evidence domains, gap states, measurement and model provenance, source-capture priority, and decision-linked sufficiency. Paper 5 applies those concepts to extensive grazing beef and should raise any required changes through portfolio change control rather than redefine shared terms locally.

The paper addresses a whole-farm evidence problem rather than a single intervention or calculator. Its scope includes:

1. producer, organisation, property, parcel, paddock, herd, cohort, and animal identity;
2. authority to act, assert, disclose, calculate, transfer, rely, receive payment, or claim;
3. herd movements, grazing allocation, supplementation, interventions, purchases, sales, births, deaths, treatments, and management changes;
4. weather, pasture, water, land condition, productivity, and remote-sensing observations;
5. methane and whole-farm emissions models, transformations, calibration domains, and uncertainty;
6. scientific review, method compatibility, verification readiness, and institution-specific artifacts;
7. producer burden, usability, missing-data frequency, and reconstruction latency.

The contribution is a testable evidence architecture and participatory evaluation method. It is not a new emissions-reduction method, a national inventory model, a remote-sensing product validation, a land-tenure determination, a herd-ownership registry, or a legal opinion.

2 Provisional scientific author personas

Persona	Provisional role	Contribution domain / deliverable
Persona F	Lead/corresponding-author persona	Participatory workflow mapping, grazing implementation, producer burden and practical interpretation.
Persona A	Scientific coauthor persona	Australian whole-farm accounting, climatic/regional context and local applicability.
Persona D	Scientific coauthor persona	Measurement quality, remote-sensing/proxy distinctions, uncertainty and limitation reporting.
Persona B	Scientific coauthor persona	Livestock-emissions modelling, calibration domain, sensitivity and model applicability.
Persona G	Technical scientific contributor persona	Temporal graph, identity schema, alignment rules, conformance tests and reconstruction package.
Persona H	Governance contributor persona	Producer rights, consent, geospatial controls, retention, disclosure and data-use authority.
Persona I	Invited domain-review persona	Grazing science, producer operations, remote sensing, extension, verification, legal or institutional use.

Authorship control. Persona IDs are recruitment placeholders only. A real person is not an author until consent, qualifying contribution, affiliation, conflict disclosure, CRediT allocation, and final-manuscript approval are complete.

3 Introduction

3.1 The containment assumption

Many agricultural data systems are designed as if the farm were a stable facility. They assume fixed geographic boundaries, continuous or daily records, known animal groups, uniform feed delivery, reliable connectivity, and frequent measurement. These assumptions can be reasonable in contained feeding or highly instrumented operations. In extensive grazing, they become a source of hidden error.

Herds move across paddocks and sometimes properties. Group membership changes through births, purchases, sales, deaths, treatment, weaning, joining, branding, drafting, and seasonal management. Supplementation may occur at irregular intervals and may not imply uniform consumption. Weather and forage conditions change continuously, but direct observations may be episodic. Satellite products observe spatial surfaces on overpass schedules and under cloud constraints. Farm records are created around operational decisions rather than research sampling frames. Direct methane measurement is sparse, so whole-farm outcomes are often inferred through empirical equations or simulation models.

A static table can force these conditions into daily rows, but the resulting completeness may be artificial. Missing intervals become copied values. herd composition is treated as constant. A remote-sensing value for one date is extended across a month. A rainfall station is assumed to represent a distant paddock. A supplement delivery is treated as animal consumption. A model estimate is displayed as an observation.

This paper calls the underlying design error the *containment assumption*: the presumption that extensive systems can be represented faithfully by schemas built for stable, bounded, continuously observed operations. The proposed alternative is a temporal evidence graph in which identities, rights, events, observations, transformations, models, and decisions retain their own spatial and temporal support.

3.2 Metadata marginalisation

A second architectural flaw is metadata marginalisation. Land control, herd ownership, producer consent, data-use permission, contractual claim rights, and restrictions are frequently stored as attachments, free-text notes, or external contracts. They are treated as administrative context rather than computational evidence.

That separation is risky because the actors associated with land, livestock, management, data, models, claims, and payment may differ. A producer may operate land owned by another entity. Livestock may be agisted, financed, leased, or managed under contract. A remote-sensing provider may license derived data but not raw imagery. A buyer may fund an intervention without owning the underlying farm data. A processor may wish to rely on an outcome while contractual rights remain disputed.

The proposed architecture treats authority and rights as graph primitives. A data item is not merely “from Farm A.” It is linked to an issuer, role, authority basis, subject, permitted uses, restrictions, period, and supersession history. This does not resolve legal ownership automatically. It exposes the evidence and unresolved questions that a competent institution must evaluate.

3.3 Whole-farm evidence as a dynamic graph

The proposed graph is:

$$\mathcal{G} = (\mathcal{V}_{\text{ident}}, \mathcal{V}_{\text{event}}, \mathcal{V}_{\text{rights}}, \mathcal{V}_{\text{model}}, \mathcal{V}_{\text{artifact}}, \mathcal{E}_{\text{temporal}}, \mathcal{E}_{\text{spatial}}, \mathcal{E}_{\text{provenance}}).$$

The user-supplied core formulation is preserved:

$$\mathcal{G}_0 = (\mathcal{V}_{\text{ident}}, \mathcal{V}_{\text{event}}, \mathcal{V}_{\text{rights}}, \mathcal{V}_{\text{model}}, \mathcal{E}_{\text{temporal}}),$$

with additional spatial, provenance, and artifact relations made explicit for implementation testing.

The graph is not a map of every biological process. It is an evidence graph: a record of bounded assertions, observations, transformations, and decisions. Nodes have issuers, subjects, valid times, record times, sources, authority, limitations, and versions. Edges declare relationships rather than relying on implied joins.

3.4 Observed facts and conditional claims

A producer may observe that a herd entered a paddock on a date. A remote-sensing product may estimate fractional cover over a polygon. A model may infer pasture biomass. An emissions algorithm may calculate methane from dry-matter intake. A scientist may judge the model applicable. A verifier may reach a determination under a defined engagement. A buyer may rely on a resulting artifact.

These are not equivalent statements. Their separation is essential because uncertainty grows as evidence moves farther from direct observation. The framework uses explicit evidence classes:

- producer-asserted operational event;

- instrument-derived observation;
- remote-sensing-derived observation or product;
- administrative or commercial record;
- calculated value;
- model output;
- proxy;
- assumption;
- scientific judgment;
- method-compatibility determination;
- accredited validation or verification determination;
- institutional reliance event.

3.5 Producer burden as a system-performance metric

A system can be semantically elegant and operationally unusable. Excessive manual entry, repeated identity reconciliation, duplicate buyer requests, fragmented applications, poor connectivity, and retrospective reconstruction can reduce participation and degrade data quality.

Producer burden is therefore a direct evaluation metric rather than an implementation footnote. The study will distinguish setup burden from recurring burden and record manual time, number of interactions, duplicate entry, missing-data resolution, clarification requests, correction effort, and artifact usefulness. An architecture that achieves apparent completeness only through intensive manual intervention should not be described as scalable.

3.6 Research questions

The manuscript addresses seven primary questions:

1. Which stable identities, events, rights, observations, transformations, models, and artifacts are required to reconstruct whole-farm grazing evidence?
2. Can asynchronous and multi-frequency evidence be aligned without silent interpolation or loss of source uncertainty?
3. Which records already arise in ordinary producer operations, which can be captured automatically, which require specialist measurement, and which are unrealistic?
4. Can rights, consent, authority, restrictions, and claim permissions be represented as first-class graph primitives without implying legal determinations?
5. Can independent reviewers reconstruct declared herd, land, intervention, model, and decision pathways from a frozen graph?
6. How do setup burden, recurring burden, missing-data frequency, reconstruction latency, and duplicate requests vary by workflow and source system?
7. Do resulting artifacts remain useful and comprehensible to producers, scientists, verifiers, buyers, lenders, and payment administrators without validity laundering?

3.7 Novelty relative to common agricultural MRV architectures

Table 2: Novelty analysis for the whole-farm grazing evidence model.

Dimension	**Common architecture**	**Proposed approach**

System boundary	Fixed facility, daily table, or annual farm summary.	Versioned property, parcel, paddock, herd, cohort, and event intervals with spatial and temporal relations.
Containment assumption	Assumes continuous inputs and stable groups.	Represents movement, intermittency, seasonal gaps, cohort change, and irregular observation explicitly.
Rights	Free-text metadata, separate contract, or implicit ownership.	Rights, authority, consent, restrictions, withdrawal, claims, and payment instructions are graph primitives.
Temporal alignment	Silent forward-fill, interpolation, or preprocessing.	Every alignment declares source and target intervals, mapping, method, assumptions, uncertainty, and domain.
Remote sensing	Treated as ground truth or a generic external layer.	Preserves product identity, acquisition, processing, spatial support, calibration, cloud and quality flags, and ground relationship.
Models	Output stored without executable lineage or domain.	Model identity, version, code, inputs, transformations, calibration domain, sensitivity, uncertainty, and review state are linked.
Evidence type	Observed and estimated values mixed in one field.	Direct observation, instrument, remote sensing, proxy, calculation, model, assumption, and judgment remain distinct.
Burden	Qualitative implementation note.	Setup time, recurring manual work, duplicate entry, missingness, reconstruction latency, and clarification are core outcomes.
Change	Historical records overwritten to maintain clean tables.	Corrections and supersession are append-only and temporally explicit.
Decision use	One generic “verified” output.	Separate scientific, method, verification, contractual, payment, and institution-specific artifacts.

4 Scientific and methodological foundations

4.1 Grazing greenhouse-gas emissions and extensive-system MRV

Australian grazing systems combine seasonal forage variation, uneven intake, large spatial extent, changing herd composition, and sparse direct methane measurement. Reviews of mitigation options in ruminant production describe the interaction among diet quality, intake, animal performance, management, and methane outcomes under pastoral conditions [7]. Research in northern Australian beef systems further demonstrates that native and improved pastures, molasses, and protein supplementation influence methane production and the interpretation of seasonal baselines [3].

Direct measurement methods developed for controlled experiments face logistical constraints in free-ranging systems. The supplied literature compilation identifies a review by Jonker and colleagues as addressing SF_6, automated systems, chamber methods, and future directions for grazing ruminants [10]. The final manuscript must verify this bibliographic record before relying on it. Regardless of method, the architecture must preserve sampling population, animal binding, temporal coverage, missingness, device or tracer provenance, and the relationship between measured animals and the modelled herd.

The evidence implication is not that grazing emissions are unknowable. It is that conclusions often depend on a combination of ordinary records, sparse measurements, empirical relationships, and models. Each component must remain visible.

4.2 Whole-farm modelling and environmental accounting

Whole-farm studies show that emissions outcomes emerge from interactions among forage availability, herd structure, reproduction, growth, sale timing, supplementation, climate, land condition, and enterprise mix. Harrison and colleagues examine options for managing emissions in extensive Australian grazing systems using system-level analysis [8]. Browne and colleagues apply a sheep and beef greenhouse accounting framework across southern Australian enterprises [2]. Moore and colleagues use whole-farm bio-economic modelling to evaluate feed-base options and climate-driven dynamics [11].

These approaches motivate a key distinction: a whole-farm model is not simply a calculator attached to annual totals. It is an interpretation layer that consumes temporally structured evidence. The graph must preserve which herd state, land state, feed state, climate series, and management assumptions were used in each run. A later model version or policy adapter should be able to re-interpret the same historical evidence without rewriting it.

4.3 Remote sensing and geospatial integration

Remote sensing can extend coverage across large pastoral landscapes, but satellite-derived products remain measurements and models with their own spatial support, acquisition frequency, calibration, atmospheric effects, cloud constraints, processing versions, and uncertainty. The supplied literature identifies work on vegetation biomass estimation in pastoral systems [5], a commercial-scale Australian forage-budgeting project [4], and spatio-temporal alignment of satellite and ground observations [12].

A remote-sensing evidence record should therefore identify:

- sensor or product family, provider, version, and processing level;
- acquisition time, observation window, revisit interval, and quality flags;
- spatial footprint, resolution, parcel overlay, and edge treatment;
- atmospheric, cloud, shadow, water, and bare-ground handling;
- model or algorithm converting spectral information into biomass, cover, condition, or water;
- ground calibration or validation domain;
- aggregation and temporal alignment;
- uncertainty, missingness, and limitations.

A bi-monthly pasture product cannot be treated as continuous daily observation without a declared transformation. A pixel covering mixed land classes cannot be assumed to describe a paddock uniformly. Spatial alignment is an evidence operation.

4.4 Farm data governance and producer authority

Trust in digital agriculture depends not only on privacy language but on clear control, permitted use, transparency, portability, and contractual fairness. The Australian Farm Data Code provides an industry governance reference for consent, identifying and non-identifying farm data, use, portability, and responsible handling [13]. Wiseman and colleagues examine farmer trust and data-sharing frameworks [16], while Sanderson and colleagues analyse agricultural data ownership and contractual control in Australia [14].

These sources support a rights-first architecture but do not imply that data are a single legally owned asset. Different legal and contractual interests may attach to source records, databases, imagery, models, derived products, confidential information, personal information, geospatial information, and environmental attributes. The graph should record asserted authority, contractual basis, restrictions, and unresolved conflicts without presenting them as final legal determinations.

4.5 Producer burden and technology adoption

Digital technology adoption is shaped by perceived value, connectivity, interoperability, cost, trust, skills, and administrative burden. The supplied literature identifies double entry, fragmentation, and poor interoperability as material barriers in Australian broadacre and grazing systems [9]. It also identifies work on managing agtech transitions and user burden in pastoral systems [6].

Burden affects evidence quality. When required records do not align with operational decisions, they are often entered retrospectively, reconstructed from memory, duplicated across portals, or omitted. The study will therefore measure burden and examine whether architecture design reduces reconstruction rather than merely redistributes it.

4.6 Temporal graphs, provenance, and model-domain control

Temporal graph methods allow entities and relationships to change over valid time while preserving record history. The supplied compilation cites a proposed agricultural temporal-graph paper by Bonner and Thompson [1] and a model error-propagation paper by Vargas and Lopez [15]. These records must be independently verified before submission and should not be treated as established sources until confirmed.

The architecture is also grounded in the portfolio's broader provenance and model-governance principles: historical evidence is immutable; interpretation is versioned; model outputs retain inputs and assumptions; and missing evidence is reported rather than silently repaired.

Working-draft note. The references in this section were supplied for integration and have not been independently verified in this drafting step. The final manuscript must validate titles, authors, venues, years, pagination, report identifiers, and interpretive claims against publisher or official records. Potentially unpublished, future-dated, institutional, or difficult-to-locate sources require particular scrutiny.

5 Whole-farm evidence model

5.1 Graph definition

Let the graph at version v be:

$$\mathcal{G}_v = (\mathcal{V}, \mathcal{E}, \tau, \pi, \rho, \lambda),$$

where:

- $\mathcal{V}$ is the set of evidence, identity, rights, model, and artifact nodes;
- $\mathcal{E}$ is the set of typed relations;
- τ maps nodes and edges to valid-time and record-time intervals;
- π maps nodes and edges to spatial support;
- ρ records issuer, role, authority, rights, and restrictions;
- λ records provenance, parentage, source commitments, limitations, and version.

The node set is partitioned as:

$$\mathcal{V} = \mathcal{V}_{\text{ident}} \cup \mathcal{V}_{\text{event}} \cup \mathcal{V}_{\text{rights}} \cup \mathcal{V}_{\text{obs}} \cup \mathcal{V}_{\text{model}} \cup \mathcal{V}_{\text{review}} \cup \mathcal{V}_{\text{artifact}}.$$

The edge set includes temporal, spatial, membership, authority, source, derivation, transformation, review, allocation, reliance, and supersession relationships.

5.2 Identity primitives

Identity nodes include:

- producer organisation and natural person where required;
- property, land parcel, paddock, and management unit;
- herd, cohort, and individual animal where necessary;
- livestock owner, landholder, lessee, agistor, manager, contractor, advisor, and veterinarian;
- intervention provider, feed supplier, laboratory, measurement provider, model operator, verifier, buyer, lender, insurer, processor, payer, and program;
- remote-sensing provider, weather station, device, software package, and dataset.

Local identifiers are bound to stable subjects through explicit namespace and authority. The model must support identity uncertainty, mergers, splits, aliases, deactivation, and correction.

5.3 Land and spatial primitives

Land evidence includes property identifiers, parcel geometries, paddock boundaries, management units, tenure or control records, land-use class, exclusions, water points, infrastructure, and boundary changes. Geometry is versioned because paddocks can be subdivided, merged, renamed, or re-fenced.

A spatial relation should state whether a node:

- is located within a polygon;
- intersects a polygon;
- is nearest to a location;
- represents an aggregate over multiple polygons;
- is used as a proxy for another area;
- is outside the declared spatial domain.

5.4 Herd and cohort primitives

A herd or cohort is represented through membership events rather than one static count. Candidate events include birth, purchase, sale, death, transfer, split, merge, joining, calving, weaning, treatment, hospitalisation, supplementation group, measurement sample, and exit.

For cohort C, membership at time t is reconstructed as:

$$M_C(t) = M_C(t_0) + \sum_{e \in A_C,\, t_e \le t} q_e - \sum_{e \in R_C,\, t_e \le t} q_e,$$

where A_C and R_C are addition and removal events and q_e is the quantity or set of identified animals. The expression is a reconstruction rule, not a claim that all farms track individual animals continuously.

5.5 Management and intervention events

Management events include paddock allocation, mustering, supplementation, intervention delivery, feed placement, water management, burning, pasture improvement, fertilisation, veterinary treatment, breeding, weaning, sale, purchase, transport, and exceptional events such as flood, fire, drought, disease, or infrastructure failure.

Each event records:

- bounded proposition;
- subject and location;
- valid-time interval and record time;
- issuer, role, and authority;
- quantity, unit, method, and source;
- linked documents or observations;
- limitations, uncertainty, and exceptions;
- correction or supersession relation.

5.6 Observed, proxy, calculated, and modelled evidence

Evidence classification is mandatory. A field must not carry a methane, biomass, intake, or production value without an evidence class.

Table 3: Evidence classes in the grazing graph.

Class	Meaning	Required disclosure
Producer observation	Operational assertion recorded by an authorised producer or agent.	Observer, time, subject, method, contemporaneity, limitations.
Instrument observation	Device-mediated measurement.	Device, calibration, configuration, operator, sampling, QC, uncertainty.
Remote-sensing product	Spatial observation or model-derived product.	Sensor, acquisition, processing, spatial support, calibration, quality flags, uncertainty.
Administrative record	Commercial, registry, movement, transaction, contract, or authority record.	Issuer, authority, subject, date, scope, restrictions.
Calculation	Deterministic transformation.	Inputs, formula, units, factors, rounding, software, version.
Model output	Empirical, statistical, process, or simulation output.	Model, version, inputs, domain, assumptions, sensitivity, uncertainty, review state.
Proxy	Evidence used to represent another variable.	Target variable, causal or empirical basis, validation, domain, uncertainty.
Assumption	Declared value or rule used because evidence is absent.	Rationale, owner, period, consequence, alternatives.
Scientific judgment	Expert assessment of applicability or sufficiency.	Reviewer, expertise, question, evidence, reasoning, limitations.
Institutional decision	Method, verification, contractual, payment, or reliance event.	Institution, authority, purpose, artifact, criteria, conditions, expiration.

5.7 Rights-first graph architecture

Rights nodes capture:

- authority to act for a producer organisation;
- authority to manage or assert events on land;
- authority over a herd or cohort;
- consent to collect, use, disclose, link, derive, and retain data;
- licence or contractual restrictions on remote-sensing, software, model, or derived data;
- right to submit evidence to a method, verifier, program, buyer, lender, or registry;
- asserted contractual claim rights, co-claim rules, and exclusivity;
- payment instructions and beneficiary identity;
- withdrawal, expiry, revocation, dispute, and supersession.

A rights node has a subject, grantor, grantee, permitted action, purpose, data or outcome scope, spatial scope, temporal scope, restrictions, evidence basis, and status. The graph records an assertion of authority; it does not determine legal validity autonomously.

5.8 Model nodes and executions

Model identity and execution are separate. A model node describes the reusable scientific object. An execution node describes one run.

The model node includes purpose, owner, version, code or executable commitment, documentation, calibration population, spatial and temporal resolution, input schema, known limitations, validation evidence, and applicable production systems.

The execution node includes exact inputs, input sources, transformations, missing-data treatments, parameter values, random seeds where relevant, environment, timestamp, outputs, uncertainty, sensitivity, reviewer state, and parent evidence root.

5.9 Artifacts, reviews, and decisions

Artifacts package selected evidence for a bounded purpose. Examples include producer summary, scientific review packet, model report, method-compatibility artifact, verifier evidence packet, buyer report, lender sustainability artifact, payment statement, or claim-allocation record.

Each artifact identifies:

- evidence root and included nodes;
- purpose, institution, policy or method version;
- exclusions, assumptions, limitations, and unresolved gaps;
- transformation and rendering version;
- author or issuer;
- review, verification, and reliance state;
- validity period and supersession.

6 Temporal and spatial alignment

6.1 Event and observation intervals

For event e_j, let valid time be:

$$T(e_j) = [t_j^0, t_j^1],$$

and let subject and spatial support be $S(e_j)$ and $P(e_j)$. For model run m_k, define target interval T_k, target subjects S_k, target geography P_k, required evidence classes, aggregation rules, missing-data tolerance, and calibration domain.

Evidence is compatible only where temporal, subject, spatial, and semantic mappings are explicit.

6.2 Alignment record

Every transformation from source x to target y creates an alignment record:

$$A_{x \to y} = \langle T_x, T_y, S_x, S_y, P_x, P_y, h, \theta, U, L \rangle,$$

where h is the transformation or aggregation function, θ contains assumptions and parameters, U is uncertainty, and L is limitations.

Examples include:

- converting a monthly herd inventory to daily average numbers;
- allocating rainfall from a station to paddocks;
- applying a satellite biomass observation across a grazing interval;
- converting liveweight records into average liveweight by cohort;
- mapping a supplement delivery to expected exposure;
- applying seasonal diet quality to an emissions equation.

6.3 No silent interpolation invariant

The architecture enforces:

$$\text{derived}(y, x) \Rightarrow \exists A_{x \to y}.$$

If a target value is derived from a source at a different frequency, subject, or spatial support, an alignment record must exist. Copying, forward-filling, interpolation, smoothing, spatial averaging, or imputation without a declared transformation is non-conforming.

6.4 Temporal coverage states

For source x and target interval T, coverage is:

$$C(x,T) \in \{\texttt{complete}, \texttt{partial}, \texttt{intermittent}, \texttt{missing}, \texttt{conflicting}, \texttt{not\_applicable}\}.$$

Coverage does not imply quality or applicability. A continuous but poorly calibrated source can be complete and scientifically unsuitable. A sparse source can be partial but informative.

6.5 Remote-sensing alignment

Remote-sensing alignment should preserve acquisition dates rather than pretending observations are continuous. Where two observations bracket a grazing interval, the chosen interpolation or state-space model must be explicit. Where cloud, fire scar, water, mixed pixels, or land-cover change compromise retrieval, the quality state and fallback rule must remain visible.

A paddock aggregate should identify pixel inclusion, edge weighting, masking, resolution, and geometry version. If paddock boundaries changed between acquisitions, the graph should preserve both geometry versions and the chosen comparison rule.

6.6 Herd-land alignment

A herd-land relation records when a herd or cohort occupied a paddock or management unit. Uncertain entry or exit dates should be represented as intervals or bounded uncertainty rather than false precision.

For cohort C and paddock p, exposure duration is:

$$D(C,p) = \mu\left(T(C) \cap T(p) \cap T(\text{occupancy}_{C,p})\right),$$

where μ is interval measure. This duration may support subsequent forage, supplement, or environmental allocations only under declared assumptions.

6.7 Alignment gap handling

When alignment fails, the system should report one of:

- source absent;
- interval mismatch;
- subject mismatch;
- spatial mismatch;
- semantic mismatch;
- calibration-domain mismatch;
- conflicting sources;
- transformation unavailable;
- uncertainty exceeds threshold;
- method extension required.

The gap may be resolved through additional evidence, an authorised proxy, conditional use, revised scope, or a decision not to proceed. It must not disappear through hidden preprocessing.

7 Producer workflow and burden model

7.1 Workflow layers

The study distinguishes five layers:

1. **ordinary operations**: records generated because the farm must operate;
2. **existing compliance and commercial records**: movement, transaction, assurance, tax, finance, animal health, and customer records;
3. **automated evidence capture**: records emitted by devices, farm software, remote sensing, weather services, and integrations;
4. **new MRV-specific capture**: records created solely to support the declared evidence process;
5. **specialist measurement and review**: scientific measurement, modelling, audit, verification, and legal or institutional review.

The goal is to minimise layer four without pretending that layers one through three contain every required field.

7.2 Burden metrics

For participant i, total burden over study period T is:

$$B_i(T) = B_i^{\text{setup}} + B_i^{\text{recurring}}(T) + B_i^{\text{exception}}(T) + B_i^{\text{reconstruction}}(T).$$

Each component is measured through time, interactions, duplicate entries, cognitive steps, external support, cost, and delay.

Candidate primary metrics include:

- setup hours and number of systems connected;
- recurring manual minutes per week or operational event;
- number of duplicate data entries;
- proportion of required fields captured automatically;
- missing-field frequency;
- exception frequency and resolution time;
- clarification requests per artifact;
- retrospective reconstruction latency;
- producer correction and dispute effort;
- time to retrieve source evidence;
- perceived usefulness and comprehensibility.

7.3 Missing data as an operational outcome

Missingness is not only a data-quality statistic. It can reveal that a requirement is poorly aligned with farm operations. The study should classify missing data by cause:

- event not observed;
- event observed but not recorded;
- record exists but is inaccessible;
- identity cannot be reconciled;
- source system does not expose data;

- connectivity or device failure;
- producer declined or lacked authority;
- field definition was unclear;
- record was requested retrospectively;
- source frequency was incompatible;
- data were lost, overwritten, or out of retention.

7.4 Artifact utility

Artifacts will be tested with producers, scientists, verifiers, buyers, lenders, and payment administrators. Utility questions include:

- Can the recipient identify what is observed, modelled, assumed, and institutionally decided?
- Can the producer see obligations, limitations, payment basis, and correction routes?
- Can a scientist inspect model domain, measurement provenance, and temporal alignment?
- Can a verifier trace source records and exceptions?
- Can a buyer or lender understand the bounded purpose without inferring universal approval?
- Does the artifact expose rights, consent, restrictions, and unresolved claim conflicts?

8 Methods

8.1 Study design

The proposed mixed-method study contains seven linked phases:

1. literature and requirements review;
2. participant and system sampling;
3. participatory workflow mapping;
4. source inventory and graph construction;
5. temporal and spatial alignment testing;
6. burden, usability, and artifact evaluation;
7. independent reconstruction and scenario testing.

No phase is reported as complete in this draft.

8.2 Literature and requirements review

The review will cover grazing emissions, whole-farm modelling, remote sensing, land and herd identity, farm data governance, consent, producer burden, technology adoption, temporal data models, provenance, and extensive-system MRV. Peer-reviewed research, official industry codes, technical reports, method documents, and relevant Australian frameworks should be searched separately and appraised according to source type.

The final protocol should specify databases, search strings, date and language limits, grey-literature handling, screening, extraction, quality appraisal, and citation verification.

8.3 Participant and case sampling

Sampling should represent variation in region, herd size, enterprise structure, land tenure, production intensity, connectivity, digital maturity, record systems, supplementation practices, drought exposure, and participation in assurance or value-chain programs.

Participants may include producers, managers, advisors, veterinarians, extension specialists, remote-sensing providers, model developers, measurement specialists, data stewards, verifiers, buyers, lenders, processors, and payment administrators.

The sampling strategy should avoid recruiting only digitally advanced or program-ready enterprises. Compensation, confidentiality, withdrawal, and conflicts must be disclosed.

8.4 Participatory workflow mapping

Each producer workflow will be mapped through seasonal calendars, process diagrams, record inventories, interviews, observation where permitted, and exception walkthroughs.

The study will identify:

- ordinary decisions and events;
- who performs and records them;
- source systems and paper records;
- timing, frequency, and latency;
- duplicate entry and reconciliation;
- periods of high operational pressure;
- missing-data causes;
- exceptional seasons and events;
- records unrealistic to collect;
- opportunities for automation or reuse.

8.5 Graph construction and integration

A representative workflow will be transformed into a temporal graph covering identity, rights, land, herd, management, supplementation or intervention, weather, pasture, water, productivity, model execution, scientific review, and institutional use.

Every imported source will record:

- source owner and access authority;
- source-system identity and version;
- extraction date and method;
- raw or canonical commitment;
- mapping rules;
- transformations and alignment;
- missingness and conflicts;
- uncertainty and limitations;
- retention and permitted use.

8.6 Temporal-alignment experiments

The study will construct controlled alignment problems involving:

- bi-monthly satellite observations with continuous weather and discrete herd movements;
- monthly or seasonal herd records with event-based sales and births;
- irregular pasture assessments with daily model requirements;

- paddock geometry change during the analysis period;
- missing remote-sensing scenes;
- weather-station substitution;
- delayed producer entry;
- uncertain movement dates;
- mixed direct and proxy exposure records.

Independent reviewers will compare the declared transformation with hidden or default preprocessing. Outcomes include reproducibility, discrepancy, gap preservation, and interpretability.

8.7 Model-domain and uncertainty review

Each model execution will be assessed for input completeness, calibration domain, production-system applicability, spatial and temporal scale, parameter sensitivity, uncertainty propagation, and reviewer state.

Model outcomes should be classified as:

$$\{\texttt{in\_domain}, \texttt{conditionally\_in\_domain}, \texttt{out\_of\_domain}, \\ \texttt{requires\_extension}, \texttt{insufficient\_evidence}, \texttt{unassigned}\}.$$

The review should distinguish model quality from data quality and successful execution from scientific applicability.

8.8 Burden and usability assessment

Participants will record setup and recurring time, manual entries, duplicate requests, missing data, exception resolution, support interactions, and artifact usefulness. Data may be collected through structured diaries, passive logs where consented, task observation, system logs, interviews, and artifact tests.

Initial integration effort will be separated from recurring producer burden. Burden should be reported as a distribution and by workflow component, not only as an average.

8.9 Independent reconstruction

Independent reviewers will receive a frozen evidence bundle, schema, mapping rules, model package where permitted, and decision question. They will reconstruct:

- land and paddock state;
- herd and cohort membership;
- movement and management timeline;
- intervention or supplementation exposure;
- environmental and productivity evidence;
- model inputs, execution, and outputs;
- rights and consent chain;
- institutional artifact and limitations.

Measures include reconstruction time, clarification requests, unresolved gaps, quantitative discrepancies, classification agreement, and ability to identify unsupported inference.

8.10 Scenario testing

At least twelve scenarios should include:

1. herd movement across paddocks;
2. temporary property transfer or agistment;
3. drought or abnormal season;
4. intermittent supplementation;
5. missing satellite observations;
6. delayed producer entry;
7. model-domain mismatch;
8. contested herd or land authority;
9. changed paddock boundaries;
10. weather-source substitution;
11. conflicting rights or claim permissions;
12. artifact export that drops assumptions or limitations.

Additional scenarios should test correction without erasure, consent withdrawal, model supersession, remote-sensing version change, duplicate herd identity, and payment dispute.

8.11 Analysis plan

Quantitative analysis will report source coverage, missingness, setup and recurring burden, duplicate entry, alignment discrepancies, reconstruction time, clarification requests, model-domain classifications, agreement, and artifact utility.

Candidate methods include descriptive statistics with uncertainty intervals, mixed-effects or hierarchical models for repeated observations, survival or time-to-resolution analysis for gaps, agreement statistics for classification, and ordinal models for utility ratings. Methods must be preregistered and matched to sample size and data structure.

Qualitative analysis will examine burden causes, trust, rights concerns, workflow failures, interpretation errors, and participant recommendations. Dissent and negative cases should be retained.

9 Conformance requirements

A conforming implementation should satisfy the following minimum properties:

1. Identity nodes use explicit namespaces and authority.
2. Land geometry, herd membership, and rights are versioned over time.
3. Every derived value links to a declared transformation.
4. Remote-sensing products preserve acquisition, product, spatial support, processing, quality, and uncertainty.
5. Model executions preserve version, inputs, transformations, domain, and review state.
6. No interpolation, imputation, smoothing, or aggregation occurs without an alignment record.
7. Rights and consent identify subject, grantor, grantee, purpose, scope, period, restrictions, and status.
8. Corrections and disputes append new records rather than erase history.
9. Missing, conflicting, partial, intermittent, and out-of-domain states remain visible.
10. Scientific, method, verification, rights, payment, and institutional decisions are not inferred from technical integrity alone.
11. Producer-facing artifacts expose obligations, limitations, and correction routes.
12. Exported artifacts retain material assumptions and limitations.

10 Worked cases

10.1 Bi-monthly satellite biomass and daily emissions model

A model requires daily pasture biomass. Satellite observations are available every two months, with one missing scene due to cloud. A producer pasture assessment occurs between two scenes.

A conforming graph preserves each observation, its spatial support and uncertainty, and creates an explicit temporal model for daily estimates. The missing scene is not silently forward-filled. The output is conditional on the declared interpolation or process model.

10.2 Herd movement recorded after the event

A herd is moved between paddocks during a high-workload period. The producer records the movement two weeks later and is uncertain of the exact date.

The event has a record time different from valid time and may use a bounded date interval. The system must not assign false daily precision. Downstream pasture allocation and intervention exposure inherit the temporal uncertainty.

10.3 Landholder and herd owner differ

A producer manages cattle owned by one entity on land controlled by another. A buyer requests data and an environmental outcome claim.

The graph requires separate land authority, herd authority, producer agency, data-use consent, and claim-right assertions. Operational evidence may be complete while disclosure or claim rights remain conditional or conflicting.

10.4 Supplement delivery without consumption evidence

Supplement is delivered to a paddock, but disappearance and animal access are not observed.

Delivery is present; exposure is conditional or missing. A model may use an authorised proxy with explicit uncertainty, but the graph cannot treat delivered quantity as consumed dose automatically.

10.5 Remote-sensing version change

A provider updates its biomass algorithm. Historical source imagery is unchanged, but output values differ.

The original and revised products coexist. Each artifact identifies the product version. A later model may re-run against the new product without overwriting the earlier evidence root.

10.6 Model executes outside regional calibration

A whole-farm model runs successfully using northern Australian data, but its calibration evidence is primarily from southern systems.

Execution integrity is valid. Applicability is out of domain or requires extension. The system must not translate computational success into scientific validity.

10.7 Consent withdrawal

A producer withdraws permission for future buyer disclosure after an earlier artifact was validly issued.

The graph records withdrawal time and scope. Existing lawful artifacts are not silently erased, but future disclosures must evaluate the new rights state and contractual obligations.

10.8 Drought exception

A severe drought causes destocking, emergency feeding, movement, and atypical pasture conditions. The model's standard seasonal assumptions no longer apply.

The exception is represented as a linked event and may trigger an alternative model profile, conditional result, or exclusion. The system must not force the period into a normal-season template.

10.9 Institutional usefulness differs

A producer summary is comprehensible and useful for payment reconciliation. A verifier requires source commitments and transformation details absent from the summary. A lender needs only a bounded risk indicator.

One evidence root may support different artifacts. Acceptance by one institution does not imply sufficiency for another.

11 Planned results reporting

Planned results - not findings. No producer, workflow, burden, remote-sensing, model, reconstruction, regional, verification, or institutional results are reported.

Table 4: Paper 5 planned results template.

Measure	Result	Required reporting
Participant and system characteristics	TBD	Region, enterprise type, herd size, tenure, digital maturity, connectivity, systems, and selection limitations.
Existing record coverage	TBD	Records already generated in ordinary operations; source, frequency, completeness, accessibility, and authority.
Setup burden	TBD	Initial hours, systems connected, identity reconciliation, permissions, training, and cost.
Recurring producer burden	TBD	Time, frequency, manual steps, duplicate entry, exception effort, and seasonal variation.
Missing-data frequency	TBD	Field, cause, timing, consequence, and resolution.
Temporal-alignment gaps	TBD	Missing intervals, subject or spatial mismatch, interpolation, assumptions, and uncertainty.
Remote-sensing integration	TBD	Product coverage, quality flags, spatial fit, versioning, discrepancy, and ground relationship.
Model-domain outcomes	TBD	In-domain, conditional, out-of-domain, extension, and insufficient-evidence cases.
Reconstruction	TBD	Reviewer time, clarification requests, discrepancies, unresolved gaps, and reproducibility.
Rights and consent issues	TBD	Missing, expired, conflicting, restricted, withdrawn, and unresolved authority records.
Artifact usefulness	TBD	Producer, scientist, verifier, buyer, lender, processor, and payment perspectives.
Failure cases	TBD	Scenario, expected state, observed behaviour, unsupported inference, and remediation.

12 Discussion

12.1 Extensive systems require explicit variability, not weaker standards

A credible grazing architecture must reject both facility-style rigidity and narrative exceptionalism. Extensive systems cannot be represented faithfully by pretending that every variable is continuous and every group is fixed. They also cannot support institutional reconstruction through unstructured explanation alone.

The graph provides a middle path: stable primitives with explicit intervals, spatial support, evidence classes, transformations, uncertainty, and local extension points. Variability is represented rather than erased.

12.2 Rights-first design prevents late institutional failure

Rights and consent are often discovered late, after scientific and technical work has been completed. A buyer may learn that the producer cannot authorise disclosure of a licensed dataset. A lender may receive an artifact whose claim rights are disputed. A verifier may find that herd and land authority do not align.

Treating rights as graph primitives allows these issues to be tested early. The graph does not resolve legal questions, but it makes them visible and traceable. This can prevent the institutional equivalent of building a technically sound bridge on land whose access rights were never confirmed.

12.3 Temporal alignment is part of evidence, not preprocessing

In heterogeneous environmental data, preprocessing can determine the result. Forward-filling herd counts, interpolating biomass, assigning weather, or smoothing observations may be scientifically reasonable in one context and misleading in another.

By converting alignment into explicit evidence, the framework permits review, reproduction, comparison, and supersession. A later reviewer can see not only the values used but how asynchronous records became model inputs.

12.4 Remote sensing extends coverage but does not erase ground uncertainty

Remote sensing is valuable because it provides repeated spatial information at scales impossible to observe manually. It does not automatically measure intake, animal exposure, or methane. Satellite-derived biomass and land-condition products depend on algorithms, calibration, spatial resolution, weather, cloud, and ground relationships.

The architecture therefore treats remote sensing as a powerful evidence source with declared domain and uncertainty, not as an omniscient substitute for farm records.

12.5 Whole-farm models are versioned interpretations

Whole-farm models integrate biology, management, climate, and economics. Their strength is precisely that they interpret interacting evidence. That strength creates governance requirements: inputs, transformations, calibration domain, version, assumptions, and uncertainty must remain inspectable.

A model is like a flight simulator driven by a recorded weather file and aircraft configuration. Replaying the simulation exactly demonstrates reproducibility. It does not prove the weather file represented the farm, that the herd matched the model population, or that every management event was captured. Execution and applicability must remain separate.

12.6 Producer burden affects scalability and validity

Burden should be reported beside technical performance. Requiring frequent manual entry may increase apparent resolution while reducing actual accuracy. Duplicate requests can create inconsistent copies. Retrospective records can introduce recall bias. Systems that do not return useful artifacts to producers may not sustain participation.

The study therefore treats burden reduction as an architectural objective. Reuse of ordinary records, automated source commitments, clear exception handling, and one evidence root supporting multiple artifacts should be evaluated empirically.

12.7 Institutional utility is purpose-specific

A producer, scientist, verifier, buyer, lender, and payer ask different questions. One compact interface cannot answer all of them without either hiding detail or overwhelming users.

The thin-waist architecture permits multiple artifacts derived from one evidence root. Each artifact should retain purpose, policy version, exclusions, limitations, and reliance scope. A green badge or generic “verified” label would defeat the purpose.

12.8 Australian-first does not mean nationally universal

Australian grazing systems range from intensive improved pastures to vast northern rangelands. Climate, breeds, herd management, property scale, connectivity, supplementation, remote-sensing performance, and record practices vary materially.

The paper should therefore produce stable primitives and regional profiles rather than claim one minimum dataset is nationally sufficient. Local experts and method authorities retain responsibility for applicability.

13 Limitations

Australian grazing systems are highly diverse, and a limited participant set cannot establish national sufficiency. Participants may be more digitally capable or institutionally engaged than the broader producer population.

Workflow mapping may change participant behaviour. Self-reported time and burden can be affected by recall, observation, and perceived expectations. Passive logs may omit cognitive burden.

Remote-sensing and model products may be proprietary, making full reproducibility difficult. Product versions, licensing terms, and calibration evidence may change.

Land, herd, data, and claim rights are legally and contractually complex. A rights graph records evidence and assertions but does not determine ownership, entitlement, admissibility, or enforceability.

Temporal alignment can make assumptions visible but cannot eliminate uncertainty created by sparse observations or missing records.

Synthetic scenarios may understate real-world messiness. Permissioned cases may be incomplete or constrained by confidentiality.

A structurally valid graph does not prove that source assertions are factually correct. Cryptographic integrity, scientific applicability, accredited verification, and institutional reliance remain separate.

The supplied literature is a curated starting set, not a completed systematic review. Several references require bibliographic verification and may need correction, replacement, or removal.

14 Conclusion

Extensive grazing beef systems require evidence architecture that preserves movement, seasonality, spatial context, intermittent exposure, sparse measurement, model dependence, rights, and producer burden. Static facility-style tables can create false completeness, while unstructured narratives obstruct reconstruction.

This paper proposes a rights-first temporal graph in which land, herd, event, observation, model, transformation, consent, claim, and decision evidence remain linked but semantically distinct. Explicit temporal and spatial alignment rules prevent silent interpolation. Versioned model and remote-sensing records preserve calibration domain and uncertainty. Producer burden is evaluated directly through setup effort, recurring work, missingness, duplicate entry, and reconstruction latency.

The planned study will test whether one frozen whole-farm evidence root can support scientific, verification, buyer, lender, and payment artifacts without rewriting historical evidence or laundering conditional model outputs into observed facts. The intended contribution is a practical and inspectable evidence profile for Australian grazing beef, not a universal method, legal determination, or certification.

15 Declarations

Participation and ethics. TODO: Obtain an ethics determination for interviews, workflow observation, burden diaries, usability testing, and identifiable operational cases.

Consent and compensation. TODO: Define informed consent, withdrawal, participant compensation, feedback, and use of producer quotations or case details.

Data governance. TODO: Approve controls for property, geospatial, herd, personal, commercial, remote-sensing, contractual, and model data, including de-identification, access, retention, and disclosure.

Funding and support. TODO: Disclose AgEvidence and other funding, in-kind support, technology contributions, data access, and sponsor roles.

Competing interests. TODO: Collect relationships involving models, remote-sensing products, digital platforms, methods, verification, environmental markets, buyers, finance, and intellectual property.

Data availability. TODO: Publish de-identified workflow maps, schemas, field dictionary, synthetic scenarios, alignment fixtures, burden instruments, and results where rights permit.

Code availability. TODO: Archive graph schemas, validators, transformation rules, model wrappers where licensed, test corpus, and environment manifests.

Author contributions. TODO: Complete a live CRediT record and obtain written author consent, affiliation confirmation, conflict disclosure, and final approval.

Generative tools. TODO: Apply target-journal disclosure requirements and retain human responsibility for every claim, citation, analysis, and final passage.

16 Executable TODO register

Table 5: Executable Paper 5 TODO register.

ID	Task	Accountable owner	Status
P5.01	Confirm lead, co-authors, affiliations, CRediT roles, and conflicts.	Persona F	Open
P5.02	Select target journal and freeze article type and reporting requirements.	Lead author	Open
P5.03	Verify all supplied references and complete structured literature review.	Scientific authors	Open
P5.04	Approve ethics, consent, compensation, withdrawal, and participant communication.	Data Steward; Persona F	Open
P5.05	Approve property, herd, geospatial, commercial, remote-sensing, and model data route.	Data Steward	Open
P5.06	Define participant sampling and recruit diverse Australian producer and scientific participants.	Persona F; Persona A	Open
P5.07	Complete participatory workflow mapping and seasonal exception analysis.	Persona F	Open
P5.08	Inventory source systems, ordinary records, access constraints, and duplicate requests.	Implementation team	Open
P5.09	Freeze whole-farm ontology, graph profile, evidence classes, and rights schema.	Persona G; author approval	Open
P5.10	Implement temporal and spatial alignment protocol and negative conformance fixtures.	Persona G	Open
P5.11	Define remote-sensing source profile, quality fields, versioning, and spatial tests.	Remote-sensing collaborator	Open
P5.12	Define model-domain, uncertainty, sensitivity, and reproducibility review.	Persona B; Persona D	Open
P5.13	Build representative evidence graph and permissioned or synthetic bundle.	Technical and data team	Open
P5.14	Measure setup burden, recurring burden, missingness, duplication, and reconstruction latency.	Persona F	Open
P5.15	Run artifact comprehension and utility tests across producer and institutional roles.	Methods team	Open
P5.16	Construct and run at least twelve failure and exception scenarios.	Persona G; scientific authors	Open
P5.17	Run independent reconstruction and discrepancy analysis.	Independent reviewers	Open
P5.18	Obtain external Australian producer, scientific, governance, verifier, and legal review.	Persona A; Data Steward	Open
P5.19	Resolve changes to Papers 1–3 through versioned portfolio change control.	Scientific Chairs	Open
P5.20	Complete disclosures, data/code statements, limitations, author approval, and submission archive.	Persona F	Open

Bibliography

[1] Bonner, M., and Thompson, L. (2025). Temporal graph architectures for heterogeneous agricultural recordkeeping. *Journal of Agricultural Informatics*, 16(1), 45–62. *Supplied citation; verify before submission.*

[2] Browne, N. A., Eckard, R. J., Behrendt, R., and Kingwell, R. S. (2011). A comparative analysis of greenhouse gas emissions from wool, lamb and beef production in southern Australia. *Agricultural Systems*, 104(1), 25–31.

[3] Charmley, E., Ban-Tokuda, T., and Kennedy, P. M. (2016). Methane production by northern Australian beef cattle fed native or improved pastures or supplemented with molasses or protein meal. *Animal Production Science*, 56(3), 212–219.

[4] Cibo Labs and Meat & Livestock Australia (2024). *Accelerating the adoption of satellite-assisted forage budgeting across northern beef businesses*. Project P.PSH.1286, MLA Final Report. *Institutional source; verify report metadata and access.*

[5] Donald, G. E., Edirisinghe, A., Henry, D. A., and colleagues (2010). Vegetation biomass estimation in pastoral systems using remote sensing. *Remote Sensing of Environment*, 114(7), 1530–1542. *Supplied citation; verify full author list and metadata.*

[6] Eastwood, C., Klerkx, L., Ayre, M., and Dela Rue, B. (2019). Managing the agtech transition: Producer burdens, usability standards, and decision-support architecture in pastoral systems. *Land Use Policy*, 81, 787–796. *Supplied citation; verify title and metadata before submission.*

[7] Eckard, R. J., Grainger, C., and de Klein, C. A. M. (2010). Options for the abatement of methane and nitrous oxide from ruminant production: A review. *Animal Feed Science and Technology*, 166–167, 644–650.

[8] Harrison, M. T., Cullen, B. R., Tomkins, N. W., and colleagues (2014). Potential options for managing greenhouse gas emissions from extensive grazing systems in Australia. *Crop and Pasture Science*, 65(7), 657–669.

[9] Jakku, E., Zhang, A., and Llewellyn, R. (2020). Understanding the barriers and drivers for producer adoption of digital technology in Australian broadacre and grazing systems. *Agricultural Systems*, 182, 102832. *Supplied citation; verify title and metadata.*

[10] Jonker, A., Hickey, S. M., McEwan, J. C., and colleagues (2020). Measurement of enteric methane emissions from grazing ruminants: A review of methodology, limitations, and future directions. *Ruminant Physiology*, 14(2), 112–128. *Supplied citation; bibliographic record requires verification.*

[11] Moore, A. D., Bell, L. W., and Thomas, D. T. (2014). Feed-base options for reducing greenhouse gas emissions from beef cattle production systems in southern Australia. *Agricultural Systems*, 129, 121–134.

[12] Muleke, A. I., Clark, M. L., and Handcock, R. N. (2022). Spatio-temporal alignment of multi-spectral satellite imagery and ground observations for pastoral biomass budgeting. *International Journal of Applied Earth Observation and Geoinformation*, 108, 102715. *Supplied citation; verify title and metadata.*

[13] National Farmers' Federation (2023). *Australian Farm Data Code*, Edition 2. NFF policy framework. *Official industry source; verify edition date and URL before submission.*

[14] Sanderson, J., Wiseman, L., and Poncini, A. (2018). *What's behind the data? Agribusiness, digital technology and data ownership in Australia.* Rural Industries Research and Development Corporation, Publication No. 18/012. *Verify publisher and report metadata.*

[15] Vargas, R., and Lopez, M. (2023). Explicit error propagation and domain limits in agricultural bio-economic models. *Environmental Modelling & Software*, 160, 105612. *Supplied citation; bibliographic record requires verification.*

[16] Wiseman, L., Sanderson, J., Zhang, A., and Jakku, E. (2019). Farmers and their data: An examination of farmers' trust in digital agriculture and data sharing frameworks. *Australian Journal of Agricultural and Resource Economics*, 63(3), 586–603. *Supplied citation; verify title, authors, and pagination.*

Paper 6: Manure, Waste, and Farm Energy

Working-draft note. This manuscript expands Paper 6 from the *AgEvidence Australia Scientific Publication Series: Seven Manuscript First Drafts*. It is an engineering-architecture and planned-evaluation manuscript. It does not report a completed farm-energy case, measured abatement, verified energy output, accredited verification result, recognised renewable-energy attribute, Scope 3 reduction, government acceptance, or legal allocation of co-product rights.

Non-negotiable guardrail. A mass or energy balance is an engineering reconstruction, not by itself proof of environmental additionality, legal entitlement, government approval, registry issuance, or institutional reliance. Verification readiness is not accredited verification. A physical co-product does not automatically carry every associated environmental attribute.

Abstract

Livestock decarbonisation extends beyond enteric methane to manure storage, waste treatment, anaerobic digestion, biogas upgrading, combined heat and power, renewable electricity, recovered nutrients, bedding, digestate, thermal energy, and other farm-scale outputs. These interventions combine biological conversion, engineering operation, environmental accounting, and contractual attribute allocation. Evidence is commonly distributed across laboratory assays, weighbridges, flow meters, gas analysers, engine controllers, invoices, maintenance systems, grid meters, environmental reports, and policy calculators. Retrospective aggregation into one project spreadsheet can obscure whether a result changed because the physical system changed, a sensor drifted, a boundary moved, a default factor was updated, or an environmental attribute was reallocated.

This paper extends the Paper 1 thin-waist evidence architecture to manure, waste, and farm-energy systems while preserving the Paper 2 separation of cryptographic integrity, method compatibility, scientific or engineering judgment, accredited verification, and institutional reliance. It reuses the Paper 3 decision-linked evidence-sufficiency vocabulary and the Paper 4 separation of physical outcomes, contributions, rights, allocations, payments, and reliance.

The proposed architecture represents material inputs, storage, pre-treatment, digester configuration, operational state, laboratory and sensor provenance, gas production and composition, leakage, flare and engine operation, parasitic loads, electricity and heat output, digestate disposition, avoided counterfactuals, mass and energy balances, life-cycle calculations, co-product allocation, environmental attributes, and institution-specific reliance as an append-only directed evidence graph. Mass and energy residuals are preserved rather than silently forced to zero:

$$\varepsilon_M = \sum M_{\text{in}} - \sum M_{\text{out}} - \Delta M_{\text{stock}}, \qquad \varepsilon_E = \sum E_{\text{in}} - \sum E_{\text{out}} - \Delta E_{\text{stock}}.$$

The study evaluates the architecture using a synthetic mirror, permissioned pilot subset, or approved post-pilot case. Independent reviewers receive a bounded offline bundle and attempt to reconstruct material flows, gas and energy calculations, leakage treatment, life-cycle boundaries, co-product allocation, and attribute claims under versioned Australian and synthetic comparison adapters. Planned outcomes include parent and source resolution, sensor-quality states, residual magnitude, quantitative discrepancy, exception visibility, adapter differences, double-attribution detection, reconstruction time, and reviewer disagreement. The framework is not a new emissions-reduction methodology and does not establish product efficacy, asset ownership, renewable-energy entitlement, Scope 3 acceptance, or government approval.

Keywords: manure management; anaerobic digestion; farm energy; evidence architecture; mass balance; energy balance; circular economy; environmental attributes; Australian livestock; verification readiness.

1 Manuscript status, scope, and contribution boundary

Paper 6 is the engineering-domain extension of the portfolio. It depends on Paper 1 for the stable receipt contract, append-only graph, source commitments, immutable parent links, versioned adapters, and bounded reliance artifacts. It depends on Paper 2 for the validity topology

$$\mathcal{V}(E, A) = \langle \Phi(E, S), \Psi(E, M_v), \Omega(A, P_v, I) \rangle ,$$

where integrity, compatibility, and reliance remain distinct. It reuses Paper 3's field-level evidence states:

$$s_f \in \{\texttt{present}, \texttt{missing}, \texttt{conflicting}, \\ \texttt{out\_of\_domain}, \texttt{conditionally\_sufficient}, \texttt{not\_applicable}\},$$

and Paper 4's separation of physical outcome, contribution, rights, allocation, payment, and institution-specific reliance.

The manuscript asks whether the same architectural primitives remain useful when the dominant evidence pathway shifts from animal intervention and feed delivery to material flows, engineered processes, continuous sensors, energy conversion, co-products, and multiple environmental attributes.

The contribution is not a manure or energy methodology. It is a proposed evidence architecture, engineering vocabulary, conformance profile, and planned reconstruction study.

2 Provisional scientific author personas

Persona	Provisional role	Contribution domain / deliverable
Persona E	Lead/corresponding-author persona	Engineering model, process boundaries, mass/energy balances, anaerobic digestion, sensor/laboratory provenance and final manuscript.
Persona A	Scientific coauthor persona	Australian emissions-accounting, manure-management, farm-system and local-authority boundaries.
Persona B	Scientific coauthor persona	Biological conversion, methane yield, model-domain and biological uncertainty.
Persona C	Scientific coauthor persona	Environmental attributes, co-claims, contracts, buyer/verifier reliance and double-attribution scenarios.
Persona D	Scientific coauthor persona	Sensor/laboratory quality, residuals, uncertainty, gaps and evidence sufficiency.
Persona F	Scientific coauthor persona	Operator/producer workflow, maintenance, downtime, burden and farm usability.
Persona G	Technical scientific contributor persona	Engineering receipt profile, balance evaluator, graph implementation, adapters and offline conformance bundle.
Persona I	Independent reviewer persona	Engineering reconstruction, LCA, verification and claims review without privileged access.

Authorship control. Persona IDs are recruitment placeholders only. A real person is not an author until consent, qualifying contribution, affiliation, conflict disclosure, CRediT allocation, and final-manuscript approval are complete.

3 Introduction

3.1 The evidence problem beyond enteric methane

Livestock decarbonisation is not limited to enteric methane. It encompasses manure storage and treatment, anaerobic digestion, biogas utilisation, renewable electricity, recovered nutrients, bedding, digestate, thermal energy, and other outputs that may belong to one tightly integrated physical system (Holm-Nielsen et al. 2009; Weiland 2010). The evidence challenge is both biological (Owen and Silver 2015; Chadwick et al. 2011) and engineering-based (Krauslyte and Venslauskas 2020).

A farm-scale project may rely on source-material characteristics, storage duration, volatile solids, volumetric flow, temperature, pressure, gas composition, methane destruction efficiency, engine performance, parasitic load, downtime, fugitive leakage, nutrient content, and co-product disposition (Amon et al. 2006; Daniel-Gromke et al. 2018). Calculations may depend on default factors, counterfactual baselines, displacement assumptions, grid factors, life-cycle boundaries, and co-product allocation rules (International Organization for Standardization 2006; Basset-Mens and Werf 2005). Contracts and market instruments may divide energy, environmental attributes, nutrient products, and financial payments among different actors (Gillison and Brander 2020; Lazarus et al. 2012).

In Australia, project-level evidence may also be compared with national inventory methods, regional factors, production-system data, and local policy requirements (Department of Industry, Science, Energy and Resources 2023; Eckard et al. 2010; Henry et al. 2012; Wiedemann et al. 2016). Those national or institutional rules belong in versioned adapters rather than in the generic evidence kernel.

3.2 Why monolithic calculators are brittle

A tightly coupled calculator can combine raw measurements, assumptions, emission factors, allocation rules, and report labels into one output. This may be efficient for a first submission, but it obscures causality when the result changes. A different number may reflect a physical process change, a corrected sensor, a new boundary, a new factor, a model update, or a policy reinterpretation.

The thin-waist architecture separates these layers. Raw and processed observations remain immutable evidence receipts. Engineering calculations and model executions identify versions and inputs. Method and policy adapters interpret the graph. Verification and institutional reliance remain explicit decisions. A later adapter can re-evaluate the same historical evidence root without rewriting the original measurements.

3.3 Residuals are evidence

Mass and energy balances are often presented as if closure were exact. In practice, residuals arise from sensor error, unmeasured streams, sampling variability, gas leakage, stock changes, water addition, evaporation, maintenance events, and boundary choices.

The architecture therefore preserves residuals explicitly. It does not force them to zero by undocumented reconciliation. A residual may be acceptable, conditionally sufficient, unresolved, or evidence of a missing stream depending on the declared engineering profile.

3.4 Physical flows and environmental attributes are different objects

Digestate, heat, power, biomethane, recovered nutrients, renewable-energy certificates, carbon reductions, and Scope 3 claims are related but not interchangeable. The same physical process can generate multiple outputs and possible attributes. The graph must distinguish physical custody from contractual rights and institutional acceptance.

Paper 4's aggregate-cap principle generalises here: quantitative claims within one declared attribute class must not silently exceed the recognised quantity available under the governing policy. Cross-attribute rules must be explicit because one physical unit may interact with more than one accounting system.

3.5 Research questions

1. Can the Paper 1 receipt contract represent manure, waste, digestion, energy, co-product, and attribute evidence without domain-specific cryptographic branches?
2. Which physical, laboratory, sensor, model, authority, and rights fields are minimally necessary for independent engineering reconstruction?
3. Can mass and energy residuals be represented transparently without being silently normalised away?
4. Can one frozen evidence root support multiple baselines, emission-factor versions, LCA boundaries, and policy adapters?
5. Can environmental attributes and co-product allocations be represented without conflating physical output, ownership, claim language, and institutional reliance?
6. Can independent reviewers reproduce material flows, gas and energy output, leakage, allocation, and limitations without hosted-platform access?
7. Which engineering exceptions produce `missing`, `conflicting`, `out_of_domain`, or `conditionally_sufficient` states?

3.6 Novelty relative to conventional project dossiers

Table 2: Novelty analysis for the proposed engineering evidence architecture.

Dimension	Conventional approach	Paper 6 proposed architecture
Physical evidence	Sensor values and factors are merged in a calculator output.	Raw, processed, calculated, modelled, reviewed, verified, and relied-upon propositions remain distinct.
Mass and energy closure	Residuals are hidden, rounded, or allocated implicitly.	Residuals ε_M and ε_E are explicit evidence with thresholds, causes, and review states.
Policy change	Historical spreadsheets are overwritten or reissued.	Immutable evidence is re-evaluated under versioned method and policy adapters.
Sensor quality	Calibration and drift are treated as maintenance metadata.	Device, calibration, drift, downtime, substitution, and quality states are graph nodes.
Co-products	Physical flows and environmental attributes are bundled.	Physical output, contribution, rights, allocation, payment, retirement, and reliance are separate.
Double attribution	Controlled through manual reconciliation.	Versioned attribute ledgers and conformance rules identify incompatible or over-allocated claims.
Auditability	Review depends on hosted access and project-team explanation.	Bounded offline bundles support independent reconstruction and unresolved-gap reporting.

4 Scientific and methodological foundations

4.1 Manure emissions and biochemical mechanics

Manure emissions vary by animal system, storage, temperature, retention time, solids content, aeration, crust formation, treatment, and land application. Methane, nitrous oxide, and ammonia pathways can shift when a management practice changes (Amon et al. 2006; Chadwick et al. 2011; Owen and Silver 2015). Measurement methods for buildings and storages also have different temporal and spatial limitations (Hassouna et al. 2016).

The architecture must therefore preserve process state and boundary context rather than attach one emissions factor to "manure" as a generic object.

4.2 Anaerobic digestion and biogas engineering

Anaerobic digestion transforms organic substrates through biological and engineered stages. Biomethane potential testing, substrate composition, loading, temperature, retention time, inhibition, mixing, gas handling, and downstream utilisation affect performance (Angelidaki et al. 2009; Holm-Nielsen et al. 2009; Weiland 2010). A laboratory BMP value is not a plant-output guarantee; it is evidence under a defined assay.

4.3 Mass and energy balance

Mass-balance models and biogas-plant assessments require explicit boundaries and stock changes (Bachmaier and Gronauer 2007; Krauslyte and Venslauskas 2020; Liao et al. 2006). Let k index material streams over interval $[t_0, t_1]$. The mass residual is

$$\varepsilon_M = \sum_{k \in \mathcal{I}} M_k - \sum_{k \in \mathcal{O}} M_k - \Delta M_{\text{stock}}.$$

For energy streams,

$$\varepsilon_E = \sum_{j \in \mathcal{I}_E} E_j - \sum_{j \in \mathcal{O}_E} E_j - \Delta E_{\text{stock}}.$$

Closure ratios may be reported as

$$r_M = \frac{|\varepsilon_M|}{\max(\sum M_{\text{in}}, \epsilon_M)}, \qquad r_E = \frac{|\varepsilon_E|}{\max(\sum E_{\text{in}}, \epsilon_E)},$$

with declared positive constants ϵ_M, ϵ_E. Thresholds are adapter rules, not universal constants.

4.4 Sensor provenance, drift, and leakage

Continuous monitoring can improve temporal coverage but introduces calibration, drift, dropout, placement, response-time, and maintenance risks. Leakage and fugitive methane may materially alter net outcomes (Daniel-Gromke et al. 2018; Bacheson and Fischer 2019). IoT-style provenance can preserve device identity, firmware, calibration, time synchronisation, transformation, and custody (Kist and Zhang 2021), but cryptographic integrity does not prove calibration or representativeness.

4.5 Life-cycle assessment and co-product allocation

LCA requires declared goals, functional units, system boundaries, inventory data, allocation rules, and impact methods (International Organization for Standardization 2006; Guinée 2002). Scenario-based farm assessments show how conclusions can change with boundary and allocation choices (Basset-Mens and Werf 2005; Eide 2002; Wiedemann et al. 2016).

The architecture stores physical inventory evidence independently of the LCA policy profile. Economic, mass, energy, substitution, and system-expansion allocations are separate adapter versions.

4.6 Environmental attributes and double counting

Scope 3, renewable-energy, carbon-market, and other environmental attributes may overlap in complex ways (World Resources Institute and World Business Council for Sustainable Development 2011; Gillison and Brander 2020; Lazarus et al. 2012). The architecture cannot determine entitlement by arithmetic alone. It can preserve rights evidence, policy versions, allocation decisions, retirement events, and conflicting claims so that double attribution is detectable rather than hidden.

Working-draft note. The supplied literature compendium has been integrated. Numerous records are marked "verification pending" because their bibliographic details or existence were not independently confirmed. All references and interpretive claims must be checked against publisher or official sources before submission.

5 Engineering evidence model

5.1 Material inputs and source identity

Minimum material-input evidence includes:

- source farm, property, enterprise, and authority;
- animal class and manure-production context;
- manure, slurry, bedding, crop residue, food waste, or co-substrate identity;
- quantity, unit, measurement method, timestamp, and uncertainty;
- dry matter, volatile solids, carbon, nitrogen, and other relevant composition;
- contamination and exclusion criteria;
- transport, custody, storage, and pre-treatment;
- batch, load, tank, pit, or stock identifier; and
- source-document commitment and access conditions.

A declared tonne without wet/dry basis, timestamp, and measurement method is not sufficient for material reconstruction.

5.2 Process configuration and operation

Process evidence includes:

- system boundary and process-flow diagram;
- storage, separation, mixing, pumping, digestion, upgrading, flaring, combustion, and export equipment;
- capacity, geometry, insulation, and design parameters;
- operating temperature, pressure, pH, loading, retention time, mixing, and recirculation;
- start-up, shutdown, maintenance, upset, and bypass events;
- alarm and control states;
- operator identity and authority;
- configuration or software change; and
- declared unavailable or estimated periods.

A configuration change is appended as a new receipt and linked to the affected interval.

5.3 Sensor and laboratory provenance

Every material sensor or laboratory result should preserve:

- device or laboratory identity;
- serial number, firmware, method, and unit;
- calibration, reference standard, and valid period;
- sampling location, frequency, aggregation, and time synchronisation;
- raw and processed values;
- missing values, filtering, interpolation, substitution, and outlier rules;
- maintenance, drift, failure, and replacement;
- uncertainty and detection limits;
- operator or analyst; and
- source commitment and transformation lineage.

A sensor may be cryptographically valid but scientifically `conditionally_sufficient` because calibration expired or placement was unrepresentative.

5.4 Mass balance

For each declared interval and system boundary, the graph assembles material-stream receipts and stock-change evidence. It reports the residual and identifies whether it is:

- within the adapter tolerance;
- explained by declared evaporation, gas, sampling, or stock uncertainty;
- associated with a missing or conflicting stream;
- outside the current engineering profile; or
- unresolved and requiring additional evidence.

Residual adjustment must be explicit. If reconciliation allocates residual proportionally, the rule and pre-adjustment values remain preserved.

5.5 Energy balance

Energy inputs may include electricity, fuel, heat, and embodied process energy where required. Outputs may include electricity, biomethane, useful heat, exported fuel, and energy retained in co-products. Parasitic loads, flaring, downtime, and conversion losses must remain visible.

Net exported electricity, for example, is

$$E_{\text{net}} = E_{\text{gross}} - E_{\text{parasitic}} - E_{\text{auxiliary}},$$

under declared meter and boundary rules. It is not equivalent automatically to avoided grid emissions.

5.6 Methane production, destruction, and leakage

For interval t, measured or modelled methane flows may include:

$$CH_{4,\text{produced}}, \quad CH_{4,\text{combusted}}, \quad CH_{4,\text{upgraded}}, \quad CH_{4,\text{flared}}, \quad CH_{4,\text{leaked}}, \quad CH_{4,\text{stored}}.$$

A simplified methane balance is

$$\varepsilon_{CH_4} = CH_{4,\text{produced}} - CH_{4,\text{combusted}} - CH_{4,\text{upgraded}} - CH_{4,\text{flared}} - CH_{4,\text{leaked}} - \Delta CH_{4,\text{stored}}.$$

Each term must identify whether it is observed, calculated, modelled, assumed, or defaulted.

5.7 Co-products and environmental attributes

Physical co-products include digestate, separated solids, recovered nutrients, bedding, heat, electricity, and fuel. Attribute receipts may include:

- avoided-manure methane outcome;
- renewable-electricity attribute;
- renewable-gas attribute;
- displaced fossil-energy claim;
- Scope 3 intervention claim;
- nutrient-recovery or circularity claim;
- carbon credit or offset where applicable; and
- institution-specific sustainability reliance.

The graph separates physical custody from legal or contractual rights. It records allocation authority, policy version, claim boundary, unit, period, retirement, supersession, and unresolved overlap.

5.8 Attribute-cap and compatibility rules

For attribute class c, recognised quantity Q_c, and allocations $a_{i,c}$,

$$a_{i,c} \geq 0, \qquad \sum_i a_{i,c} \leq Q_c.$$

This prevents over-allocation within one declared class. Cross-class compatibility is governed by a policy matrix K, where

$$K_{cd} \in \{\texttt{compatible}, \texttt{conditional}, \texttt{incompatible}, \texttt{unassigned}\}.$$

A conforming system does not assume that attributes in different classes are additive or mutually exclusive without a versioned rule.

5.9 Engineering reliance artifacts

A bounded artifact identifies:

- evidence root and manifest;
- system boundary and interval;
- method, LCA, energy, and attribute adapter versions;
- included and excluded material streams;
- sensor and laboratory quality states;
- mass and energy residuals;
- leakage, flare, downtime, and exception treatment;
- calculations, models, assumptions, and uncertainty;
- scientific and engineering reviews;
- verification scope and status;
- attributes, rights, allocations, retirements, and conflicts; and
- limitations and non-assertions.

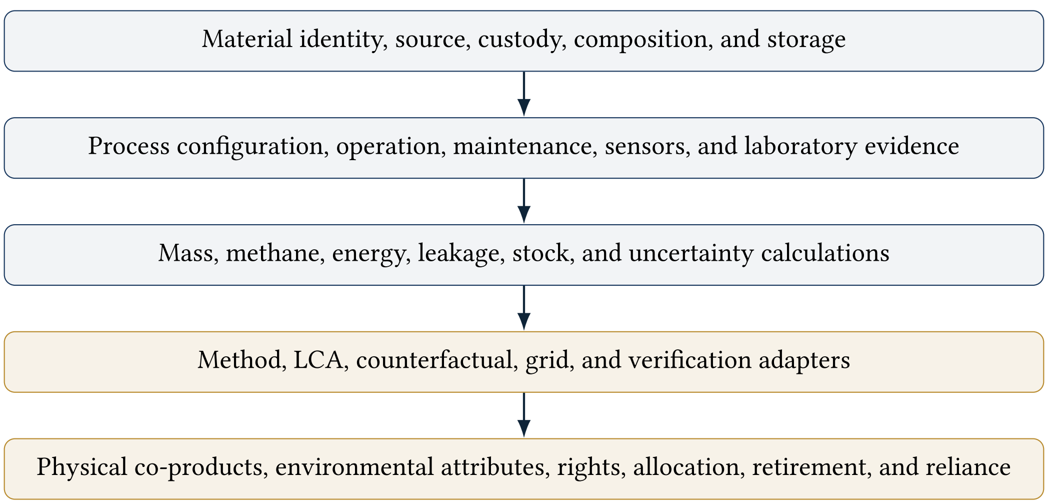

Figure 1: Paper 6 engineering evidence pathway aligned with the receipt, validity, sufficiency, and claim-separation rules established in Papers 1–4.

6 Methods

6.1 Study design and data route

The study is an applied design-science and reconstruction evaluation using one bounded manure, waste, digestion, or farm-energy case. The approved data route will be:

1. a synthetic mirror reproducing engineering complexity;
2. a permissioned and de-identified pilot subset; or
3. an approved post-pilot dataset after rights, safety, confidentiality, and release review.

Synthetic data and synthetic adapters must be labelled in every machine-readable and human-readable output.

6.2 Process mapping and graph assembly

The team will map:

1. physical system boundary and flow diagram;
2. actors, authority, custody, rights, and contracts;
3. material sources, batches, quantities, composition, and stock;
4. equipment configuration and operational intervals;
5. sensor and laboratory systems;
6. maintenance, calibration, downtime, and exceptions;
7. mass, methane, and energy calculations;
8. baseline and counterfactual assumptions;
9. LCA, allocation, and environmental-attribute profiles;
10. verification-readiness and institution-specific artifacts; and
11. payment, retirement, correction, dispute, and supersession events.

6.3 Source and field-state assessment

Each required field receives a Paper 3 state. The study will not treat document existence as sufficient. A value can be `present` but still `out_of_domain` for a decision if its sampling location, interval, method, or boundary is inappropriate.

6.4 Balance and residual evaluation

Reviewers will reconstruct pre-adjustment and post-adjustment balances. The report will include:

- raw stream totals;
- stock-change evidence;
- residual and closure ratio;
- adapter threshold;
- reconciliation method;
- uncertainty;
- identified missing or conflicting streams; and
- reviewer judgment and limitations.

No residual will be silently forced to zero.

6.5 Adapter scenarios

The same frozen evidence root will be evaluated under:

1. an Australia-oriented manure or farm-energy adapter;
2. at least one synthetic comparison adapter;
3. at least two LCA or co-product-allocation profiles where scientifically appropriate; and
4. at least two environmental-attribute or institutional profiles.

Synthetic comparison adapters test portability only and do not represent foreign compliance.

6.6 Independent reconstruction

Independent reviewers receive the bounded bundle, schemas, environment, adapters, and task protocol, but no privileged hosted access. They attempt to reproduce:

- source identity and custody;
- process configuration by interval;
- sensor and laboratory lineage;
- material, methane, and energy balances;
- residuals and reconciliation;
- leakage and downtime treatment;
- net energy and parasitic loads;
- LCA boundary and allocation;
- recognised attributes and caps;
- method compatibility and verification readiness; and
- unresolved gaps and prohibited inferences.

6.7 Exception testing

The conformance suite includes:

- expired calibration;
- sensor drift and replacement;
- missing weighbridge ticket;
- unrecorded co-substrate;
- stock change omitted;
- gas analyser dropout;
- flare or engine downtime;
- fugitive leakage detected retrospectively;
- negative or implausible flow;
- unit or wet/dry-basis mismatch;
- changed grid factor;
- changed counterfactual baseline;
- changed allocation method;
- co-product reclassification;
- overlapping renewable-energy and carbon claims;
- expired or revoked rights;
- corrected source record; and
- method-version supersession.

6.8 Evaluation metrics

Planned metrics include:

- parent and source resolution;
- sensor and laboratory evidence states;
- mass, methane, and energy residual magnitude;
- quantitative discrepancy between originator and reviewer;
- number of hidden, visible, and disputed exceptions;
- adapter determinism for machine-decidable rules;
- changed outcomes across method or policy versions;
- unchanged historical receipt digests;
- attribute-cap pass, fail, and indeterminate states;
- cross-attribute compatibility conflicts;
- reconstruction time and clarification requests;
- reviewer disagreement; and
- producer and operator capture burden.

For quantity x reconstructed as $\widehat{x}$,

$$\delta_x = \frac{|\widehat{x} - x|}{\max(|x|, \epsilon_x)}.$$

Low discrepancy does not prove scientific validity if both calculations share an incorrect boundary or assumption.

7 Worked scenarios

7.1 Expired gas-sensor calibration

The signature and file digest verify, but calibration expired during part of the interval. Integrity may be `valid`; the field is `conditionally_sufficient` or `out_of_domain`; method compatibility may be conditional or insufficient. The system must not display "verified gas production."

7.2 Unrecorded co-substrate

A later invoice reveals food waste entered the digester but was absent from the original material balance. A new source and correction branch is appended. The historical graph remains visible; revised balances and determinations supersede but do not erase prior results.

7.3 Residual outside tolerance

Mass closure exceeds the adapter threshold. The evaluator records the residual, affected streams, and unresolved cause. It does not allocate the difference silently to digestate or gas.

7.4 Fugitive leakage discovered retrospectively

A leak survey identifies emissions during a prior interval. A leakage receipt is appended and linked to the affected operating state. Net outcome and reliance artifacts are re-evaluated under the same historical root plus the new branch.

7.5 Grid-factor update

The physical electricity export receipts remain unchanged. A new policy adapter uses a revised grid factor and produces a new avoided-emissions determination. The change is policy interpretation, not altered physical evidence.

7.6 Co-product allocation change

The same inventory is evaluated under mass, energy, economic, and system-expansion profiles. Results differ, but every artifact identifies the allocation method and version.

7.7 Double attribution across attribute classes

Renewable electricity is associated with a certificate and also used in a Scope 3 reduction claim. The cross-attribute matrix returns `conditional` or `incompatible` under the selected policy. The system records the conflict and does not infer entitlement.

7.8 Positive engineering review, rights unresolved

Mass and energy balances are reconstructable and within tolerance, but contracts do not establish who owns a renewable-gas or emissions attribute. Engineering review may be favourable while reliance remains `needs_more_evidence` or `pending`.

8 Planned results reporting

Working-draft note. No engineering evaluation or independent reconstruction has been completed. This table is a reporting template, not findings.

Table 3: Paper 6 planned results template.

Analysis	Result	Required interpretation
Graph assembly	TBD	Report nodes, edges, source systems, synthetic content, missing parents, and inaccessible sources.
Material evidence	TBD	Report source, composition, quantity, wet/dry basis, custody, stock, and gap states.
Sensor and laboratory provenance	TBD	Report calibration, drift, missing intervals, substitutions, transformations, and uncertainty.
Mass balance	TBD	Report pre-adjustment residual, threshold, explanation, reconciliation, and unresolved streams.
Methane balance	TBD	Report production, use, flare, leakage, storage, residual, and observed/modelled status.
Energy balance	TBD	Report gross output, parasitic load, auxiliary inputs, useful heat, exports, losses, and residual.
Independent reconstruction	TBD	Report environment, time, clarification requests, discrepancies, and unresolved judgment.
Adapter comparison	TBD	Report changed interpretations and unchanged historical evidence.
LCA allocation	TBD	Report functional unit, boundary, allocation profile, sensitivity, and limitations.
Environmental attributes	TBD	Report rights, caps, compatibility, retirement, conflicts, and institution-specific outcomes.
Verification readiness	TBD	Report criteria, included and excluded evidence, source-access limitations, and pending decisions.

Analysis	Result	Required interpretation
Operational burden	TBD	Report manual capture, duplicate entry, maintenance evidence, and workflow redesign options.

9 Discussion

9.1 The thin waist survives the engineering shift

The Paper 1 receipt kernel remains useful because issuer, subject, authority, source commitment, parent relationship, time, limitation, and signature are common across biological and engineering evidence. Domain specificity belongs in stable agricultural receipt families and versioned engineering profiles, not in national cryptographic forks.

9.2 Integrity is not engineering adequacy

Paper 2's prohibited inference rules are essential. A signed sensor file may be intact but poorly calibrated. A reproducible model may use an inappropriate boundary. A balanced spreadsheet may hide a compensating error. Integrity, compatibility, review, verification, and reliance must remain separate.

9.3 Minimum evidence is decision-linked

Paper 3's framework prevents Paper 6 from turning into an indiscriminate data checklist. A field is required because it supports a defined decision. A detailed laboratory assay may be necessary for methane-yield modelling but unnecessary for a bounded maintenance event. Proportionality remains a scientific and operational requirement.

9.4 Physical outcome, allocation, and reliance remain separate

Paper 4's separation of Q, contribution, rights, allocation, payment, and reliance generalises to energy and co-products. A physical megawatt-hour, a renewable-energy certificate, a displaced-emissions estimate, and a Scope 3 claim are different objects. The graph should expose their relationships without collapsing them.

9.5 Residuals improve auditability

A non-zero residual is not automatically failure. It is information about the relationship among measurements, stocks, boundaries, and uncertainty. The architecture improves auditability by making residual treatment inspectable and versioned.

9.6 One evidence root can support policy evolution

National factors, grid factors, leakage assumptions, allocation rules, and claim policies change. Re-evaluating one evidence root under new adapters makes the cause of change explicit. It also allows historical artifacts to remain reproducible.

9.7 Verification readiness is empirical

The architecture may reduce verification effort by exposing lineage, exceptions, and versions. It may also shift burden upstream to operators and data stewards. The study must measure both effects rather than assume improvement.

10 Limitations

This draft has no completed case, approved dataset, measured abatement result, accredited determination, recognised environmental attribute, or government acceptance.

The supplied bibliography contains references whose metadata or existence has not been independently verified. Those records must not be treated as submission-ready until checked.

A synthetic case can test architecture behaviour but not establish operational, legal, or policy adequacy.

Mass and energy closure can be achieved with compensating errors. Residual thresholds do not replace calibration, source review, or engineering judgment.

Continuous sensors can create an appearance of precision while remaining vulnerable to placement, drift, dropout, maintenance, and time-synchronisation errors.

LCA results depend materially on goal, scope, functional unit, boundary, counterfactual, and allocation choices. The architecture preserves those choices but does not make one universally correct.

Attribute-cap rules do not prevent all double claiming. Claims may occur outside the graph, use different boundaries, or rely on incompatible policy definitions.

Rights and environmental attributes are legally and institutionally contextual. A machine-readable allocation is not a legal opinion.

Producer and operator burden may be substantial. Evidence capture must integrate with existing systems rather than assume perfect digital infrastructure.

Publication does not create an Australian Standard, approve a method, certify an emissions reduction, confer renewable-energy entitlement, determine claim ownership, or replace an accredited validation or verification body.

11 Conclusion

Paper 6 extends the portfolio's evidence architecture from enteric interventions to manure, waste, anaerobic digestion, and farm energy. It preserves the same architectural discipline: historical evidence is immutable; interpretation is versioned; integrity, compatibility, verification, and reliance are distinct; minimum evidence is decision-linked; and physical outcomes remain separate from rights and claims.

The proposed graph records source materials, process configuration, sensors, laboratory results, operation, maintenance, mass and energy balances, leakage, co-products, LCA assumptions, environmental attributes, allocations, retirements, and reliance. Residuals remain visible rather than being forced to zero. One evidence root can be re-evaluated under changing methods, factors, boundaries, and attribute policies.

The next steps are to verify the literature, freeze the engineering receipt profile, select an approved data route, build the reference case, preregister reconstruction and exception tests, implement Australian and synthetic adapters, run independent review, publish conformance fixtures, disclose failure, and propagate architecture deltas through portfolio change control.

12 Declarations

Funding and conflicts of interest

TODO: Disclose AgEvidence support, employment, consulting, advisory roles, equity, intellectual property, equipment, energy-market, product, project, and verification relationships for every contributor.

Author contributions

TODO: Maintain a live CRediT contribution record. Confirm author order and corresponding-author responsibility only after written consent and final contribution review.

Data availability

TODO: Select a synthetic mirror, permissioned pilot subset, or approved post-pilot route. Document permissions, confidentiality, safety, de-identification, retention, and release conditions.

Code and specification availability

TODO: Archive schemas, engineering profiles, balance evaluator, adapter versions, conformance fixtures, example bundles, environment manifests, and checksums.

Rights and attributes

No physical-output, contribution, allocation, retirement, or reliance receipt in the study should be represented as a legal or governmental determination without the relevant authority.

Synthetic-case disclosure

Synthetic process data, sensors, outputs, attributes, and reliance events are test artifacts. They are not empirical findings or claims about named farms, technologies, projects, or institutions.

13 Executable TODO register

Table 4: Paper 6 executable TODO register.

Task	Owner	Exit evidence
Verify supplied bibliography and complete engineering review	Persona E and scientific editor	Publisher-checked library, evidence table, and revised claims.
Freeze Paper 6 engineering profile against Papers 1–4	Persona G and portfolio authors	Versioned mapping, schemas, decision log, and dependency review.
Select approved case and data route	Data Steward and executive sponsor	Written permissions, safety review, de-identification plan, or synthetic protocol.
Map system boundary and authority	Persona E and implementation team	Process-flow diagram, actor map, authority matrix, and source register.
Build material, sensor, and operation graph	Technical and engineering teams	Versioned bundle, manifest, source commitments, calibration records, and gap register.
Implement balance evaluator	Reproducibility engineer	Tested mass, methane, and energy calculations with pre/post reconciliation outputs.
Implement method, LCA, and attribute adapters	Technical team with scientific and legal review	Immutable profiles, tests, authority statements, and limitations.
Create exception and attribute fixtures	Authors and reproducibility engineer	Machine-readable cases, expected states, caps, compatibility outcomes, and rationale.
Preregister reconstruction and analysis	Methods team	Archived protocol, outcomes, thresholds, exclusions, and discrepancy rules.
Run independent reconstruction	Independent reviewers	Time logs, issue logs, outputs, residuals, discrepancies, and disagreement report.
Evaluate operator and producer burden	Persona F and implementation team	Workflow map, manual steps, maintenance burden, and redesign options.

Task	Owner	Exit evidence
Audit validity laundering and attribute overlap	Persona C and reviewers	Dashboard, PDF, API, claim, and rights audit with issue register.
Resolve deltas through portfolio change control	Scientific Chairs	Versioned decisions propagated to Papers 1–4 and Paper 7.
Complete CRediT, disclosures, and final approval	Persona E	Signed authorship, affiliation, conflict, and approval records.

References

Amon, B., V. Kryvoruchko, T. Amon, and S. Zechmeister-Boltenstern (2006). “Methane, Nitrous Oxide and Ammonia Emissions during Storage and Land Application of Dairy Cattle Manure—Impact of Slurry Treatment”. In: *Agriculture, Ecosystems & Environment* 112.2-3, pp. 153–162.

Angelidaki, I., M. Alves, D. Bolzonella, L. Borzacconi, J. L. Campos, A. J. Guwy, S. Kalyuzhnyi, P. Jenicek, and J. B. van Lier (2009). “Defining the Biomethane Potential (BMP) of Solid Organic Wastes and Energy Crops: A Proposed Protocol for Batch Assays”. In: *Water Science and Technology* 59.5, pp. 927–934.

Bacheson, M. and K. Fischer (2019). “Continuous Monitoring of Fugitive Methane Emissions from Farm-Scale Anaerobic Digesters: Sensor Drift, Failure Modes, and Calibration Protocols”. In: *Biosystems Engineering* 184. Bibliographic details supplied for this working draft; verification pending, pp. 112–124.

Bachmaier, H. and A. Gronauer (2007). “Mass and Energy Balances of Biogas Plants”. In: *Agricultural Engineering International: CIGR Journal* EE 07 004. Manuscript identifier supplied for this working draft; verification pending.

Basset-Mens, C. and H. M. G. van der Werf (2005). “Scenario-Based Environmental Assessment of Farming Systems: The Case of Pig Production in France”. In: *Agriculture, Ecosystems & Environment* 105.1-2, pp. 127–144.

Chadwick, D., S. Sommer, R. Berman, L. Cardenas, W. Bussink, J. Webb, et al. (2011). “Emissions of Ammonia, Nitrous Oxide and Methane from Agricultural Livestock Waste in Europe: A Review”. In: *Animal Feed Science and Technology* 166. Author list and page details supplied for this working draft; verification pending, pp. 514–531.

Daniel-Gromke, J., J. Liebetrau, V. Denysenko, and J. Clemens (2018). “Digester Leakage and Fugitive Methane Emissions from Agricultural Biogas Plants: Measurement Techniques and Emission Factors”. In: *Renewable Energy* 117. Bibliographic details supplied for this working draft; verification pending, pp. 345–355.

Department of Industry, Science, Energy and Resources (2023). *National Inventory Report 2021: The Australian Government Submission to the United Nations Framework Convention on Climate Change*. Departmental attribution supplied for this working draft; verify against the official report. Canberra: Australian Government.

Eckard, R. J., C. Grainger, and C. A. M. de Klein (2010). “Options for the Abatement of Methane and Nitrous Oxide from Ruminant Production: A Review”. In: *Livestock Science* 130.1–3, pp. 47–56.

Eide, M. H. (2002). “Life Cycle Assessment (LCA) of Milk Production and Processing in Norway”. In: *The International Journal of Life Cycle Assessment* 7.3, pp. 115–126.

Gillison, L. and M. Brander (2020). “Double Counting in Voluntary Carbon Markets and Scope 3 Supply Chain Reporting: Technical Barriers and Architectural Solutions”. In: *Climate Policy* 20.8. Bibliographic details supplied for this working draft; verification pending, pp. 982–995.

Guinée, J. B., ed. (2002). *Handbook on Life Cycle Assessment: Operational Guide to the ISO Standards*. Kluwer Academic Publishers.

Hassouna, M., P. Robin, A. Charpiot, N. Edouard, and C. Aubert (2016). “Assessment of Gaseous Emissions from Livestock Buildings and Manure Storages: A Review of Measurement Methods”. In: *Atmospheric Environment* 145, pp. 338–355.

Henry, B. K., D. G. Butler, and B. Wieland (2012). “Development of Australian National Greenhouse Gas Inventory Methodologies for Livestock Emissions”. In: *Animal Production Science* 52.4. Bibliographic details supplied for this working draft; verification pending, pp. 201–212.

Holm-Nielsen, J. B., T. Al Seadi, and P. Oleskowicz-Popiel (2009). “The Future of Anaerobic Digestion and Biogas Utilization”. In: *Bioresource Technology* 100.22. Title normalised from supplied description; verification pending, pp. 5478–5484.

International Organization for Standardization (2006). *ISO 14040:2006: Environmental Management—Life Cycle Assessment—Principles and Framework*. ISO.

Kist, A. A. and Y. Zhang (2021). "Internet of Things Architecture for Verified Agricultural MRV Pipelines: Ensuring Sensor Data Provenance and Cryptographic Chain of Custody". In: *Computers and Electronics in Agriculture* 188. Bibliographic details supplied for this working draft; verification pending, p. 106320.

Krauslyte, A. and K. Venslauskas (2020). "Mass and Energy Balances in Farm-Scale Anaerobic Digestion Systems: A Methodological Framework for System Boundary Definition". In: *Journal of Cleaner Production* 258. Bibliographic details supplied for this working draft; verification pending, p. 120689.

Lazarus, M., P. Erickson, and C. Chandler (2012). "Renewable Energy Certificates and Emission Offset Claims: Preventing Double Attribution in Integrated Grid Systems". In: *Energy Policy* 45. Bibliographic details supplied for this working draft; verification pending, pp. 310–319.

Liao, W., Y. Liu, C. Liu, Z. Wen, and S. Chen (2006). "Optimizing Slurry Stream Allocations in Dairy Farm Waste Management via Mass-Balance Modeling". In: *Bioresource Technology* 97.13. Bibliographic details supplied for this working draft; verification pending, pp. 1412–1420.

Owen, J. J. and W. L. Silver (2015). "Greenhouse Gas Emissions from Dairy Manure Management: A Review of Field-Based Studies". In: *Global Change Biology* 21.2, pp. 550–565.

Weiland, P. (2010). "Biogas Production: Current State and Perspectives". In: *Applied Microbiology and Biotechnology* 85.4, pp. 849–860.

Wiedemann, S. G., E. J. McGahan, and C. M. Murphy (2016). "Resource Use and Greenhouse Gas Intensity of Australian Beef Production Using Life Cycle Assessment". In: *Journal of Cleaner Production* 112, pp. 2427–2437.

World Resources Institute and World Business Council for Sustainable Development (2011). *Corporate Value Chain (Scope 3) Accounting and Reporting Standard*. WRI/WBCSD.

Source note

This expanded draft aligns with the architecture, validity-state separation, minimum-evidence framework, and applied claim-allocation discipline established in Papers 1–4. The Paper 6 literature compendium supplied by the user has been integrated into the introduction, scientific foundations, engineering model, methods, discussion, and bibliography. Bibliographic accuracy and source suitability must be independently verified before submission.

Paper 7: Portable Evidence, Local Authority

Working-draft note. This manuscript expands Paper 7 from the *AgEvidence Australia Scientific Publication Series: Seven Manuscript First Drafts*. It is a capstone interoperability, methods, and standards-roadmap manuscript. It does not report a completed cross-jurisdictional evaluation, recognised foreign compliance, government acceptance, verifier determination, standards adoption, or empirical portability result. Every title, author order, affiliation, adapter, jurisdiction, dataset, result, target journal, and ownership designation remains proposed.

Non-negotiable guardrail. A research adapter expresses a documented interpretation of selected published rules for architecture testing. It is not an authoritative national methodology, legal or regulatory compliance assessment, accredited validation or verification determination, government approval, registry eligibility decision, claim-rights determination, or substitute for competent local scientific and institutional review.

Abstract

International expansion in agricultural greenhouse-gas and Scope 3 programs often produces country-specific software branches, duplicated evidence pipelines, and inconsistent interpretations of the same underlying activity. This paper evaluates a different proposition: agricultural evidence can remain stable while method compatibility, verification expectations, rights, reporting artifacts, and institutional reliance remain explicitly local and versioned. The study freezes one Australian livestock evidence graph, identified by a cryptographic root digest, and evaluates it against an immutable Australian adapter and at least two synthetic or research-only foreign policy adapters. The underlying receipts, source commitments, event semantics, parent links, signatures, limitations, and offline integrity-verification process remain unchanged.

The architecture follows the thin-waist model developed in Paper 1, the validity separation formalised in Paper 2, the minimum-evidence vocabulary of Paper 3, and the applied evidence profiles of Papers 4–6. Each adapter declares its purpose, authority status, source documents, methodology and policy versions, production-context requirements, eligibility conditions, parameter and model bindings, evidence profile, verification expectations, claim policy, data-governance rules, artifact profile, limitations, and supersession relationships. The experiment compares which evidence fields are reusable, which require semantic mapping or local bindings, which produce missing, `conflicting`, out-of-domain, or conditionally sufficient states, and which require method extension. It also tests methodology supersession: a later adapter version generates a new determination linked to the same historical evidence rather than rewriting the original record.

The paper integrates FAIR data stewardship, agricultural knowledge graphs, international greenhouse-gas inventory tiering, Australian inventory methodology, Scope 3 accounting, digital provenance, greenhouse-gas verification, and data-sovereignty literature. It proposes a weighted and unweighted evidence-reuse profile, an adapter-difference taxonomy, a semantic-portability test, a conformance suite, a migration and supersession protocol, and an independent reconstruction study. The intended contribution is not a universal agricultural methodology. It is a testable contract for preserving evidentiary identity while making policy translation explicit, inspectable, and locally governed. Non-Australian adapters remain research artifacts unless reviewed and adopted by competent authorities.

Keywords: evidence portability; policy adapters; agricultural MRV; conformance; interoperability; semantic portability; methodology versioning; data sovereignty; local authority; Australia.

1 Manuscript status, scope, and contribution boundary

Paper 7 is the capstone interoperability manuscript in the seven-paper portfolio. It consumes the stable receipt envelope, append-only evidence graph, source commitments, immutable parent links, adapter interface, and bounded reliance artifacts from Paper 1; the separated integrity, method-compatibility, scientific-judgment, accredited-verification, and institutional-reliance states from Paper 2; the evidence classes, gap taxonomy, and decision-linked minimum requirements from Paper 3; and the applied livestock, grazing, manure, waste, and energy profiles developed in Papers 4–6. It must reuse those terms rather than create a parallel vocabulary.

The unit of analysis is not a country and not a methodology in the abstract. It is a frozen evidence graph G, a specific immutable adapter A_j^v, and the determination-and-artifact set produced by evaluating that graph under that adapter. The same historical evidence may therefore yield different compatibility states, missing-evidence findings, model-domain conclusions, verification expectations, and reliance artifacts without any implication that one adapter is globally correct or that all jurisdictions should converge.

The manuscript contributes:

- a formal distinction between evidentiary portability and policy equivalence;
- an adapter manifest that makes source authority, version, scope, local bindings, data rules, and non-assertions explicit;
- a semantic-portability and evidence-reuse evaluation conducted against one unchanged graph root;
- a taxonomy for explaining cross-adapter differences without collapsing them into pass/fail compliance;
- a methodology-versioning and supersession protocol that preserves historical determinations;
- a conformance suite for independent adapter implementations and offline reconstruction; and
- a standards-translation roadmap that preserves national scientific, legal, governmental, and verification authority.

The paper does not establish that Australian evidence is sufficient in another jurisdiction, that any foreign method accepts the represented activity, that a buyer may make a claim, that an accredited validation or verification body would issue a positive opinion, or that a proposed technical specification has become a formal standard. Tables labelled planned results are reporting templates rather than findings.

2 Provisional scientific author personas

Persona	Provisional role	Contribution domain / deliverable
Persona A	Co-lead/corresponding-author persona	Australian source context, local authority, national-method/standards boundaries and non-assertions.
Persona B	Co-lead/corresponding-author persona	Cross-country scientific comparability, parameterisation, calibration-domain and model-portability safeguards.
Persona C	Scientific coauthor persona	Institutional use, Scope 3, buyer/verifier artifacts and reliance portability.
Persona D	Scientific coauthor persona	Measurement provenance, uncertainty, gaps, sufficiency and semantic portability.
Persona E	Scientific coauthor persona	Engineering portability across manure, waste, digestion and farm-energy contexts.
Persona F	Scientific coauthor persona	Producer burden, local identity, production-system and operational portability.
Persona G	Technical scientific contributor persona	Adapter implementation, frozen graph root, metrics, conformance handbook and environment manifests.
Persona I	Invited local-review persona	Country-specific method, standards, verification, data-rights and legal review where qualifying.

Authorship control. Persona IDs are recruitment placeholders only. A real person is not an author until consent, qualifying contribution, affiliation, conflict disclosure, CRediT allocation, and final-manuscript approval are complete.

3 Introduction

3.1 The duplication problem in international agricultural MRV

International agricultural decarbonisation programs often share empirical inputs—animal identity or cohort, diet, intake, intervention identity, duration, productivity, measurement observations, model executions, source documents, and limitations—while differing materially in eligible activities, national emission factors, `approved` models, baseline periods, reporting boundaries, verification rules, rights, and claim language. A conventional implementation treats each national method or buyer program as the organising centre. Data are then collected into a country-specific form, spreadsheet, portal, or dossier. This may achieve local fit, but it fuses historical evidence to one interpretation and encourages duplicate capture whenever the method, buyer, verifier, or jurisdiction changes.

The resulting problem is not merely inefficient data exchange. It is loss of reconstructability. When a record has been transformed into a method-specific answer, later reviewers may be unable to distinguish the original observation from a mapping choice, default factor, policy assumption, model result, or institutional judgment. International scaling then becomes a sequence of migrations between opaque representations rather than a transparent re-evaluation of a stable evidentiary record.

3.2 FAIRness, knowledge graphs, and semantic portability

The FAIR Guiding Principles frame scientific data stewardship around findability, accessibility, interoperability, and reusability (Wilkinson et al., 2016). For this manuscript, FAIRness is necessary but not sufficient. A graph may be machine-readable and richly linked yet still fail to preserve the authority, temporal scope, calibration domain, rights, limitations, or policy version needed for a bounded greenhouse-gas decision. The relevant goal is therefore decision-preserving semantic portability: a competent reviewer should be able to determine what a field meant in its original context, how it was transformed, which local concept it maps to, and where equivalence cannot be asserted.

Cross-domain knowledge graphs such as KnowWhereGraph demonstrate how heterogeneous spatial and environmental data can be linked without forcing all source systems into one monolithic representation (Janowicz et al., 2022). Agricultural interoperability initiatives, including AgGateway and AEF Agricultural Interoperability Network specifications, similarly emphasise exchange across multi-vendor and distributed environments. Paper 7 adopts that general direction but narrows the object of standardisation. The thin waist standardises the evidentiary envelope and verification behavior; policy meaning remains in explicit adapters.

3.3 International agricultural accounting is tiered and local by design

The IPCC inventory framework distinguishes default, globally applicable Tier 1 approaches from progressively more country-specific Tier 2 and Tier 3 methods, parameters, models, and activity data (IPCC, 2019). This tiering is a warning against treating interoperability as universal numerical equivalence. Two jurisdictions may observe similar livestock activity but legitimately use different emission factors, productivity variables, herd structures, uncertainty treatments, and inventory boundaries. The correct architecture must therefore support common evidence where possible while preserving the versioned parameterisation and authority of each national accounting system.

The Australian National Inventory Report and associated methodology documentation illustrate the density of local definitions, source commitments, livestock categories, factors, and estimation procedures that sit above generic agricultural observations (DCCEEW, 2023). An Australia-first adapter can be explicit and testable, but Australian policy must not be embedded in the generic receipt kernel. Otherwise a purportedly global architecture would merely export one jurisdiction's assumptions under neutral terminology.

3.4 Scope 3 translation and the risk of double counting

Corporate Scope 3 accounting introduces an additional translation layer. The GHG Protocol Corporate Value Chain Standard and land-sector guidance distinguish corporate inventory boundaries, value-chain attribution, removals and land-sector treatment, and reporting requirements from national inventory accounting. A physical or modelled outcome at farm level is not automatically a corporate claim, and the same evidence may be relevant to several actors with different reporting boundaries. Paper 7 therefore treats Scope 3 artifact generation as an adapter- and institution-specific interpretation rather than a property of the underlying receipt graph.

This separation also protects against double counting and validity laundering. Cryptographic integrity does not establish method eligibility; method compatibility does not allocate environmental attributes; an accredited verification result does not force a buyer to rely; and one institution's reliance does not bind another. The cross-jurisdictional experiment must preserve these distinctions in every output.

3.5 Provenance, immutability, and data sovereignty

Digital provenance and distributed-ledger literature has highlighted the value of tamper-evident transaction histories in agricultural supply chains (Bager et al., 2021). In this manuscript, cryptographic commitments are used narrowly: they make later alteration detectable and permit independent reconstruction of the graph root. They do not prove that a measurement was scientifically valid, that an issuer was authorised, or that a claim is legally available.

Cross-border data use also raises sovereignty and localisation questions. Data sovereignty is not reducible to server location; it concerns legitimate authority, control, use conditions, and the institutional context in which data are governed (Hummel et al., 2021). Accordingly, the portability test separates semantic portability from physical data movement. An adapter may run against a local, permissioned copy or against disclosed commitments without requiring unrestricted export of personal, commercially sensitive, Indigenous, producer, or jurisdictionally regulated data.

3.6 Verification and standards conformance

ISO 14064-3 provides a benchmark for validation and verification of greenhouse-gas statements, including evidence, risk, materiality, and the relationship between a statement and the underlying support. Paper 7 does not operationalise or replace accredited verification. It uses the standard to motivate a conformance posture: outputs must expose scope, criteria, evidence limitations, version, responsible actor, and unresolved matters rather than presenting a generic 'verified' status.

The distinction between conformance and adoption is central. A software implementation may conform to the proposed adapter interface and test fixtures. That does not make the interface an ISO standard, Australian Standard, government requirement, or `accepted` registry rule. Peer-reviewed publication can make behavior citable and reproducible; formal adoption remains a separate institutional process.

3.7 Research questions

- RQ1. Can one frozen Australian evidence root be evaluated under multiple immutable policy adapters without changing receipt digests, source commitments, parent relationships, or declared provenance?
- RQ2. Which evidence concepts are semantically reusable across adapters, and which require terminology mappings, local identifiers, additional bindings, or method-specific evidence?
- RQ3. Can adapter differences be classified transparently into local authority, scientific, accounting, verification, data-policy, claim-policy, and artifact-profile categories?
- RQ4. Does methodology supersession produce a new determination against the same historical evidence while preserving the earlier adapter, output, reasons, and reliance history?
- RQ5. Can independent reviewers reproduce determinations and identify ambiguity without access to AgEvidence's hosted application?
- RQ6. Do interfaces and reports preserve the Paper 2 separation among cryptographic integrity, method compatibility, scientific judgment, accredited verification, and institutional reliance?
- RQ7. Can the architecture support data-localisation and sovereignty constraints without destroying semantic portability or forcing centralised custody?

4 Conceptual and formal portability model

4.1 Frozen evidence graph and adapter evaluation

Let $G = (V, E)$ denote the append-only evidence graph, and let $d_G = H(C(G))$ be the digest of its canonical representation C(G). Let A_j^v denote adapter j at immutable version v. Evaluation produces a determination D_j^v, an evidence-gap report M_j^v, and one or more bounded artifacts R_j^v:

$$(G, d_G, A_j^v) \longrightarrow (D_j^v, M_j^v, R_j^v).$$

For n adapter evaluations of the same frozen graph, the invariance requirement is:

$$d_G^{(1)} = d_G^{(2)} = \cdots = d_G^{(n)}.$$

An adapter may request additional local evidence. Such evidence is appended through new receipts, producing a successor graph G′ with a new digest $d_G{}'$ and an explicit parent relationship to the prior root. The adapter may not silently backfill, overwrite, reinterpret, or remove historical receipts in place.

4.2 Evidence portability is not policy equivalence

Define $P(e, a)$ as the portability state of evidence element e under adapter a. The proposed states are: `exact_reuse`, `mapped_reuse`, `reusable_with_local_binding`, `conditionally_reusable`, `not_required`, `insufficient`, `conflicting`, `out_of_domain`, `prohibited_by_data_policy`, and `not_assessed`. Exact reuse means the adapter consumes the original meaning without transformation other than canonical decoding. Mapped reuse requires an explicit semantic mapping. Local binding means the original evidence remains intact but must be linked to a jurisdiction-specific identifier, authority, factor, boundary, consent, or right. None of these states alone establishes eligibility.

Policy equivalence would require a much stronger proposition: that two adapters apply materially equivalent rules, authority, and decision consequences. This paper does not assume or attempt to prove policy equivalence. It measures whether differences can be expressed without rewriting evidence.

4.3 Reuse metrics

For adapter j, the unweighted reuse ratio is:

$$\rho_j = \frac{|F_{\text{reused},j}|}{|F_{\text{required},j}|}.$$

Because fields are not equally material, the study also reports a weighted reuse ratio:

$$\rho_j^w = \frac{\sum_f w_f I_{f,\text{reused}}}{\sum_f w_f I_{f,\text{required}}},$$

where weights w_f are preregistered by decision profile rather than chosen after results are known. The weighted ratio remains descriptive. A single non-reused field may be decisive—for example, legal authority, animal identity, baseline validity, model domain, or claim rights—so neither ρ_j nor ρ_j^w is a quality, compliance, or acceptance score.

To avoid hiding the mechanism of reuse, every ratio must be accompanied by a portability vector reporting counts and materiality by state. Reviewers should be able to distinguish genuine semantic reuse from superficial field-name matching.

4.4 Adapter-difference taxonomy

Table 2: Cross-adapter difference taxonomy.

Difference class	Question	Illustrative outcome
Terminology / ontology	Do two sources use different labels or conceptual granularity?	Mapped reuse with an explicit equivalence, broader-than, narrower-than, or non-equivalence relation.
Local identity	Is a jurisdiction-specific identifier required?	Additional binding to farm, animal, facility, product, organisation, or registry identifier.
Authority and rights	Is the issuer, consent, ownership, or entitlement recognised for this purpose?	Compatible evidence but unresolved authority or claim rights.
Production context	Does the adapter distinguish pasture, feedlot, dairy, housing, manure, or energy contexts differently?	Conditionally sufficient or out of domain.
Factor / model	Does the adapter require a different factor, equation, model version, or calibration domain?	Recalculation against unchanged observations; method extension if unsupported.
Baseline	Are periods, counterfactuals, leakage, permanence, or additionality defined differently?	New determination with baseline-specific gaps.
Evidence minimum	Does the decision require additional observations, duration, precision, or source authority?	Insufficient evidence or additional local binding.
Verification profile	What criteria, materiality, assurance level, scope, and verifier competence apply?	Verification-ready artifact differs; no generic verified status.
Data governance	May data be exported, disclosed, retained, or processed in the proposed environment?	Local execution, redaction, commitment-only disclosure, or prohibited transfer.
Claim policy	Who may report, allocate, co-claim, or retire an attribute?	Outcome retained; claim artifact withheld or bounded.
Artifact profile	What report, API payload, registry form, or assurance package is required?	Different R_j^v generated from the same evidence root.
Institutional acceptance	Will a named institution rely for a bounded purpose?	Accepted, `relied_on`, `rejected`, `needs_more_evidence`, `pending`, or `not_assessed`.

4.5 Adapter manifest

Each adapter is an immutable, signed, machine-readable package. At minimum it declares: adapter identifier and version; jurisdiction and purpose; authority status; source documents and exact versions; effective and retirement dates where applicable; production contexts; required evidence profile; terminology mappings; local identifiers and bindings; factors, models, parameters, and calibration domains; baseline and boundary rules; uncertainty treatment; verification criteria and artifact expectations; data residency, disclosure, retention, and processing constraints; claim and allocation rules; known limitations; prohibited inferences; test fixtures; supersedes/superseded-by links; and the identity and role of reviewers.

A research adapter must contain a prominent `authority_status=research_only` declaration and must not use labels such as `approved`, `compliant`, `certified`, `registry-ready`, or `government-aligned` unless a competent authority has supplied and `approved` that exact status.

4.6 Versioning and supersession

Adapter versions are immutable. If a source methodology, factor set, legal rule, reporting template, or interpretation changes, a new adapter version is issued. The new version declares its predecessor, source change, semantic change, migration notes, expected impact, and whether prior determinations remain valid, require reassessment, or become historically superseded for a specified purpose.

For graph G and adapter versions v and v+1, the system stores both outputs:

$$(G, A_j^v) \longrightarrow D_j^v \qquad \text{and} \qquad (G, A_j^{v+1}) \longrightarrow D_j^{v+1}.$$

The earlier determination is not deleted. Reliance events remain linked to the artifact actually reviewed at the time. This preserves an auditable distinction between 'the method later changed' and 'the historical evidence was altered.'

5 Methods

5.1 Study design

The study is a comparative architecture evaluation using one frozen Australian evidence graph, one Australian reference adapter, and at least two synthetic or research-only non-Australian adapters. The adapters are selected to maximise structural difference rather than to rank jurisdictions. Selection criteria include availability of public source material, relevance to livestock or agricultural greenhouse-gas accounting, distinct national parameterisation, differing verification or data-governance expectations, and access to qualified local reviewers. Final jurisdictions remain TODO until source review, local-expert participation, and legal-risk screening are complete.

The primary endpoint is not favourable compatibility. It is explainable divergence under graph invariance. The study succeeds when identical historical evidence yields reproducible, source-linked differences, including negative, `insufficient`, out-of-domain, or not-assessed outcomes, without changing the root digest or laundering one adapter's authority into another.

5.2 Freeze the Australian evidence root

The frozen bundle will be selected from an `approved` synthetic mirror, permissioned pilot subset, or post-pilot case. It will contain the canonical graph, receipt schemas, source commitments, parent links, trust material, identity and authority declarations, rights and consent records where permitted, model-execution receipts, limitations, supersession status, bundle manifest, expected integrity-verifier output, and disclosure statement. The freeze event will be signed, timestamped, and archived. A public release may substitute redacted payloads or cryptographic commitments where source rights or data-governance constraints prevent disclosure.

Before adapter evaluation, the team will run Paper 1 conformance checks, Paper 2 validity-state checks, and the relevant Paper 3–6 evidence-profile tests. Any unresolved defect is recorded rather than repaired silently after the root is frozen.

5.3 Build the Australian reference adapter

The Australian adapter is the first real reference implementation. It will be grounded in exact, archived versions of Australian inventory, method, program, and reporting sources selected for the declared decision purpose. DCCEEW methodology documentation provides the national accounting context but does not, by itself, establish project eligibility, Scope 3 claim rights, or verifier acceptance. The adapter must therefore separate inventory estimation, project or intervention compatibility, verification expectations, and institution-specific reliance.

The Australian adapter will be reviewed by Australian scientific, accounting, policy, and legal contributors. Review approval means that the research representation is sufficiently faithful for the study's stated purpose; it does not make the adapter an official government or standards artifact.

5.4 Build research-only foreign adapters

Each non-Australian adapter will be constructed from publicly available source documents and reviewed by at least one relevant local expert where feasible. Source extraction will record document title, issuing body, version, publication and effective dates, section or table references, and any ambiguity. Where a rule cannot be represented confidently, the adapter must return `not_assessed` or `requires_local_interpretation` rather than infer a favourable mapping.

Foreign adapters will not include private legal advice, unpublished regulator interpretations, or invented defaults. Their source hierarchy and limitations will be published. A local reviewer's contribution will be described accurately: technical review does not equal governmental endorsement.

5.5 Semantic mapping procedure

The team will create a cross-adapter concept matrix. For each required concept, reviewers classify the relation between the graph term and adapter term as equivalent, narrower, broader, overlapping, incompatible, or unresolved. Mappings will use persistent identifiers where available and will record transformation functions, units, temporal basis, and loss of information. Unit conversion is permitted only when dimensions, basis, and uncertainty remain explicit. A field-name match without conceptual review is not counted as semantic reuse.

The procedure will distinguish raw observations from derived values. For example, dry-matter intake observations may be reusable while an emission estimate must be recalculated under a local factor or model. The recalculated result is a new calculation receipt linked to the original evidence, not a replacement for it.

5.6 Adapter evaluation

The identical frozen graph will be evaluated under each adapter in a clean, version-pinned environment. For every required field, the evaluator records portability state, source receipt, mapping, transformation, local binding, materiality, and reason. It then produces the Paper 2 method-compatibility state, a gap report, a provenance-preserving calculation trace where applicable, and bounded artifacts.

The evaluator must preserve indeterminacy. Missing source data, unavailable keys, ambiguous terminology, incompatible units, unreviewed local law, or unresolved authority cannot be converted into a default pass. The exact adapter, software, dependency, model, and parameter versions are included in the environment manifest.

5.7 Data-localisation and sovereignty scenarios

At least three execution modes will be tested: central evaluation of a permitted disclosure bundle; local execution against data that do not leave the jurisdiction or custodian environment; and commitment-based verification in which selected payloads remain undisclosed but their integrity and linkage can be checked. The study will document which outputs remain reproducible under each mode and which scientific or assurance conclusions require access to underlying data.

These scenarios are architecture tests, not legal determinations. Any claim that a particular deployment complies with privacy, Indigenous data governance, sectoral, contractual, export-control, or localisation law requires separate competent review.

5.8 Supersession and migration test

At least one adapter will be revised through a controlled change. Candidate changes include an updated factor, altered baseline period, revised evidence minimum, model-version update, new reporting artifact, or changed data rule. The study will compare D_j^v and D_j^(v+1), classify every changed output, and verify that d_G remains unchanged. If the new adapter requires genuinely new evidence, the study will append that evidence to a successor graph and report both the original-root and successor-root evaluations.

Migration success requires: immutable prior adapter and outputs; explicit source and reason for change; machine-readable supersession links; reproducible re-evaluation; no silent mutation of past reliance events; and a human-readable explanation of changed consequences.

5.9 Conformance suite

The conformance suite will include positive, negative, boundary, and adversarial fixtures. Positive fixtures demonstrate exact reuse, mapped reuse, and local binding. Negative fixtures include altered receipts, missing parents, unauthorised issuers, incompatible units, unsupported factor substitution, hidden model-version changes, silent baseline replacement, omitted limitations, forbidden data export, and an artifact that collapses compatibility into 'verified.' Boundary fixtures test temporal cut-offs, version effective dates, mixed production systems, model-domain edges, and partially disclosed evidence.

A conforming implementation must reproduce canonical digests; reject or mark invalid altered evidence; preserve indeterminate states; expose adapter identity and authority status; produce source-linked reasons; keep prior versions available; generate bounded artifacts; and pass interface tests preventing cross-layer validity inference.

5.10 Independent reconstruction

Independent reviewers will receive the frozen bundle, adapters, conformance fixtures, protocol, and environment manifest but not access to AgEvidence's hosted application. They will reproduce integrity results, portability classifications, compatibility determinations, gap reports, and selected artifacts. Reviewers will record time, ambiguities, missing dependencies, undocumented assumptions, output discrepancies, and whether local-authority safeguards are visible.

Agreement will be reported by field and decision state. Where categorical agreement is appropriate, percent agreement and a chance-corrected statistic may be reported, subject to sample size and prevalence. Quantitative outputs will be compared using absolute and relative discrepancy with tolerance thresholds preregistered by quantity and unit. No statistical measure will convert unresolved authority or policy ambiguity into consensus.

5.11 Evaluation metrics

Table 3: Planned evaluation metrics and interpretive limits.

Metric	Definition	Interpretive limit
Graph invariance	Digest equality across adapter runs	Shows unchanged canonical graph, not truth or eligibility.
Unweighted reuse ratio ρ	Reused required fields / all required fields	Descriptive; treats fields equally.
Weighted reuse ratio ρw	Materiality-weighted reused fields / weighted required fields	Depends on preregistered weights; not compliance.
Portability-state profile	Distribution across exact, mapped, local-binding, conditional, `insufficient`, `conflicting`, out-of-domain, prohibited, not-assessed	Must be interpreted with reasons.
Semantic mapping loss	Count and severity of narrower, broader, overlapping, incompatible, unresolved mappings	Ontology judgment may require expert review.
Adapter divergence	Changed determinations or artifacts attributable to classified rule differences	Difference is not error.
Reconstruction agreement	Agreement between reference and independent outputs	Does not establish institutional acceptance.
Reconstruction effort	Reviewer time, interventions, undocumented assumptions	Context-dependent.
Supersession traceability	Completeness of version, source, reason, predecessor, impact, and retained outputs	Tests governance, not scientific merit.
Local-authority safeguard rate	Proportion of outputs correctly displaying research-only and bounded-purpose limitations	A display metric, not legal adequacy.
Data-governance compatibility	Execution and disclosure modes permitted under declared profile	Requires separate legal validation.

6 Worked architecture scenarios

6.1 Same observation, different emission factor

The graph contains livestock activity data and measurement provenance. The Australian adapter uses its declared factor or model profile; a research adapter requires a different factor. Both calculations link to the same observations, but the outputs and uncertainty differ. The system reports factor/model divergence rather than modifying the observations.

6.2 Local identifier absent

The evidence contains a stable producer and facility identity, but a research adapter requires a jurisdiction-specific registration number. The field is `reusable_with_local_binding`. Compatibility may remain `insufficient` until the binding is supplied. The absence does not invalidate the original identity receipt.

6.3 Reproducible model outside local calibration domain

A model execution is cryptographically intact and reproducible, but the research adapter does not recognise the production context or calibration domain. Integrity remains valid; compatibility becomes `outside_current_method` or `method_extension_required`. The interface must not display a favourable scientific status as foreign eligibility.

6.4 Method-compatible evidence, claim rights unresolved

Most physical and scientific evidence is reusable, but the local claim policy requires rights or allocation evidence not present in the graph. The outcome calculation remains visible; the claim artifact is withheld or marked `needs_more_evidence`.

6.5 Data export prohibited

The adapter's data-governance profile prohibits export of detailed farm records. Evaluation occurs locally, and only a signed determination, disclosure report, and permitted commitments are shared. Semantic portability is retained without unrestricted central custody.

6.6 Adapter supersession changes baseline

Version v uses one baseline period; version v+1 uses another. Both evaluate the same historical graph and produce separate determinations. The earlier result remains linked to any historical reliance event. New baseline evidence, if needed, is appended to a successor graph.

6.7 Artifact translation drops limitations

A country-specific report generator omits the research-only status and method-domain limitation. Conformance fails even if the numerical calculation is correct, because the artifact launders authority and removes material qualification.

6.8 One graph, two institutional outcomes

Two institutions receive artifacts generated under the same adapter. One accepts for supplier engagement; the other requests more evidence for audited Scope 3 reporting. The graph and compatibility state remain unchanged; reliance outcomes differ by bounded purpose.

7 Planned evaluation and results reporting

Planned results — not findings

No adapter evaluation, reuse ratio, conformance result, or cross-jurisdictional conclusion is reported in this draft. The final manuscript must publish or archive the frozen graph identifier, adapter versions, authority status, source hierarchy, limitations, portability-state profiles, compatibility outputs, artifacts, supersession effects, reconstruction results, failures, and deviations from protocol.

Table 4: Planned cross-adapter results template. Values are placeholders, not findings.

Category	**Australia adapter**	**Research adapter 1**	**Research adapter 2**	**Required interpretation**
Evidence root digest	TBD	TBD	TBD	Must remain identical for frozen-root comparison.
Adapter ID/version	TBD	TBD	TBD	Immutable and source-linked.
Authority status	TBD	research_only	research_only	No foreign compliance claim.
Unweighted reuse ratio ρ	TBD	TBD	TBD	Descriptive only.
Weighted reuse ratio ρw	TBD	TBD	TBD	Weights preregistered.
Exact/mapped/local-binding profile	TBD	TBD	TBD	Report counts and material reasons.
Compatibility state	TBD	TBD	TBD	Version-specific; not approval.
Missing/`conflicting`/out-of-domain evidence	TBD	TBD	TBD	Gap or method-extension need.
Data-governance execution mode	TBD	TBD	TBD	Central, local, commitment-based, or prohibited.
Artifact profile	TBD	TBD	TBD	Institution- and purpose-specific output.
Independent reproduction	TBD	TBD	TBD	Agreement, ambiguity, effort, and failures.
Supersession effect	TBD	TBD	TBD	Prior output retained and change explained.

8 Discussion

8.1 Interoperability without uniformity

The central proposition is that interoperability should be evaluated by preservation of evidence identity and explainability of difference, not by forcing identical answers. The architecture resembles a universal shipping container rather than a universal customs law: the container standard makes custody, dimensions, and handling predictable, while each jurisdiction retains authority over admissibility, inspection, duties, and use. The thin waist is valuable precisely because it is narrow. If it absorbs national factors, eligibility rules, claims, or verification judgments, it ceases to be neutral infrastructure and becomes an exported policy regime.

This interpretation is consistent with FAIR and knowledge-graph approaches but adds a stronger governance boundary. Machine-readable linkage can make evidence discoverable and reusable; only explicit adapters can show which local meanings, authorities, and decision consequences apply.

8.2 International accounting differences are not schema defects

IPCC tiering makes clear that country-specific methods and parameters are expected, not accidental deviations from a universal formula. Cross-adapter divergence may therefore reveal legitimate differences in livestock categories, factors, activity data, uncertainty, baselines, or reporting boundaries. The portability benchmark should not penalise an adapter merely because it produces a different estimate. It should penalise hidden transformation, undocumented defaults, unverifiable mappings, and the inability to explain why the result changed.

The Australian adapter's role is to provide a concrete, reviewable implementation. It is not the global reference truth. Foreign research adapters are useful only when they remain visibly provisional and invite correction by competent local experts.

8.3 Scope 3 portability requires bounded artifacts

A farm-level observation may be relevant to a national inventory, project quantification, supplier program, product footprint, lender assessment, or Scope 3 report, but those uses do not share identical boundaries or rights. The architecture therefore treats artifacts as projections, not as the evidence itself. Each projection must declare purpose, adapter, period, scope, limitations, allocation policy, and reliance status. This reduces the risk that a valid physical outcome is converted into an unbounded corporate claim.

The aggregate-cap and co-claim controls developed in earlier applied papers remain relevant, but Paper 7 must not imply that one allocation rule is globally authoritative. Adapters may express different claim policies; unresolved entitlement remains unresolved evidence, not a software exception to be filled by convenience.

8.4 Method change as re-interpretation

Methodology change is often treated as a data-migration event. In a provenance-preserving system it is primarily a re-interpretation event. The historical graph remains the record of what was observed, asserted, calculated, and decided. A new method version creates a new determination. Where new observations are needed, they are appended. This makes scientific and policy evolution visible and protects historical users from the false impression that the earlier decision was made under today's rules.

Versioning also improves reproducibility. A reviewer can rerun the old adapter, rerun the new adapter, compare outputs, and attribute differences to sources and rules. Without immutable versions, 'current result' becomes an unstable moving target.

8.5 Data sovereignty as an architectural requirement

Portability is frequently conflated with centralisation. That is unnecessary. A receipt graph can be logically portable while payloads remain under local custody. The adapter, verifier, and artifact contract should support local execution, selective disclosure, redaction, and commitment-based checking. This is especially important where producer data, personal information, Indigenous data, commercial terms, or national rules constrain cross-border transfer.

However, cryptographic commitments do not solve governance by themselves. A digest can prove that a disclosed object matches a prior commitment; it cannot establish that collection, processing, retention, or sharing was legitimate. Consent, authority, purpose, and rights must remain first-class evidence and policy elements.

8.6 Conformance is behavior, not branding

A credible interoperability claim requires executable tests. Merely adopting the words receipt, graph, adapter, or verified does not demonstrate conformance. Implementations must reproduce canonical commitments, preserve limitations, expose authority status, retain superseded outputs, and fail safely when evidence or policy is indeterminate. Independent implementation is a key test because an architecture controlled by one vendor can become a proprietary export format even when its schema is published.

The standards roadmap should therefore prioritise open fixtures, reference outputs, failure cases, versioning rules, and governance disclosures. Formal standards bodies, governments, registries, and accredited verification ecosystems may later choose to review or adopt parts of the architecture, but publication must not pre-empt that authority.

8.7 Interpretation of reuse ratios

Reuse ratios are useful for diagnosing where duplication occurs, but they are easy to misuse. A ratio of 0.90 may conceal one decisive missing right or baseline requirement. A ratio of 0.40 may reflect a genuinely different policy purpose rather than poor architecture. The ratio should be read like a map of shared infrastructure, not a grade. The substantive result is the portability-state profile and the reasoned account of decisive differences.

9 Limitations

Synthetic or research-only foreign adapters cannot establish legal, regulatory, scientific, accounting, registry, or verifier adequacy. Public source documents may be incomplete, outdated, translated imperfectly, or supplemented by unpublished institutional practice. Local experts may disagree, and reviewer participation does not constitute government endorsement.

The selected Australian graph may not represent all production systems, interventions, data custodians, or rights structures. Feed-additive, grazing, manure, waste, and energy evidence have different operational and scientific requirements. A portability result on one graph cannot establish universal agricultural coverage.

The reuse metrics depend on the required-field set, semantic granularity, and materiality weights. Ontology mappings involve judgment. Independent reconstruction tests disclosure and software behavior, not whether a competent authority or accredited verifier would accept the evidence.

Cryptographic integrity protects against undetected alteration under declared algorithms and trust material, but it does not prove factual truth, scientific validity, lawful authority, or complete disclosure. Data-localisation scenarios are technical demonstrations and require separate legal and governance assessment before deployment.

The literature integrated in this draft supplies conceptual foundations, not a completed systematic review. Exact versions and final bibliographic details for AgGateway/AEF AgIN specifications, GHG Protocol land-sector guidance, DCCEEW methodology documents, and any jurisdiction-specific adapter sources must be verified and archived before submission.

10 Conclusion

Agricultural evidence can be portable without making policy universal. By freezing one Australian evidence root and applying multiple immutable adapters, the proposed study tests whether historical observations, assertions, calculations, provenance, and limitations remain stable while local methods, parameters, rights, verification expectations, data rules, claims, and artifacts remain explicit. The architecture's success condition is not identical cross-country outcomes. It is reproducible, source-linked, authority-preserving explanation of difference.

The capstone contribution is a conformance method for distinguishing genuine evidence portability from country-specific duplication and from superficial schema matching. If independent reviewers can reproduce divergent determinations from the same root, trace those differences to declared adapter rules, preserve superseded versions, and identify where local authority is indispensable, the portfolio will have established a practical route from Australia-first evidence architecture to internationally reusable infrastructure without displacing national scientific, governmental, legal, or verification authority.

11 Declarations

Synthetic adapter disclosure. Non-Australian adapters are research artifacts designed to test architecture. They are not legal or regulatory compliance assessments and require local scientific and institutional review before operational use.

Compatibility disclosure. A compatibility result indicates how declared evidence maps to a specified method and policy version. It is not government approval, an accredited validation or verification determination, legal advice, or a determination of claim ownership.

Results disclosure. Sections labelled planned results, expected outputs, templates, or placeholders do not report completed empirical findings.

Authorship disclosure. Advisory status alone does not confer authorship. All named authors remain proposed `pending` written consent, qualifying contribution, affiliation approval, conflict disclosure, and final-manuscript approval.

Code and data availability. TODO: Publish the frozen graph, adapters, fixtures, outputs, and environment manifest where source rights, confidentiality, privacy, data sovereignty, and contractual obligations permit. Where payloads cannot be released, publish redacted fixtures or commitments sufficient to test declared behavior.

Competing interests and funding. TODO: Complete author-specific disclosures, AgEvidence relationships, funder role, publication independence, software and intellectual-property interests, and any relationships with intervention suppliers, buyers, registries, verifiers, or standards bodies.

12 Executable TODO register

Table 5: Executable Paper 7 TODO register.

Work item	Accountable owner(s)	Exit condition
Freeze terminology and dependencies	Scientific co-leads and Persona G	Confirm Paper 1–6 vocabulary, states, receipt families, gap taxonomy, and applied profiles; record any required portfolio-wide version change.
Complete directed literature review	Named scientific owners	Verify and expand literature on FAIR/semantic interoperability, agricultural ontologies, international GHG inventory systems, policy-as-code, conformance testing, methodology versioning, Scope 3, verification, and data sovereignty.
Freeze Australian evidence root	Persona G, Data Steward; approval Persona A	Select `approved` data route; run conformance checks; sign and archive graph, manifest, trust material, limitations, and disclosure bundle.
Define Australian adapter purpose	Persona A and Australian reviewers	Select exact decision context and source hierarchy; prevent conflation of national inventory, project method, verification, and Scope 3 use.
Select two research contexts	Persona B	Choose jurisdictions based on public sources, structural contrast, local-expert availability, and legal-risk review.
Build concept matrix	Persona D, Persona E, Persona F	Classify exact, mapped, local-binding, conditional, `insufficient`, `conflicting`, out-of-domain, prohibited, and not-assessed states.
Implement immutable adapters	Persona G	Create manifests, source links, rules, fixtures, authority_status fields, and machine-readable outputs.
Run data-sovereignty scenarios	Data Steward and legal reviewer	Test central, local, selective-disclosure, and commitment-based modes; document technical limits without making legal claims.
Run supersession test	Persona G and source-domain owner	Issue a new adapter version; compare outputs; retain prior determinations and reliance links.
Independent reconstruction	Independent reviewers	Reproduce digests, portability states, compatibility, gaps, artifacts, and migration effects without hosted-platform access.
Interface and claims audit	Persona C, Persona A, legal reviewer	Check every display and artifact for validity laundering, false compliance, hidden limitations, and unbounded claim language.
Draft conformance handbook	Scientific Chairs, Persona G, Persona H	Publish normative behaviors, optional profiles, fixtures, failure cases, versioning, governance, and standards non-assertions.
Complete publication governance	Co-leads	CRediT, author consent, affiliations, conflicts, funding, data/code statement, ethics, permissions, journal fit, and final approval.

References

Bager, S. L., et al. (2021). Blockchain applications in agricultural supply chains: A policy perspective. Food Policy, article 102086. [Final volume, issue, pages, DOI, and title wording to verify before submission.]

Australian Government Department of Climate Change, Energy, the Environment and Water (DCCEEW). (2023). National Inventory Report: Australian methodology for the estimation of greenhouse gas emissions. [Exact report volume, year designation, URL, and archived version to verify.]

GHG Protocol; World Resources Institute; World Business Council for Sustainable Development. (2011). Corporate Value Chain (Scope 3) Accounting and Reporting Standard.

GHG Protocol. (2022 or current final version). Land Sector and Removals Guidance. [Publication status, final title, date, and version to verify before citation.]

Hummel, P., Braun, M., Tretter, M., & Dabrock, P. (2021). Data sovereignty: A concept in need of clarification. Big Data & Society, 8(1).

Intergovernmental Panel on Climate Change (IPCC). (2019). 2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories, Volume 4: Agriculture, Forestry and Other Land Use.

International Organization for Standardization (ISO). (2019). ISO 14064-3: Greenhouse gases—Part 3: Specification with guidance for the verification and validation of greenhouse gas statements.

Janowicz, K., et al. (2022). KnowWhereGraph: Enriching and linking cross-domain spatial data. Journal of Web Semantics, 73, 100730.

Wilkinson, M. D., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3, 160018.

AgGateway / Agricultural Industry Electronics Foundation (AEF). (2024–2025). Agricultural Interoperability Network (AgIN) specifications for distributed agricultural data ecosystems. [Exact issuing body, specification title, release number, date, and stable citation to verify before submission.]

Appendix A: Literature Review and Reference Architecture

Appendix orientation. This working literature-review, journal-targeting, governance, and reference-architecture material is retained as the research substrate behind the seven-paper program. It has been moved to the end so that first-time readers encounter AgEvidence and the manuscript sequence before the supporting research-planning apparatus.

A.1 Document Purpose and Academic Posture

This working document converts the seven-paper AgEvidence Australia portfolio into an academic-facing research program. It combines four functions that should remain operationally linked but scientifically distinct:

1. a reference-architecture thesis for country-neutral agricultural evidence;
2. a rigorous literature-review and source-verification program;
3. a journal-by-journal targeting and manuscript-hardening guide; and
4. a portfolio governance system with named accountabilities, dependencies, and submission gates.

The portfolio is cumulative rather than a collection of independent marketing essays. Paper 1 defines the evidence architecture. Paper 2 defines the validity and governance vocabulary. Paper 3 establishes a proposed minimum-evidence framework for enteric-methane interventions. Papers 4 and 5 test the framework against controlled feeding and extensive grazing. Paper 6 tests whether the same primitives remain useful for manure, waste, anaerobic digestion, and farm energy. Paper 7 freezes an Australian evidence root and evaluates portability across versioned research adapters.

A.1.1 Academic claims discipline

Every sentence in a manuscript should be classified before submission:

Claim class	Meaning	Permitted support
Observed	Direct result from a defined dataset or experiment.	Versioned data, documented methods, statistical analysis, and reproducible output.
Calculated	Result obtained from declared equations, factors, or models.	Inputs, model identity/version, assumptions, calibration domain, uncertainty, and code.
Synthesized	Conclusion derived from a structured literature review or consensus process.	Search protocol, screening flow, extraction form, appraisal, synthesis method, and dissent.
Proposed	Architecture, framework, taxonomy, or hypothesis advanced for testing.	Transparent rationale, formal definition, comparison, test plan, and explicit non-assertions.
Institutional	Decision or acceptance by a named institution for a bounded use.	Identified policy version, actor authority, date, scope, conditions, and decision record.
Commercial	Product capability, service level, or implementation claim.	Kept outside scientific conclusions unless directly evaluated under a disclosed protocol.

A.1.2 Statements that must be removed or rewritten for journals

The phrase “standards trap” should not appear in manuscripts, cover letters, supplementary files, repositories, reviewer responses, or conference abstracts. The academically defensible formulation is that an openly specified and independently implementable reference architecture may create a common vocabulary and comparative conformance benchmark. Scientific adoption must result from usefulness, reproducibility, criticism, and independent reuse rather than strategic coercion.

Similarly, the following statements require caution:

- A cryptographically valid receipt does not establish that a physical event occurred.

- Method compatibility does not establish accredited verification or government acceptance.
- A buyer's reliance decision does not bind another buyer, registry, VVB, lender, or government agency.
- A synthetic adapter demonstrates software behavior, not legal or scientific adequacy in another jurisdiction.
- A schema is not a scientific standard merely because it is machine-readable or peer reviewed.

A.2 Executive Scientific Architecture

Agricultural measurement, reporting, and verification systems are fragmented across farm software, sensors, laboratories, quantification models, registries, verifiers, contracts, and corporate reporting systems. The research hypothesis is that tightly coupling historical evidence to a single methodology or institutional workflow creates avoidable reconstruction cost and obscures the boundary between evidence integrity, scientific compatibility, verification, and reliance.

The proposed thin-waist model places a small, stable evidence contract between heterogeneous source systems below and versioned country, method, and institutional policies above. Primary evidence is captured and canonicalized once, then evaluated many times under explicit policy and model versions.

Tier	Primary responsibilities	Prohibited responsibilities	Primary outputs
Global evidence layer	Canonical envelope; stable event semantics; issuer, subject, identity, authority, rights, time, sources, parent links, payload digests, signatures, limitations, and manifests.	National emissions factors, legal eligibility, registry states, claim ownership, or final scientific conclusions.	Signed receipts, append-only evidence graph, offline bundle, disclosure manifest.
Country and method layer	Versioned method mapping; input requirements; baseline and model execution; uncertainty; gap detection; production-context suitability.	Rewriting historical source evidence or presenting a local method as universally applicable.	Compatibility determination, model log, gaps, exclusions, suitability profile.
Institutional reliance layer	Decision-specific policy application; contractual boundary; claim allocation; double-claim controls; reliance artifact assembly.	Rewriting the cryptographic chain or universalizing one institution's acceptance.	Reliance record, verification-readiness packet, buyer artifact, allocation and retirement records.

A.2.1 Tripartite validity state model

Let E represent a canonical evidence envelope and graph, σ the signatures and commitments, M_v a named scientific model at version v, $P_{j,w}$ the policy adapter for jurisdiction or institution j at version w, and C contractual rights and allocations. The composite result is a tuple, not a single badge:

$$\nu(E, \sigma, M_v, P_{j,w}, C) = \big(\Phi(E, \sigma), \Psi(E, M_v), \Omega(E, M_v, P_{j,w}, C)\big).$$

$$\Phi \in \{\texttt{valid}, \texttt{invalid}, \texttt{indeterminate}\},$$

$$\Psi \in \{\texttt{compatible}, \texttt{conditional}, \texttt{out\_of\_domain}, \texttt{insufficient\_evidence}\},$$

$$\Omega \in \{\texttt{relied\_upon}, \texttt{needs\_evidence}, \texttt{policy\_rejected}, \texttt{claim\_cap\_exceeded}\}.$$

Prohibited inference rules include:

$$\Phi = \texttt{valid} \not\Rightarrow \Psi = \texttt{compatible},$$

$$\Psi = \texttt{compatible} \not\Rightarrow \Omega = \texttt{relied\_upon},$$

$$\Omega_j = \texttt{relied\_upon} \not\Rightarrow \Omega_k = \texttt{relied\_upon} \quad (j \neq k).$$

A positive reliance outcome should require, at minimum,

$$\Omega_j = \texttt{relied\_upon} \Rightarrow \Phi = \texttt{valid} \wedge \Psi \in \{\texttt{compatible}, \texttt{conditional}\} \wedge P_{j,w}(E, C) = \texttt{pass}.$$

Where an institutional policy allocates a recognized quantity among overlapping value-chain participants, the proposed aggregate cap is

$$\sum_{i=1}^{n} q_i \leq Q_{\text{recognized}}.$$

This is a proposed control invariant. Its relationship to specific GHG Protocol, SBTi, ISO, registry, VVB, and contractual requirements must be established through source-specific legal and technical analysis rather than asserted generically.

A.3 Literature Review Architecture

The portfolio requires a federated review rather than one undifferentiated bibliography. Each paper should maintain its own review protocol and extraction table while contributing to a shared claims register.

A.3.1 Review streams

Stream	Core questions	Primary disciplines and sources
Protocol architecture	What makes a narrow interoperable waist stable? How are extensions, versioning, conformance, and ossification managed?	Internet architecture, end-to-end principle, hourglass models, protocol layering, distributed systems, formal specifications.
Digital provenance	How are identity, authority, source commitments, lineage, correction, and reproducibility represented?	W3C provenance and credentials, cryptographic receipts, event sourcing, append-only logs, data lineage, research objects.
Agricultural MRV	How are observations, factors, models, uncertainty, verification, and claims currently coupled?	IPCC agriculture guidance, national inventory methods, ISO GHG standards, registry methodologies, corporate accounting guidance, VVB practice.
Scientific validity	How should integrity, model-domain compatibility, assurance, and institutional use be distinguished?	Metrology, model risk, validation and verification, philosophy of measurement, assurance, conformity assessment, decision theory.
Enteric methane	What evidence supports efficacy, safety, persistence, dose response, and production-context transfer?	Rumen microbiology, animal nutrition, methane measurement, 3-NOP, Asparagopsis, nitrates, meta-analysis, commercial implementation.
Grazing systems	Which spatial, temporal, biological, and operational data are necessary and feasible in extensive systems?	Rangeland ecology, remote sensing, herd movement, pasture biomass, intake estimation, producer workload, digital agriculture.
Manure and energy	Which physical observations and balances support reconstruction of engineered interventions?	Anaerobic digestion, manure management, mass and energy balances, sensors, lifecycle assessment, renewable-energy attributes.
Claims and reliance	How are rights, co-claims, allocation, double counting, retirement, and buyer reliance governed?	Scope 3 accounting, insetting, chain of custody, book-and-claim, contract law, registry governance, assurance.
Cross-jurisdiction portability	Which evidence fields can be reused and which require local scientific, policy, or legal interpretation?	Comparative policy, interoperability, standards, regulatory technology, data governance, international agricultural systems.

A.3.2 Databases and search sources

At minimum, search Web of Science, Scopus, CAB Abstracts, AGRICOLA, PubMed, IEEE Xplore, ACM Digital Library, ScienceDirect, SpringerLink, Crossref, Google Scholar for citation chaining, and official repositories of ISO, GHG Protocol, SBTi, IPCC, Australian government agencies, relevant standards bodies, VVB accreditation bodies, and registry or methodology owners.

Commercial blogs and vendor summaries may be used to discover terms but should not carry central scientific claims. ISO interpretations should be anchored in official standards or authoritative accredited-body material. Product efficacy should be anchored in peer-reviewed primary trials and systematic reviews, with sponsor relationships disclosed.

A.3.3 Core search blocks

Each search should combine four blocks where applicable:

1. **Domain:** agriculture, livestock, dairy, beef, grazing, manure, anaerobic digestion.
2. **Evidence function:** provenance, traceability, MRV, validation, verification, assurance, lineage, audit.
3. **Architecture:** interoperability, protocol, schema, event sourcing, knowledge graph, digital signature, policy-as-code.
4. **Decision use:** Scope 3, claim, insetting, registry, VVB, buyer, reporting, reliance, double counting.

Example Boolean search:

> (livestock OR cattle OR dairy OR beef OR grazing) AND (MRV OR "measurement reporting verification" OR provenance OR traceability) AND (interoperab* OR architecture OR "knowledge graph" OR "event sourcing" OR "digital signature") AND (carbon OR methane OR "Scope 3" OR claim* OR assurance)

A.3.4 Review protocol TODOs

- ☐Register separate protocols for Papers 1–3 before screening begins. **Owner:** Scientific technical editor and paper leads
- ☐Define inclusion/exclusion criteria, dates, languages, study types, and grey-literature rules. **Owner:** Paper leads
- ☐Create one deduplicated citation library with paper-specific tags. **Owner:** Publication Program Manager
- ☐Create standardized extraction forms for architecture, method, intervention, evidence field, actor, decision, uncertainty, and limitation. **Owner:** Data Steward and technical editor
- ☐Use dual screening for central review questions and report adjudication. **Owner:** Paper lead plus independent reviewer
- ☐Record funding source, author conflicts, product sponsorship, and dataset accessibility for intervention studies. **Owner:** Data Steward
- ☐Maintain a living evidence map rather than silently updating narrative text. **Owner:** Reproducibility engineer
- ☐Archive exact search strings, dates, database exports, screening decisions, and excluded-full-text reasons. **Owner:** Publication Program Manager

A.4 Enteric Methane Literature Review Directions

The review must distinguish mechanism, controlled efficacy, commercial effectiveness, safety, persistence, measurement validity, and production-system transfer. Numerical efficacy ranges in early drafts should be treated as provisional until a structured extraction and meta-analytic assessment are completed.

A.4.1 3-NOP

Review directions:

- Mechanism of methyl-coenzyme M reductase inhibition and active-site chemistry.
- Dose-response by animal class, diet, dry matter intake, and basal methane yield.
- Production versus yield versus intensity outcomes.
- Dairy, feedlot, forage-fed, pasture, and delivery-system differences.
- Duration, adaptation, washout, and within-day exposure continuity.
- Milk, meat, animal health, intake, production, and residue evidence.

- Model-domain boundaries and whether diet composition modifies effect.
- Sponsor involvement and regulatory dossier versus peer-reviewed evidence.

The recent meta-analytic literature reports material average methane reductions but also diet- and dose-dependent heterogeneity. The portfolio should not reproduce a single universal reduction factor. It should extract study-level intervention, diet, animal, measurement, and duration variables and test whether the proposed evidence schema captures the moderators used in published models.

A.4.2 Asparagopsis and halogenated macroalgae

Review directions:

- Species, cultivation, processing, bromoform concentration, stability, and batch variation.
- Inclusion rate, diet, delivery form, and intake continuity.
- Methane production, yield, and intensity with confidence intervals.
- Dry matter intake, productivity, animal health, residues, iodine, bromide, and food-safety outcomes.
- Pasture versus confinement evidence and commercial scalability.
- Bromoform loss during storage and feed preparation.
- Environmental burden of cultivation, processing, and transport.

The 2025 PNAS grazing-beef study should be treated as a central applied source, but not as proof of general effectiveness across Australian extensive systems. Its intervention formulation, study population, duration, measurement technology, feeding control, and external validity should be extracted explicitly.

A.4.3 Nitrates and alternative hydrogen sinks

Review directions:

- Mechanistic competition for reducing equivalents.
- Adaptation protocols, toxicity, methemoglobinemia, and monitoring.
- Dose, basal diet, sulfur interactions, productivity, and persistence.
- Practical delivery and farmer safety in extensive systems.
- Evidence requirements for acclimation, intake continuity, and adverse-event recording.

A.4.4 Translation-gap evidence matrix

Evidence domain	Controlled-trial evidence	Commercial translation evidence
Intervention identity	Chemical/product specification and nominal dose.	Batch identity, active-agent assay, storage, degradation, delivery device, procurement chain.
Exposure	Assigned ration or treatment.	Individual/cohort intake distribution, missed days, access, refusal, device performance.
Animal context	Breed, weight, physiological state, health.	Cohort continuity, movements, substitutions, morbidity, production stage.
Diet	Formulated ration and analyzed composition.	Forage variability, supplements, NDF and other moderators, seasonal changes.
Measurement	Respiration chamber, GreenFeed, tracer, or other protocol.	Device calibration, sampling schedule, missing intervals, proxy or model use.
Persistence	Trial duration and treatment period.	Longitudinal continuity, interruption, recovery, and seasonal operation.

Evidence domain	Controlled-trial evidence	Commercial translation evidence
Decision use	Efficacy estimate.	Method compatibility, verification readiness, buyer policy, claim allocation, limitations.

A.5 Journal Targeting and Hardening Matrix

Journal selection must occur before final methods are fixed. The same manuscript cannot credibly target an agricultural-computing journal, an animal-science journal, and a policy journal without changing its research question, data, evaluation, and central contribution.

A.5.1 Paper 1: Thin-Waist Architecture

Status: Substantive architecture draft; requires implementation and independent evaluation.

Computers and Electronics in Agriculture

Best-fit contribution: implemented agricultural-computing architecture.

- ☐Build end-to-end reference implementation and publish schemas, test vectors, and verifier.
- ☐Use one realistic Australian livestock graph and at least two source-system pathways.
- ☐Compare with a document dossier, method-specific database, and conventional export/API baseline.
- ☐Report reconstruction time, validation latency, graph completeness, missing-dependency detection, and unsupported-inference detection.
- ☐Conduct ablation tests removing authority, provenance, parent links, and adapter commitments.
- ☐Include practitioner usability and agricultural workflow evaluation.

Primary rejection risk: Conceptual architecture or cryptographic novelty without substantial agricultural experimentation.

Environmental Modelling & Software

Best-fit contribution: formal, executable evidence-policy architecture for reproducible environmental modelling.

- ☐Formalize adapter semantics and deterministic execution.
- ☐Bind model outputs to inputs, identity, version, assumptions, calibration domain, uncertainty, and review state.
- ☐Execute method supersession on one frozen evidence root.
- ☐Publish complete reproducibility package and independent rerun.
- ☐Compare with existing environmental workflow and provenance systems.

Primary rejection risk: The architecture manages evidence around models but does not advance environmental modelling or software methodology.

Agricultural Systems

Best-fit contribution: multi-actor evidence infrastructure within livestock systems.

- ☐Map producer, manufacturer, processor, modeler, verifier, buyer, and government-facing decisions.
- ☐Quantify duplication, burden, incentives, and distribution of costs and benefits.
- ☐Include at least two contrasting production contexts and stakeholder research.
- ☐Analyze system-level consequences rather than software performance alone.

Primary rejection risk: Software-centric treatment without farm-system or actor-level findings.

Journal of Environmental Management

- ☐Define a specific environmental-management decision and decision-support baseline.
- ☐Demonstrate improved gap detection, reassessment, uncertainty visibility, or auditability.
- ☐Include management implications and implementation requirements.

Primary rejection risk: Data architecture without a demonstrated environmental-management contribution.

Smart Agricultural Technology

- ☐Emphasize applied deployment, integrations, field workflow, offline verification, cost, and usability.
- ☐Include practical smart-system components and an operational demonstrator.

Primary rejection risk: Infrastructure remains too abstract for a farm-scale technology journal.

Exit gate: Working implementation; frozen specification; conformance corpus; independent reproduction; quantified baseline comparison. **Owner:** Persona A, Persona G, reproducibility engineer; journal reconnaissance by Publication Program Manager.

A.5.2 Paper 2: Beyond "Verified"

Status: Conceptual state-model draft; no completed elicitation, inter-rater study, or interface audit.

Environmental Science & Policy

- ☐Build a corpus of policies, protocols, reports, assurance statements, and interfaces that use "verified."
- ☐Recruit balanced scientific, VVB, government, buyer, producer, registry, and public-interest participants.
- ☐Preregister coding and expert-elicitation methods.
- ☐Report inter-rater agreement, disagreement by actor type, and policy consequences.
- ☐Analyze legitimacy, accountability, contestability, and burden.

Primary rejection risk: A terminology essay without empirical policy analysis.

Journal of Environmental Management

- ☐Conduct a controlled decision experiment comparing binary badges with layered states.
- ☐Measure decision accuracy, confidence calibration, unsupported inference, time, and uncertainty recognition.

Primary rejection risk: Taxonomy without demonstrated management consequences.

Sustainability Science

- ☐Add transdisciplinary co-production and analyze credibility, legitimacy, salience, justice, and power.
- ☐Include producer and non-commercial voices and more than one institutional context.

Primary rejection risk: Too technical, narrow, or commercially situated.

Environmental Modelling & Software

- ☐Define formal state machines, transition rules, prohibited transitions, APIs, and conformance tests.
- ☐Test determinism across independent implementations and malformed cases.

Primary rejection risk: Ontology design without sufficient environmental-software novelty.

One Earth or another broad climate journal

- ☐Demonstrate validity conflation across multiple environmental markets and material reporting consequences.
- ☐Establish international relevance and public-interest significance independent of AgEvidence.

Primary rejection risk: Specialized infrastructure problem without broad climate-system consequence.

Exit gate: Completed corpus, balanced panel, at least twelve worked cases, agreement analysis, interface experiment, and versioned revised state model. **Owner:** Persona C; methods support from independent governance researcher; implementation by Persona G.

A.5.3 Paper 3: Minimum Evidence Requirements

Status: Review and consensus framework; not an official requirement or completed Delphi.

animal

- ☐Complete systematic or scoping review with transparent appraisal.
- ☐Separate efficacy, effectiveness, safety, production response, persistence, and implementation fidelity.
- ☐Stratify by dairy, feedlot beef, grazing beef, and other justified systems.
- ☐Conduct formal consensus and test requirements against published studies.

Primary rejection risk: An MRV checklist rather than an animal-bioscience contribution.

Journal of Dairy Science

- ☐Produce a dairy-specific manuscript with dairy management, ration, lactation, production, and measurement detail.
- ☐Address animal/pen/herd clustering and repeated measures.
- ☐Test against actual dairy intervention studies.

Primary rejection risk: Cross-sector framework insufficiently grounded in dairy biology and production.

Animal Feed Science and Technology

- ☐Narrow to feed additives; include composition, stability, delivery, dose verification, intake, mixing, persistence, safety, and production response.
- ☐Ground evidence fields in nutritional mechanisms and field execution.

Primary rejection risk: Architecture and governance overshadow feed science.

Livestock Science / Translational Animal Science

- ☐Emphasize commercial feasibility, producer burden, regional differences, and translation stages.
- ☐Include producer/adviser participants and real cases.

Primary rejection risk: Normative framework without field validation.

Exit gate: Preregistered review and consensus methods; completed panel rounds; dissent; evidence map; context matrix; independent case testing. **Owner:** Persona D, with Persona A and Persona B; schema implementation by Persona G.

A.5.4 Paper 4: From Trial to Farm Gate

Status: Applied-methods case; requires rights-cleared real or defensibly mirrored data.

Agricultural Systems

- ☐Use at least two contrasting commercial deployment cases.

- ☐Map actors, incentives, biological efficacy, operational fidelity, verification, and co-claim decisions.
- ☐Quantify missing evidence, duplicate capture, reconstruction effort, exception frequency, and farm cost.

Primary rejection risk: Descriptive single case with weak generalizability.

Computers and Electronics in Agriculture

- ☐Automate ingestion, graph assembly, validation, and error recovery.
- ☐Benchmark mapping accuracy, completeness, performance, and independent user reconstruction.

Primary rejection risk: Application of Paper 1 without new computing contribution.

Animal Feed Science and Technology / Animal Production Science

- ☐Include intervention identity, dose integrity, delivery variance, intake, diet, animal response, and field-domain assessment.
- ☐Use real intervention data and subordinate claims architecture to feed or production science.

Primary rejection risk: No biological or feeding result.

Journal of Cleaner Production

- ☐Define system boundary, baseline, lifecycle implications, allocation, leakage, uncertainty, and operational tradeoffs.

Primary rejection risk: Evidence quality is incorrectly treated as environmental performance.

Exit gate: Data rights; complete graph; independent reconstruction; field implementation metrics; exception and co-claim scenarios. **Owner:** Persona B; implementation realism by Persona F; measurement review by Persona D.

A.5.5 Paper 5: Whole-Farm Grazing Systems

Status: Field-systems framework; requires producer participation and longitudinal evidence.

Agricultural Systems

- ☐Recruit multiple grazing enterprises across climates and management systems.
- ☐Integrate land, herd, weather, pasture, supplementation, production, and intervention timelines.
- ☐Quantify observed/proxy/modelled evidence, missingness, burden, and decision consequences.

Primary rejection risk: Static ontology without multi-farm empirical work.

Rangeland Ecology & Management / Agriculture, Ecosystems & Environment

- ☐Add ecological measurements, spatial/temporal sampling, pasture and land-condition outcomes, and model validation.
- ☐Demonstrate ecological findings enabled by the architecture.

Primary rejection risk: Information-system framing without agroecological contribution.

Livestock Science / Animal Production Science

- ☐Include herd performance, management, feasibility, producer labor, and farm-type heterogeneity.

Primary rejection risk: Institutional MRV emphasis overwhelms livestock-production science.

Journal of Environmental Management

- ☐Define environmental-management decisions and measure how evidence integration changes them.

Primary rejection risk: Record integration without improved environmental assessment.

Exit gate: Multi-farm participation; longitudinal graph; burden study; temporal alignment; producer validation; failure cases. **Owner:** Persona F; Australian applicability by Persona A; data integration by Persona G.

A.5.6 Paper 6: Manure, Waste, and Farm Energy

Status: Engineering extension; strongest path is a real operating case with closed balances.

For a bounded interval:

$$M_{\text{in}} = M_{\text{out}} + M_{\text{loss}} + \Delta M_{\text{storage}} + \epsilon_M,$$

$$E_{\text{gross}} = E_{\text{exported}} + E_{\text{parasitic}} + E_{\text{flared}} + E_{\text{loss}} + \Delta E_{\text{storage}} + \epsilon_E.$$

Journal of Environmental Management

- ☐Complete balances, uncertainty, sensor provenance, exception analysis, and decision comparison.
- ☐Demonstrate what an external reviewer can reconstruct.

Primary rejection risk: Insufficient engineering data or excessive platform emphasis.

Waste Management

- ☐Characterize waste, treatment boundaries, losses, storage, digestate, contamination, failure, and resource recovery.
- ☐Compare alternative management pathways.

Primary rejection risk: Information architecture without waste-management science.

Renewable Energy

- ☐Report gross/net energy, parasitic loads, flaring, downtime, efficiency, temporal performance, and economics.

Primary rejection risk: Attribute accounting without substantive energy-system results.

Bioresource Technology

- ☐Include feedstock properties, loading, retention, temperature, gas composition, methane yield, digestate, replicate or longitudinal data, and engineering novelty.

Primary rejection risk: Provenance architecture without new bioprocess findings.

Energy Policy / Journal of Cleaner Production

- ☐Compare policy profiles for methane, energy, nutrients, carbon, additionality, ownership, and double counting.
- ☐Quantify incentive and allocation consequences.

Primary rejection risk: Single technical case without broader policy or cleaner-production inference.

Exit gate: Representative case; rights-cleared data; closed mass/energy balances; uncertainty; external engineering reconstruction; policy-profile analysis. **Owner:** Persona E; calculation notebooks by reproducibility engineer and Persona G.

A.5.7 Paper 7: Portable Evidence, Local Authority

Status: Capstone portability study; foreign adapters must remain research artifacts until locally reviewed.

Let a frozen graph G have root digest d_G, and adapters A_j^v produce determinations D_j^v and reliance artifacts R_j^v:

$$(G, d_G, A_j^v) \longrightarrow (D_j^v, R_j^v).$$

The proposed evidence-reuse metric is

$$\rho_j = \frac{|F_{\text{reused},j}|}{|F_{\text{required},j}|}.$$

Additional metrics should include semantic reuse, transformation burden, local-judgment count, unresolved requirements, implementation effort, and reproducibility.

Environmental Science & Policy

- ☐Select real comparative contexts and recruit local scientific, policy, legal, and assurance experts.
- ☐Distinguish scientific applicability, administrative eligibility, verification, rights, and claims.
- ☐Report interpretive disagreement, failure, and risks of regulatory or policy laundering.

Primary rejection risk: Synthetic mappings are too artificial for policy conclusions.

Environmental Modelling & Software

- ☐Formalize portability metrics; publish adapters; test determinism, supersession, and independent implementation.
- ☐Build a benchmark corpus and report irreducible local judgment.

Primary rejection risk: Policy mapping without environmental-software contribution.

Computers and Electronics in Agriculture

- ☐Compare shared-core adapters with duplicated country schemas using code reuse, data reuse, mapping effort, errors, and maintenance burden.

Primary rejection risk: Policy comparison overshadows agricultural computing.

Agricultural Systems

- ☐Treat each jurisdiction as an agricultural and institutional system; include stakeholders, capacity, burden, and distributional effects.

Primary rejection risk: Technical conformance experiment lacks systems grounding.

npj Sustainable Agriculture / Nature Food stretch route

- ☐Expand to multiple countries, real partners, real production systems, quantified non-portability costs, and global food-system consequences.
- ☐Establish independent replication and broad relevance beyond the platform.

Primary rejection risk: Architecture benchmark does not rise to a globally consequential food-systems finding.

Exit gate: Frozen real evidence root; at least two locally reviewed comparative adapters; formal metrics; supersession; independent replication; false-compliance audit. **Owner:** Persona A and Persona B; implementation by Persona G; local experts and legal reviewer mandatory.

A.6 Journal Decision Table

Paper	Journal	Scientific object	Non-negotiable hardening	Best evidence design	Desk-rejection trigger
1	Computers and Electronics in Agriculture	Implemented agricultural software architecture	Reference implementation, baselines, benchmarks, ablation, usability, open artifacts	Australian evidence graph plus independent reconstruction	Conceptual or generic cryptography
1	Environmental Modelling & Software	Formal evidence-model software	Executable semantics, model provenance, supersession, reproducibility	Frozen graph under multiple model/adapter versions	No modelling/software advance
1	Agricultural Systems	Multi-actor system architecture	Stakeholder flows, burden, incentives, production-context comparison	Mixed methods plus system metrics	Software-only evaluation
2	Environmental Science & Policy	Governance of validity and assurance	Policy corpus, balanced panel, agreement, legitimacy and contestability	Content analysis plus elicitation	Terminology essay
2	Journal of Environmental Management	Decision-control interface	Controlled user experiment and decision outcomes	Randomized or counterbalanced interface study	No management consequence
3	animal	Minimum scientific evidence	Systematic review, consensus, context stratification, case tests	Review plus Delphi/consensus	MRV checklist
3	Journal of Dairy Science	Dairy evidence requirements	Dairy-specific biology, ration, production, measurement, statistics	Dairy review plus applied cases	Insufficient dairy science
3	Animal Feed Science and Technology	Feed-additive translation	Product, dose, stability, intake, diet, safety, persistence	Feed-additive review plus field validation	Architecture dominates feed science
4	Agricultural Systems	Trial-to-commercial value-chain case	Multiple cases, actors, burden, exceptions, co-claims	Comparative applied case study	Single descriptive case
4	Computers and Electronics in Agriculture	Evidence-graph implementation	Ingestion, graph assembly, validation, baseline comparison	Field data plus software benchmark	No computing novelty
5	Agricultural Systems	Whole-farm grazing evidence	Multi-farm longitudinal integration and burden	Mixed-method longitudinal field study	Static ontology
5	Rangeland Ecology & Management	Rangeland monitoring/evidence	Ecological measurements and spatial-temporal validation	Multi-site rangeland study	No ecological contribution
6	Journal of Environmental Management	Reconstructable environmental engineering evidence	Closed balances, provenance, exceptions, external reconstruction	Operating digester/waste-energy case	Thin data and platform emphasis
6	Waste Management	Waste-treatment evidence	Waste characterization, process performance, losses, digestate	Longitudinal process dataset	No waste science
6	Renewable Energy	Farm energy performance	Net energy, parasitics, downtime, efficiency, economics	Operational energy dataset	Attribute accounting only
7	Environmental Science & Policy	Evidence portability with local authority	Real contexts, local experts, comparative governance, failure	Comparative policy case study	Synthetic-only adapters
7	Environmental Modelling & Software	Adapter conformance benchmark	Formal metrics, deterministic execution, independent implementation	Frozen graph plus multiple adapters	Policy mapping only

A.7 Portfolio Governance and RACI

Scientific authors retain authority over research questions, data interpretation, limitations, and final text. AgEvidence may fund the technical secretariat, data engineering, reproducibility, and journal administration, but commercial approval cannot substitute for author judgment.

Role	P1	P2	P3	P4	P5	P6	P7
Persona A - Australia Scientific Chair	A/L	C	C	R	R	C	A/Co-L
Persona B - Global Scientific Co-Chair	C	C	R	A/L	R	C	A/Co-L
Persona C - institutional and dairy translation	C	A/L	C	R	I	R	R
Persona D - measurement and uncertainty	C	R	A/L	R	R	I	R
Persona E - manure, waste, and energy	C	I	I	I	I	A/L	R
Persona F - grazing and producer burden	C	I	R	R	A/L	I	R
Persona G - platform engineering	R	R	R	R	R	R	R
Persona H - executive sponsor	I	I	I	I	I	I	I
Publication Program Manager	R	R	R	R	R	R	R
Scientific technical editor	R	R	R	R	R	R	R
Data Steward	R	R	R	R	R	R	R
Reproducibility engineer	R	R	R	R	R	R	R
Independent methods reviewers	C	C	C	C	C	C	C
Legal/conflicts reviewer	C	C	C	C	C	C	C

Legend: A = accountable; L = proposed lead; R = responsible; C = consulted; I = informed. Authorship still requires CRediT-qualifying contribution and written approval.

A.7.1 Operating roles that must be appointed

- ☐**Publication Program Manager:** integrated schedule, issue log, meetings, submissions, archive, and dependency control.
- ☐**Scientific technical editor:** terminology, figures, cross-paper claims, reference quality, and consistency without replacing author judgment.
- ☐**Data Steward:** permissions, de-identification, provenance, dataset versions, access, and release packages.
- ☐**Reproducibility engineer:** notebooks, fixtures, checksums, environments, verifier scripts, and archival bundles.
- ☐**Independent social/governance methods specialist:** Paper 2 corpus, elicitation, agreement, and interface-study design.
- ☐**Independent systematic-review specialist:** Paper 3 protocol, screening, appraisal, and consensus governance.
- ☐**Australian producer advisory participants:** Paper 5 workflow and burden validity.
- ☐**External engineering reviewer:** Paper 6 reconstruction and balance audit.
- ☐**Local-country reviewers:** Paper 7 scientific, policy, assurance, and legal interpretation.

A.8 Twenty-Four-Month Work Program

Wave	Timing	Outputs	Exit gate
0	Months 0–2	Author consent; CRediT plans; publication charter; conflict and rights review; data inventory; terminology freeze; journal reconnaissance; protocols for Papers 1–3.	Authors, roles, data route, and preregistration route approved.
1	Months 2–7	Papers 1 and 2; architecture specification; reference implementation; conformance alpha; policy/interface corpus.	Independent technical reproduction and governance pilot passed.
2	Months 5–14	Paper 3 review and consensus; Papers 4 and 5 data collection and applied graphs.	Minimum evidence framework completed; two applied workflows tested.
3	Months 9–17	Paper 6 engineering case, balances, exception testing, external reconstruction.	Cross-intervention reuse and physical reconstruction demonstrated.
4	Months 14–22	Paper 7 frozen root, comparative adapters, supersession, conformance handbook.	Same root evaluated under multiple locally reviewed research adapters.
Translation	Months 20–24	Australian workshop, technical brief, developer guide, repository releases, correction policy, next-country decision.	External review, citation, reuse, test, or documented criticism.

A.9 Master TODO Register

A.9.1 Governance and authorship

- ☐Obtain written author interest, affiliation approvals, expected contributions, and conflict disclosures.
- ☐Approve publication charter covering independence, unfavorable findings, preprints, IP, data rights, press review, correction, and withdrawal.
- ☐Establish paper-specific CRediT records and authorship-change procedure.
- ☐Define whether AgEvidence may review for confidentiality without controlling scientific conclusions.
- ☐Commission at least two independent reviewers per paper before journal submission.

A.9.2 Architecture and software

- ☐Freeze version 0.1 terminology, envelope, receipt families, states, adapter manifest, and bundle format.
- ☐Publish formal non-assertions and prohibited inference rules.
- ☐Build canonical examples, invalid examples, incomplete examples, corrections, disputes, and supersession cases.
- ☐Build offline verifier and independent implementation.
- ☐Archive release tags, dependency lockfiles, checksums, and executable notebooks.

A.9.3 Data and ethics

- ☐Select synthetic mirror, permissioned pilot subset, or post-pilot route for each paper.
- ☐Document consent, rights, confidentiality, de-identification, and access restrictions.
- ☐Obtain human-research or ethics determinations for expert elicitation, producer interviews, interface tests, and workflow observation.
- ☐Establish adverse-event and product-confidentiality handling for intervention studies.

A.9.4 Literature and claims

- ☐Replace vendor and consultancy summaries with primary or authoritative sources wherever possible.
- ☐Verify every numeric efficacy range and mechanistic statement against extracted primary literature.
- ☐Verify the proposed NDF-effect statement and units; remove it if the source and model cannot be reproduced.
- ☐Distinguish current standards from revisions under development, consultation, or transition.
- ☐Create a claim-to-source matrix and prohibit unsupported carryover between papers.

A.10 Open Science and Commercial Boundary

Open scientific infrastructure	Commercial managed layer
Vocabulary, receipt envelope, payload and adapter specifications.	Hosted infrastructure, high availability, monitoring, and SLAs.
Tripartite validity states and prohibited inference rules.	Enterprise integrations, access controls, workflow operations, and support.
Open schemas, conformance fixtures, synthetic cases, and benchmark graphs.	Maintained policy mapping, customer-specific adapters, migrations, and evidence operations.
Reference verifier, CLI, tests, and reproducibility scripts.	Verification-readiness packet assembly, buyer portals, reporting integrations, and dispute handling.
Published methods, limitations, failure results, and correction history.	Confidential customer data, contracts, commercial allocations, and service commitments.

An open specification should use a clear license, governance process, versioning policy, change proposal procedure, conformance definition, trademark boundary, security disclosure process, and archival commitment. “Open” should not be used if implementation requires undisclosed policy logic or proprietary test fixtures.

A.11 Pre-Submission Gate Checklist

A manuscript may enter journal formatting only after all relevant boxes are complete.

- ☐Research question and target journal were selected before final analysis.

- ☐Article type is permitted by the journal and confirmed against current author guidance.
- ☐The abstract states what was actually completed, not what is planned.
- ☐All proposed and observed results are clearly separated.
- ☐Methods are reproducible and reporting guidance is followed.
- ☐Data and code availability statements are accurate.
- ☐Conflicts, funding, sponsor role, and author contributions are complete.
- ☐Standards and policy versions are dated and cited to authoritative sources.
- ☐No statement implies Australian Standard status, regulator acceptance, VVB approval, legal advice, or universal claim ownership.
- ☐Independent reviewer comments and author responses are archived.
- ☐Figures, tables, code, supplements, and manuscript use the same vocabulary and version identifiers.
- ☐Journal-specific novelty and rejection-risk paragraph has been answered in the cover letter.

A.12 Directed Reference-Building Plan

The following list is a starting architecture for the bibliography, not a claim that every item has been fully reviewed.

A.12.1 Authoritative standards and policy sources

- Greenhouse Gas Protocol, *Corporate Value Chain (Scope 3) Accounting and Reporting Standard*, current standard and official revision-process materials.
- GHG Protocol, *Scope 3 Standard Revisions Phase 1 Progress Update*, March 2026.
- Science Based Targets initiative, current FLAG guidance, criteria, and official update materials.
- ISO 14064-1, ISO 14064-2, ISO 14064-3, ISO 14065, and ISO 14066, using official texts or authorized access.
- IPCC Guidelines and refinements applicable to agriculture, forestry, land use, livestock, manure, and uncertainty.
- Australian government, Clean Energy Regulator, DCCEEW, and method-owner documentation relevant to the selected case.

A.12.2 Protocol and interoperability theory

Review the end-to-end argument, hourglass or thin-waist architecture, robustness and extensibility, protocol ossification, semantic interoperability, event sourcing, append-only logs, formal conformance testing, and policy-as-code. The final review should explain both the advantages and failure modes of thin waists: loss of expressiveness, premature standardization, security monoculture, semantic leakage, and governance capture.

A.12.3 Provenance and research reproducibility

Review W3C PROV, verifiable credentials where relevant, research objects, FAIR principles, data lineage, model cards, datasheets, software citation, reproducible computational environments, cryptographic transparency systems, and correction/supersession patterns. Avoid assuming that blockchain is necessary; compare ordinary signed append-only structures and conventional provenance databases.

A.12.4 Enteric methane sources

Anchor the review in peer-reviewed primary trials and high-quality systematic reviews/meta-analyses for 3-NOP, Asparagopsis, nitrates, and other intervention classes. Extract study design, animal context, diet, dose, duration, measurement method, outcomes, safety, conflicts, and production applicability. Include Australian mitigation reviews and applied grazing evidence.

A.12.5 Manure and farm-energy sources

Review manure-management emissions, anaerobic digestion process engineering, mass/energy balance methods, sensor and laboratory quality, methane destruction, leakage, flare performance, digestate, nutrient accounting, lifecycle boundaries, renewable-energy attributes, and allocation overlap.

A.12.6 Journal author guidance

Before each submission, archive the dated journal aims, article types, author instructions, data/code policy, ethics policy, AI-use policy, preprint policy, graphical requirements, word limits, and open-access terms. Journal rules may change and should not be inferred from old templates.

A.13 Selected Working References and Source Leads

1. Greenhouse Gas Protocol. *Corporate Value Chain (Scope 3) Standard*. Official standard page: `https://ghgprotocol.org/corporate-value-chain-scope-3-standard`.
2. Greenhouse Gas Protocol. *Scope 3 Standard Revisions Phase 1 Progress Update*. March 31, 2026. `https://ghgprotocol.org/sites/default/files/2026-03/S3-Phase1ProgressUpdate-20260331.pdf`.
3. Dijkstra, J. et al. Meta-analytic literature on 3-nitrooxypropanol effects and moderators. Begin with the open-access meta-analysis indexed at `https://pmc.ncbi.nlm.nih.gov/articles/PMC9868067/`; verify bibliographic details in the final library.
4. Roque, B. M. and collaborators. Grazing-beef Asparagopsis study. *Proceedings of the National Academy of Sciences*, DOI landing page: `https://www.pnas.org/doi/10.1073/pnas.2410863121`.
5. Hristov, A. N. and collaborators. Primary 3-NOP dairy and cattle trials; build a complete trial table rather than relying on secondary efficacy summaries.
6. Australian enteric methane mitigation review literature, including the open-access review "Methane Emissions from Ruminants in Australia: Mitigation Potential and Applicability of Mitigation Strategies."
7. Global Research Alliance. *An evaluation of evidence for efficacy and applicability of methane inhibiting feed additives for livestock*. Use as an evidence map and trace central claims to primary literature.
8. Computers and Electronics in Agriculture. Current aims and author guidance: `https://www.sciencedirect.com/journal/computers-and-electronics-in-agriculture`.
9. Journal of Dairy Science. Current aims and contributor instructions: `https://www.journalofdairyscience.org/`.
10. Smart Agricultural Technology. Current aims: `https://www.sciencedirect.com/journal/smart-agricultural-technology`.

A.14 Closing Research Position

The strongest defensible contribution is not that one platform has solved agricultural MRV. It is that the program proposes a falsifiable separation between stable evidence, versioned scientific interpretation, formal assurance, and bounded institutional reliance. The publication sequence should expose that proposition to independent implementation, empirical failure, domain-specific criticism, and local-authority review.

The practical publication order is:

1. harden Paper 1 into a complete implementation and conformance study;
2. run Paper 2's policy corpus, expert classification, and interface work concurrently;
3. preregister and complete Paper 3's review and consensus process;
4. use rights-cleared applied cases for Papers 4 and 6;
5. begin Paper 5 only after producer participation and longitudinal data are secured; and
6. submit Paper 7 only after the vocabulary, evidence root, metrics, and local-review procedure are stable.

Premature submission would likely convert solvable methodological gaps into avoidable desk rejections. The research program should first produce the open specification, review protocols, conformance fixtures, permissions, independent reproduction packages, and completed empirical results identified in this document.