Chapter 3: The SSOT Architecture

Building the Single Source of Truth system that makes SUI verification possible.

- What is a Single Source of Truth (SSOT)?
- The Three-Tier Validation Pipeline
- The Digital Twin for Impact Verification

What is a Single Source of Truth (SSOT)?

What is a Single Source of Truth (SSOT)?

> **Definition:** A Single Source of Truth (SSOT) is a system of record in which every piece of data relevant to an enterprise's impact claims exists in exactly one canonical location — structured, timestamped, access-controlled, and audit-ready — such that any authorised party can independently verify the enterprise's SUI claims without relying on summaries prepared by the enterprise itself.

Why "Single Source"?

Most early-stage companies manage their data across a fragmented set of tools: spreadsheets emailed between team members, production records in an ERP system, customer data in a CRM, lab results in PDFs stored in Dropbox, and impact metrics calculated in a separate Excel model. Each of these is a source of data — but none is authoritative. When an investor asks "show me how you calculated 102.4 kg CO?e per hectare," the answer cannot be found in any single place.

An SSOT eliminates this fragmentation. It does not necessarily mean one database — it means one *canonical layer* through which all relevant data flows, where every claim is traceable to its source, and where the chain of custody is documented.

What the SSOT Must Contain

For SUI verification purposes, the SSOT must hold:

1. **Input records:** Evidence of each product application event (batch production records, delivery confirmations, IoT sensor readings, GPS coordinates of deployment)
2. **Baseline data:** The counterfactual reference data, with source documentation and version history
3. **Calculation engine outputs:** The intermediate steps in converting input records to SUI magnitudes (the Digital Twin layer)
4. **Outcome data:** Third-party measurement results (lab analyses, satellite observations, auditor field reports)

5. **Aggregated SUI ledger:** A time-series of verified SUI events, each linked back to its source input record and outcome measurement
6. **Version history:** All changes to the above, with timestamps and the identity of who made each change

SSOT vs. Data Warehouse vs. ERP

System Type	Purpose	Audit-Ready?	Role in SUI Pipeline
ERP (SAP, Odoo, QuickBooks)	Business operations records	Partially	Source of input data (production, sales)
Data Warehouse (Snowflake, BigQuery)	Analytics and reporting	No (mutable)	Intermediate processing layer
BI Tool (Tableau, Metabase)	Visualisation	No	Output layer for stakeholder dashboards
SSOT (SUI Architecture)	**Impact claim verification**	**Yes (immutable audit trail)**	**Canonical record of all SUI events**

The Four Properties of a SUI-Grade SSOT

Property 1: Immutability

Once a SUI event is recorded and verified, it cannot be modified without creating a new, linked record that documents the correction. This is achieved through append-only data structures, cryptographic hashing of records, or blockchain anchoring (for the highest assurance levels). The principle: you can correct errors, but the original record and the correction are both permanently visible.

Property 2: Traceability

Every SUI magnitude in the impact ledger must be traceable back to its source input records. A verifier must be able to ask: "Show me the raw data behind SUI event #4,721" and receive a complete chain: batch record ? production quantity ? application event ? calculation log ? outcome measurement ? verified SUI value.

Property 3: Access Control with Audit Logging

The SSOT must implement role-based access control: company staff can write new records; investors can read aggregated data; independent verifiers can access underlying records during audit windows. Every access event is logged — who accessed what, when, and what they downloaded.

Property 4: Structured for External Consumption

The SSOT must be able to produce a standardised data export in a format specified by the relevant verification standard (e.g., ISAE 3000, ISO 14064-3, or the auditor's own format). This is not a spreadsheet dump — it is a structured, machine-readable dataset that maps to the SUI parameter specification.

SSOT Maturity Levels

Not every startup needs a full enterprise SSOT from day one. CTH recognises four maturity levels:

Level	Name	Description	Typical Stage
0	Fragmented	Data in multiple disconnected tools; no single version of truth	Pre-seed, < 12 months
1	Consolidated	Data centralised in one tool (even a well-structured spreadsheet); no automation	Seed, pilot phase
2	Automated	Data flows automatically from source systems to a central repository; version control in place	Series A, growth phase
3	Audit-Ready	Full immutability, access control, traceability, and structured export; third-party verified	Series B+, pre-MDB engagement

A startup at SSOT Level 1 can still define and communicate a SUI — the specification document is the foundation. Verification becomes possible at Level 2 and full financial instrument eligibility at Level 3.

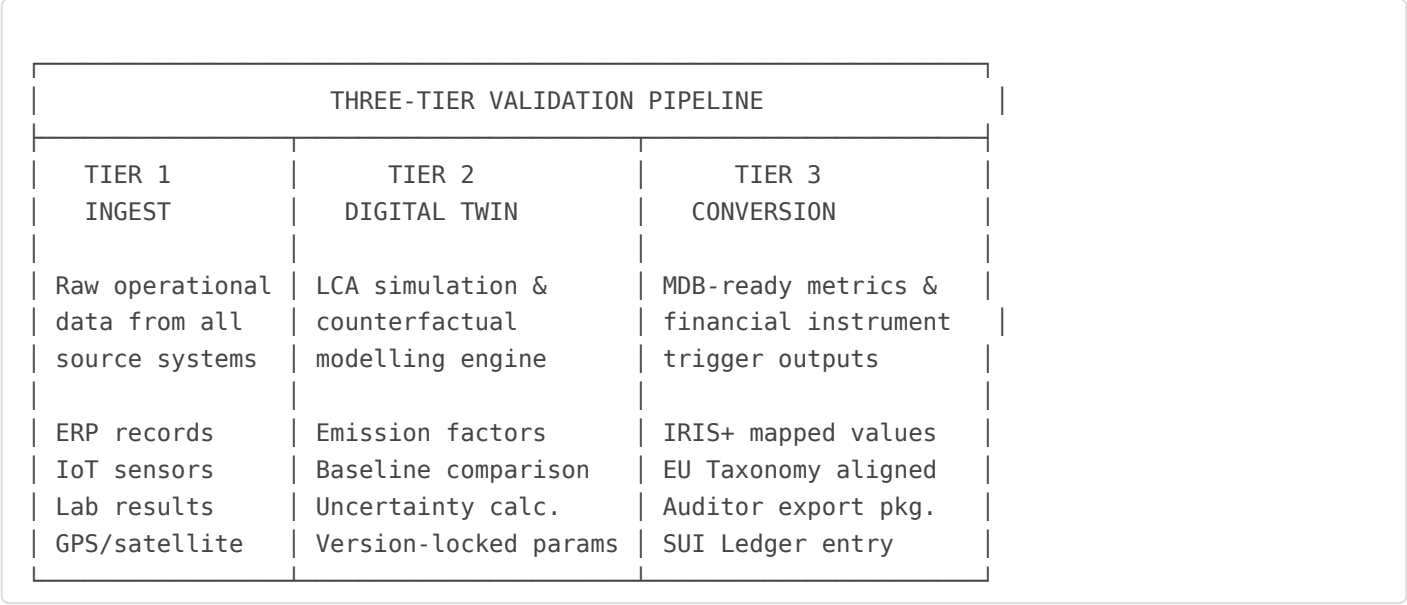
Next: The Three-Tier Validation Pipeline — how data flows from raw inputs to verified SUI events.

The Three-Tier Validation Pipeline

The Three-Tier Validation Pipeline

The Three-Tier Validation Pipeline is the data architecture that converts raw operational records into verified SUI events. It is named for its three sequential layers: Ingest, Digital Twin, and Conversion. Each tier has a defined responsibility, a defined data format, and a defined handoff protocol to the next tier.

Pipeline Overview



Tier 1: Ingest

Purpose

Capture all raw evidence of product application events in structured, timestamped form. The Ingest tier is the intake valve of the SSOT — every piece of data relevant to a SUI claim must enter through it.

Data Sources by Sector

Sector	Typical Ingest Sources
AgTech / Bio-inputs	Batch production records (ERP), field application GPS logs, customer delivery confirmations, soil lab analysis PDFs
Clean Energy / EV	IoT charging session data (kWh, duration, vehicle ID), grid connection records, utility meter readings
Water Treatment	Flow meter readings, water quality sensors (turbidity, pH, pathogen count), treatment plant operational logs
Circular Economy	Material inflow/outflow manifests, weight measurements, recycler receipts, chain-of-custody certificates
Built Environment	BMS (Building Management System) data, energy audit reports, occupancy sensors, utility bills

Ingest Requirements

- **Timestamp:** Every record carries a machine-generated UTC timestamp, not a manually entered date
- **Source ID:** Every record carries the identifier of the system or device that generated it
- **Immutability flag:** Once ingested, records are locked; corrections create new records with a "supersedes" link to the original
- **Schema validation:** Records are validated against a defined schema on ingest; malformed records are quarantined, not silently dropped

Tier 2: Digital Twin

Purpose

Apply the SUI calculation logic to the ingested records — comparing observed outcomes to the counterfactual baseline, applying emission factors, calculating uncertainty, and producing a per-application SUI magnitude.

What "Digital Twin" Means Here

In this context, "Digital Twin" refers to a computational model of the enterprise's impact mechanism — not a real-time operational simulation. The Digital Twin encodes:

- The Life Cycle Assessment (LCA) model for the product's impact pathway
- The baseline values (counterfactual scenario) and their uncertainty ranges
- The emission factors or conversion coefficients (e.g., IPCC AR6 values for N?O emission from synthetic nitrogen)
- The aggregation rules (how individual application events are summed to period totals)

Version Control for Model Parameters

Every change to the Digital Twin model — a new emission factor, an updated baseline, a revised LCA boundary — must be version-controlled. Each SUI event in the ledger is tagged with the model version that produced it. This allows historical SUI calculations to be reproduced exactly, even after model updates.

Digital Twin Outputs

For each ingested application event, the Digital Twin produces:

- SUI magnitude (central estimate)
- Uncertainty range (±N%, at 95% confidence)
- Model version tag
- Calculation audit log (step-by-step computation)
- Data quality flag (complete data vs. estimated vs. proxy)

Tier 3: Conversion

Purpose

Transform the Digital Twin outputs into the formats required by different stakeholders — investors, MDBs, auditors, regulators, and the SUI Ledger itself.

Output Formats

Output	Format	Audience
SUI Ledger Entry	Structured JSON record in the SSOT	SSOT system, auditors
IRIS+ Report	Indicator values mapped to IRIS+ codes	Impact investors, GIIN reporting
EU Taxonomy Contribution Statement	% revenue / capex / opex aligned	European institutional investors
MDB Project Brief	AIMM-compatible impact narrative + data table	IFC, IDB Invest, ADB co-investors
Auditor Export Package	CSV + methodology PDF + raw data links	Third-party verifiers (ISAE 3000)
Investor Dashboard	Aggregated charts + drill-down to unit level	Board, VCs, DFI monitors

Conversion Layer Controls

The Conversion tier must enforce several data integrity controls:

- **No manual overrides:** Conversion outputs are computed, not manually adjusted. Any "rounding" or "presentation formatting" must be documented and must not change material values.
- **Uncertainty propagation:** Uncertainty from the Digital Twin is carried through to all Conversion outputs, not dropped at the reporting layer.
- **Audit trail linkage:** Every Conversion output includes a reference back to the Tier 2 records that produced it, and through those, to the Tier 1 raw inputs.

The Pipeline in Practice: Becaps Example

1. **Ingest:** Becaps ERP records batch #BC-2024-0441: 500 kg product shipped to cooperative Finca Verde, La Calera, Colombia. GPS delivery confirmed. Cooperative confirms application to 500 ha (1 kg/ha). Soil lab reports uploaded: pre/post nitrogen content for 12 sample plots.
2. **Digital Twin:** Model applies: Baseline = 220 kg N/ha (DANE 2023 Colombian synthetic fertiliser use). Observed = 85 kg N/ha (lab-confirmed). Net displacement = 135 kg N/ha. IPCC AR6 conversion: 135 kg N × 0.758 CO?e/kg N = 102.3 kg CO?e/ha. Uncertainty: ±12.3 kg CO?e (±12%). Model version: DT-Becaps-v2.1.
3. **Conversion:** 500 SUI events recorded in ledger (one per hectare). IRIS+ PI5765 report: 51,150 kg CO ?e avoided this batch. Auditor export package generated. MDB project brief updated with cumulative totals.

Next: The Digital Twin for Impact Verification — building and validating the computational model.

The Digital Twin for Impact Verification

The Digital Twin for Impact Verification

The Digital Twin is the computational core of the SUI verification system. It is a version-controlled, auditable model that takes raw operational data as input and produces per-application SUI magnitudes as output. This page describes how to build, validate, and maintain a Digital Twin suitable for third-party impact verification.

What Goes Into the Digital Twin Model

1. The Impact Pathway Model

The impact pathway describes the causal chain from product application to environmental outcome. It answers the question: "Through what mechanism does one application of our product produce the claimed SUI magnitude?"

For a biostimulant company like Becaps:

```
Product application (1 kg biostimulant / ha)
    → Microbial inoculation (≥10¹¹ CFU/g colonise root zone)
    → Biological nitrogen fixation (BNF increases N availability)
    → Reduced synthetic N application requirement (−135 kg N/ha)
    → Avoided N₂O emissions from synthetic fertiliser production (×0.758 CO₂e/kg N)
    → Avoided N₂O emissions from soil application of synthetic N (×0.01 kg N₂O/kg N)
    → Total GHG displacement: 102.4 kg CO₂e/ha
```

Each arrow in this chain must be supported by either (a) peer-reviewed scientific literature, (b) field trial data from the company's own operations, or (c) certified emission factors from recognised sources (IPCC, EPA, DEFRA).

2. Emission Factors and Conversion Coefficients

The Digital Twin relies on external data — emission factors, conversion ratios, global warming potential values — that are updated periodically by standard-setting bodies. The model must:

- Specify the exact source and version for every external coefficient (e.g., "IPCC AR6 WGI Table 7.SM.7, GWP100 value for N?O: 273 CO?e")
- Implement version locking — historical SUI events use the factor version in effect at the time of calculation
- Trigger recalculation reviews when major updates are released (e.g., when IPCC publishes a new Assessment Report)

3. Baseline Model

The counterfactual baseline is a model in its own right. It specifies:

- The reference activity that would occur without the enterprise's product (e.g., "conventional synthetic nitrogen application at regional average rates")
- The data source and geographic scope of the baseline (e.g., DANE 2023 Colombia agricultural census)
- The temporal validity of the baseline (baselines degrade as markets change — a 2019 baseline for synthetic fertiliser use in a region that has since adopted sustainable agriculture practices will overstate impact)
- Baseline update triggers: conditions under which the baseline must be recalculated (e.g., if market penetration of competing biostimulants exceeds 20% in the region)

4. Uncertainty Model

Impact uncertainty has multiple sources, each of which must be quantified:

- **Measurement uncertainty:** Variability in field measurements (soil nitrogen content measured in 12 plots out of 500 ha — sampling error)
- **Model uncertainty:** Uncertainty in the emission factor values (IPCC provides ranges, not point estimates)
- **Baseline uncertainty:** The baseline is an average — individual farms may deviate significantly from the average
- **Attribution uncertainty:** The portion of observed N reduction attributable specifically to the biostimulant (vs. weather, management changes, etc.)

The Digital Twin propagates these uncertainties using Monte Carlo simulation or analytical uncertainty propagation and reports the combined 95% confidence interval on every SUI magnitude output.

Building the Digital Twin: Minimum Viable Version

For a pre-Series A startup, the minimum viable Digital Twin can be a well-structured, version-controlled Excel or Python model. The key requirements are:

1. **Documented inputs:** Every input cell or variable has a source citation
2. **Auditable calculations:** No black boxes — every formula is visible and reviewable

3. **Version control:** The model is stored in a git repository or equivalent with change history
4. **Reproducibility:** Given the same inputs, a second analyst running the model independently produces the same outputs within rounding error
5. **Sensitivity analysis:** The model includes a sensitivity analysis showing which inputs have the largest impact on the SUI magnitude

Digital Twin Validation

Before the Digital Twin is used for investor reporting or financial instrument design, it must be validated by an independent third party. Validation means:

1. **Model review:** The validator reviews the impact pathway logic, the source citations for all emission factors, and the baseline methodology
2. **Calculation audit:** The validator independently recalculates a sample of SUI events using the model and confirms they match the company's reported values
3. **Field data audit:** The validator reviews the raw field data (lab reports, IoT records) and confirms they match what was ingested into the model
4. **Written validation statement:** The validator issues a written statement (following ISAE 3000 or equivalent) confirming the model is fit for purpose

Scaling the Digital Twin

As the company scales, the Digital Twin must evolve from a spreadsheet model to a software system. Key milestones:

Scale Milestone	Digital Twin Requirement
< 1,000 SUI events/year	Excel/Python model with manual data input, annual validation
1,000–50,000 events/year	Automated data pipeline from SSOT to model; semi-annual validation
50,000–1M events/year	Real-time or near-real-time computation; continuous monitoring; annual third-party audit
> 1M events/year	Enterprise-grade system with SOC 2 Type II certification; continuous auditing by major accounting firm

---

Continue to Chapter 4: Financial Mechanisms — how a verified SUI translates into reduced cost of capital.