



The Blast Radius Problem: Genealogy Closure, Containment Queries, and Why a Lot Number Is Not Traceability

Architecture, Computational Methods, and Reported Outcomes from MileSoft Material Traceability

MileSoft Engineering Research Group
MileSoft Software Technologies

<https://milessoft.net/>

Corresponding authors: *Deepak Daryani* (deepakdaryani@milessoft.net), *Tanuj Daryani* (tanujdaryani@milessoft.net)

Working paper — Version 1.0 — 23 August 2026

Permanent link: <https://milessoft.net/research/products/material-traceability>

Abstract

A recall is a set-membership problem before it is anything else. The customer names a defect; the supplier must return the set of units that could carry it. Every hour spent computing that set is spent shipping and installing more of them, and every unit wrongly included in it is contained, recovered and replaced at full cost for nothing.

This paper presents the architecture and computational methods of MileSoft Material Traceability, which records receipt of raw lots, the sub-assembly events that consume them, final assembly into a serialised unit, and shipment to a customer, as a single directed acyclic graph. Containment is then a forward-closure computation on that graph and per-unit audit evidence is a backward closure — two traversals of one structure rather than two reconstructions from unlinked records.

We state the closure computations, the blast-radius ratio that quantifies what genealogy is worth in a given recall, a commingling model for the resolution loss that occurs when lots are mixed at a bulk feed, and a completeness model showing that a genealogy degrades by collapsing to its weakest captured tier rather than proportionally. The work is anchored in ISO/IEC 19987 and 19988 for the event and vocabulary model, ISO 22005 for traceability system design, IATF 16949 and AS9100D for the automotive and aerospace evidence obligations, 21 CFR Part 11 for audit-trail integrity, EU Regulation 2017/745 for unique device identification, and the FDA Food Traceability Rule for critical tracking events.

The product's published record reports containment queries returning affected units with customer and ship date in under a minute. A four-module deployment at a Tier-1 automotive supplier, in which this module was one of four, reported IATF 16949 audit preparation falling from three weeks to one day; that figure belongs to the programme and is not attributed to this module alone.

Keywords: Material traceability software; Genealogy tracking; Recall containment query; Forward and backward traceability; Lot genealogy; Barcode and RFID capture; EPCIS event model; IATF 16949 traceability; AS9100D counterfeit prevention; 21 CFR Part 11 audit trail; EU MDR unique device identification; FSMA critical tracking events

1 Introduction

The hardest call in manufacturing begins with a customer saying they found a defect and need to know which other units are affected. Everything that follows — the cost, the exposure, the relationship — is decided by how large the answer is and how long it takes to produce.

Both are properties of a data structure that had to exist before the call. If receipt, consumption, assembly and shipment were recorded as linked events, the answer is a graph traversal returning a specific set of units with their customers and ship dates. If they were recorded as four unrelated logs, no traversal exists, and the only defensible answer is the whole exposure window.

The product's own material states this contrast as the difference between containing 200 units and containing 20,000. That is an illustrative figure rather than a measured incident, and this paper treats it as such — but the ratio it points at is real, computable, and the subject of Section 4.2.

1.1 Why a lot number is not traceability

Almost every manufacturer records lot numbers. Very few can compute a containment set. The gap between those two facts is the subject of this paper.

A lot number on a receiving document is a label on a node. Traceability requires edges: this lot was consumed by that sub-assembly, that sub-assembly went into this unit, this unit shipped to that customer on that date. A node without its edges cannot be traversed, and edges cannot be inferred afterwards from timestamps and part numbers without assuming the very consumption order that was never recorded.

The practical test is not whether lot numbers are stored. It is whether, given one lot number, the system returns finished units rather than a list of documents mentioning that lot.

1.2 Contributions

1. A four-tier genealogy model stated as a directed acyclic graph, with containment and audit evidence as its forward and backward closures (Sections 3 and 4.1).
2. A blast-radius ratio that makes the value of genealogy computable for a specific recall rather than asserted in general (Section 4.2).
3. A commingling model for resolution loss at bulk feeds, which bounds achievable traceability precision independently of software (Section 4.3).

4. A completeness model showing that one uncaptured tier collapses the containment set to that tier's whole population (Section 4.4).
5. An eight-dimension capability reference framework, each dimension demonstrable on a running line (Section 6.3).

2 Background and Related Work

Traceability is one of the few manufacturing capabilities with a settled international vocabulary, a settled event model, and a set of sector regulations that all describe the same structure in different words. The engineering problem is not defining traceability; it is capturing enough of it.

2.1 What the standards actually require

ISO 22005 defines traceability as the ability to follow the movement of a material through specified stages of production, processing and distribution, and states the design requirements a traceability system must satisfy. It was written for feed and food, but its formulation is domain-neutral: specified stages, documented links, defined start and end points.

The event model that carries this is EPCIS, standardised as ISO/IEC 19987, which records each event along four dimensions — what object, when, where, and under which business step. Its companion vocabulary, ISO/IEC 19988, fixes the permitted values for business step and disposition, so that packing at one site and packing at another are the same recorded fact rather than two local conventions.

Sector regulation adds obligations rather than a different structure. IATF 16949 requires evidence of product traceability as part of evidence of process control. AS9100D adds configuration management and explicit counterfeit-part prevention, which is a traceability requirement in all but name. The Food Traceability Rule at 21 CFR Part 1 Subpart S names its links critical tracking events and its fields key data elements, and requires them to be produced as a sortable electronic record within twenty-four hours of a request. EU Regulation 2017/745 requires a unique device identifier under Article 27 and traceability obligations on economic operators under Article 25.

Four regulators, four vocabularies, one structure: identified objects, captured links between them, and a bounded time in which the graph must be queryable.

2.2 Record integrity as a separate obligation

A traceability graph that can be edited after the fact answers questions but proves nothing. 21 CFR Part 11 states the obligation directly: audit trails must be secure, computer-generated and time-stamped, must independently record the operator entries and actions that create, modify or delete a record, and must not obscure previously recorded information.

This is stricter than it first appears. It rules out in-place correction of a genealogy edge: a mis-scanned lot is corrected by recording a superseding event, never by overwriting the original. The graph is therefore append-only with respect to history, and any query can be asked as of a date.

2.3 Identification keys and where they come from

Genealogy edges connect identifiers, so the identifier scheme determines what can be connected. GS1 provides the Global Trade Item Number for a trade item class, the Serial Shipping Container Code for a logistic unit, and label placement rules governing where those symbols may be applied; ISO/IEC 15459-1 provides unique identification of transport units where a GS1 key is not in use.

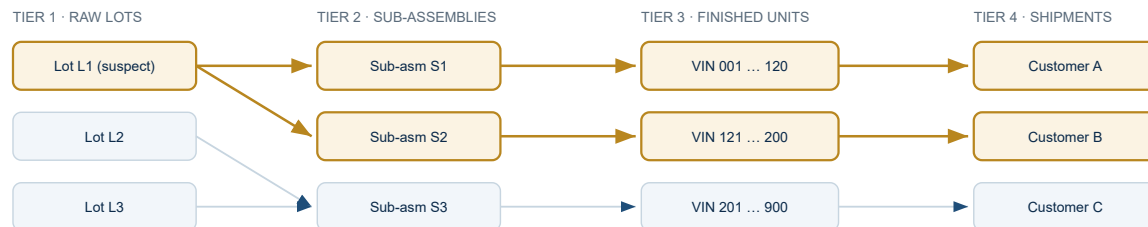
The distinction that matters for genealogy is class against instance. A trade item number identifies what something is; a serial number or vehicle identification number identifies which one. A genealogy built on class identifiers alone cannot separate two units of the same part number, which is exactly the separation a containment set requires.

3 System Overview

The system is one graph with four kinds of node and three kinds of edge, fed by hardware-agnostic capture and queried in two directions. Everything else — dispositions, reports, regulatory profiles — is a projection of that graph.

THE GENEALOGY GRAPH, AND THE TWO CLOSURES A CONTAINMENT QUERY NEEDS

A directed acyclic graph over four tiers. Every edge is one captured event; an edge that was never captured is a query that silently returns the wrong answer.



THE TWO QUERIES

Forward closure $D(L1)$ — every descendant of the suspect lot. Highlighted above: 200 units across two customers. This is the containment set.

Backward closure $A(v)$ — every ancestor of one unit. Answers what went into this vehicle, which is the audit and warranty question.

Both are reachability computations on the same graph. Neither is reconstructable afterwards from records that were never linked at capture.

Without the graph there is no closure to compute, so containment falls back to the only defensible envelope: everything shipped in the exposure window.

In this diagram that is all 900 units and all three customers — 4.5 times the population that is actually affected.

Figure 1. Schematic. Schematic (not product UI): the four-tier genealogy graph, with the forward closure of a suspect raw lot highlighted. The containment set is 200 units and two customers; without the graph, the defensible envelope is all 900 units and three customers.

3.1 The four tiers

Table 1. The four node tiers, what identifies each, and the event that creates it.

Tier	Identity	Creating event
Raw lot	Supplier lot number plus receipt reference	Receipt: supplier, lot, certificate, timestamp
Sub-assembly	Serial or internal batch identifier	Consumption: which lots were drawn, at which station, by which operator
Finished unit	Serial number or vehicle identification number	Final assembly: which sub-assemblies were fitted, with quality data attached
Shipment	Logistic unit code and delivery reference	Despatch: which finished serials, to which customer, on which date

Edges run only downstream in time, so the structure is acyclic by construction and both closures terminate. Rework is not a cycle: a reworked unit records a new event referencing the prior state, which keeps the ordering intact and satisfies the append-only requirement of Section 2.2.

3.2 Capture, and why it is hardware-agnostic

Handheld barcode scanners, fixed RFID readers and RFID gate readers all feed the same spine. The system is deliberately indifferent to which produced a given edge, because the two technologies are good at different things and a real plant uses both.

- RFID is used for bulk pallet movements and high-value sub-assemblies, where reads happen without line of sight and without a person stopping to aim.
- Barcode is used for serialised components and operator-confirmed steps, where the act of scanning is itself the confirmation that a specified step was performed.

The engineering consequence is that the capture layer must normalise two quite different failure modes. A barcode read fails loudly — nothing is recorded and the operator sees it. An RFID gate read fails quietly, by reading fourteen of fifteen tags on a pallet. The second is the dangerous one, because it produces a genealogy that looks complete and is not, which is precisely the condition Section 4.4 models.

3.3 The two queries and the disposition layer

Forward containment takes a lot, part number or date range and returns affected finished units with customer and ship date. Backward genealogy takes a unit and returns every lot that entered it, with the quality data captured along the way — torque events, test results, inspection sign-offs where those systems are integrated.

On top of both sits lot-level disposition: hold, non-conformance report, rework, release. A disposition is itself an event on the graph rather than a flag on a row, which is what makes the state of a lot answerable as of any past date — the question an auditor actually asks.

What-if recall modelling runs the forward closure against a hypothetical defect signature before any customer call. Its purpose is not the answer but the exposure: a plant that has never run one does not know its own blast radius.

4 Computational Methods

Four computations do the work. The first two are what the system runs; the third and fourth are what a plant should run on itself before it trusts the first two.

4.1 Containment and audit evidence as graph closures

Let $G = (V, E)$ be the genealogy graph, with V partitioned into the four tiers of Section 3.1 and each edge (u, v) recording that u was consumed by, fitted into, or despatched as v .

$D(u) = \text{every node reachable from } u \text{ (the forward closure)}, A(v) = \text{every node that reaches } v \text{ (the backward closure)}$ (closure)

$u \leadsto v$ denotes the existence of a directed path from u to v . The containment set for a suspect lot L is $D(L)$ restricted to the finished-unit and shipment tiers; the audit genealogy for a unit is $A(\text{unit})$ restricted to the raw-lot tier.

Both are single-source reachability computations, linear in the edges actually touched rather than in the size of the graph. This is why the published sub-minute containment figure is architecturally credible: the traversal visits the affected subgraph, not the year of production it sits in.

The cost of a containment query is a property of the answer, not of the database. A recall touching 200 units costs the same to compute whether the plant has built 200,000 units or 20 million.

4.2 The blast-radius ratio

What genealogy is worth in a given recall is the difference between the set that must be contained with it and the set that must be contained without it. Both are computable, which turns a marketing contrast into an engineering quantity.

$\text{beta} = (\text{units shipped in the exposure window}) / (\text{units in the forward closure of the suspect lot}), \text{cost avoided} = \text{unit containment cost} \times \text{the difference between them}$ (blast)

$N(\text{window})$ is every unit shipped between the earliest and latest possible use of the suspect lot — the envelope a plant must contain when it cannot compute $D(L)$. $c(\text{unit})$ is the fully loaded cost of containing, recovering and replacing one unit.

Beta is a property of the plant, not of the incident: it rises with production rate, with lot size relative to consumption rate, and with the delay between consumption and the defect being noticed. A plant can compute its own beta today from its shipping records and its bill of materials, without deploying anything.

The 200-against-20,000 contrast published for this product is a beta of 100. Appendix B works a smaller and more typical case and still reaches a beta of 4.5.

4.3 Commingling, and the resolution a genealogy can actually reach

Traceability precision is bounded before any software is chosen, by the points at which lots are physically mixed. A hopper fed from two lots produces parts that belong to both; a paint batch, a bulk fastener bin and a resin silo all behave the same way.

ρ = units implicated per unit of suspect lot consumed, and the implicated set is at least the sum of the mixing-vessel populations the lot passed through (commingle)

$q(L)$ is the quantity of the suspect lot consumed, K the set of mixing points it passed through, and $m(k)$ the number of units drawn from vessel k between the lot entering it and the vessel being emptied. $\rho = 1$ is perfect resolution; $\rho > 1$ is the dilution the physical process imposes.

The engineering consequence is that improving traceability is often a process change rather than a software change: emptying a hopper between lots, reducing vessel size, or recording a purge event all reduce ρ directly. A system that captures mixing events at least makes ρ visible; one that does not lets a plant believe its resolution is higher than the physics allows.

A supplier claiming unit-level traceability through a bulk feed is claiming something the process does not support. The honest claim is lot-to-vessel-population traceability, and the number that matters is ρ .

4.4 Why an incomplete genealogy fails discontinuously

It is tempting to treat genealogy completeness as a percentage, on the assumption that a 90% complete graph gives a 90% good answer. It does not, and the reason is structural.

$\gamma(t)$ = captured edges at tier t / true edges at tier t , and the computed containment set equals the true one only if every tier is complete - otherwise it is the whole population of the first broken tier (completeness)

t^* is the earliest tier at which $\gamma < 1$ along the suspect path. Below that tier the traversal has no edges to follow, so every node in it is a candidate and must be contained.

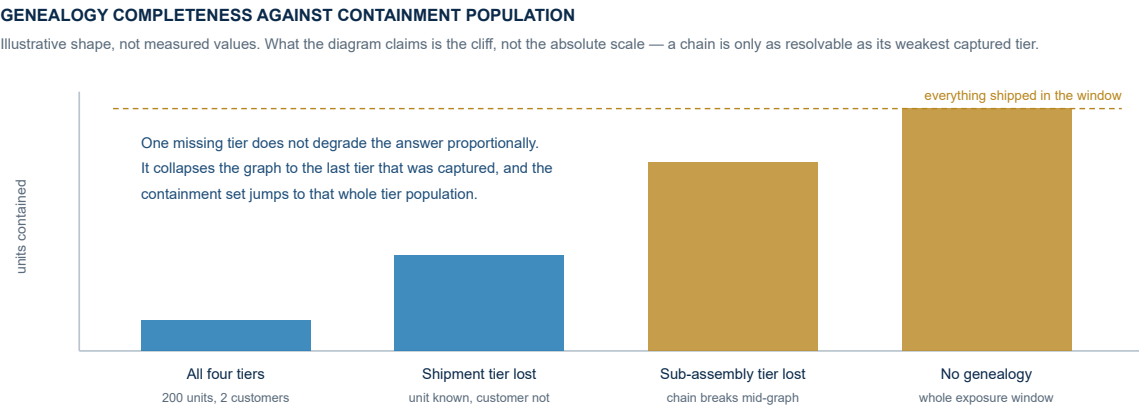


Figure 2. *Schematic.* Schematic (not measured data): containment population against which tiers were captured. The claim is the discontinuity, not the absolute scale — losing one tier does not cost a proportional amount of precision, it collapses the answer to that tier's whole population.

This is why the quiet RFID failure of Section 3.2 matters more than its read rate suggests. A gate reading fourteen of fifteen tags has $\gamma = 0.93$ at that tier, and any recall whose path crosses the missing edge is answered at the resolution of the tier above.

5 Reported Outcomes and Field Evidence

This section separates three kinds of claim, because they carry different weight and a reader is entitled to know which is which.

5.1 Figures published for this module

Table 2. Figures published for this module, with their provenance.

Figure	Value	Provenance
Containment query time	under one minute	Product material — vendor-reported, no sample or distribution given
Containment population contrast	200 units against 20,000	Product material — explicitly illustrative, not a measured incident
Regulatory profiles supported	IATF 16949, AS9100, 21 CFR Part 11, EU MDR, FSMA	Product material — configurable per domain and contract

The first is a system property that Section 4.1 makes architecturally credible — a closure traversal is bounded by the answer, not by the archive. The second is presented in its own source as a contrast rather than a measurement, and this paper does not upgrade it.

5.2 A deployment in which this module was one of four

A Tier-1 automotive supplier deployed four MileSoft modules together across two plants — Torque Traceability, Andon, this module, and ERP. This module supplied genealogy from raw lot through sub-assembly to finished vehicle identification number, closing the loop for IATF 16949 audits.

The programme reported IATF 16949 audit preparation falling from three weeks to one day, with audit evidence — torque history, process control charts, traceability links — available directly from the system rather than manually consolidated.

Shared attribution. That result belongs to the four-module programme. Torque history and process control charts came from a different module, and the individual contribution of material genealogy was not isolated. It is reported here as programme evidence, not as this module's measured effect.

What can be said specifically is narrower and more useful: the audit-preparation obligation this module addresses is the traceability link evidence, and that evidence is a projection of the graph in Section 3 rather than a document assembled for the auditor.

5.3 Modelled scrap and recall avoidance

The published return model for this module is stated as a scrap and recall reduction, and every assumption in it is printed here so a reader can substitute their own.

Table 3. Modelled scenario — not a deployment result. Assumptions are the published defaults of the return-on-investment model for this product.

Assumption	Value
Units produced per month	50,000 (600,000 per year)
Baseline scrap and recall rate	1% (6,000 events per year)
Reduction attributed to the module	50%
Modelled events avoided per year	3,000
Cost per scrap or recall event	site-specific; the published default is used only to illustrate

The 50% figure is an assumption, not a measurement, and it is the one a buyer should interrogate hardest. Genealogy does not prevent a defective lot from being consumed. What it changes is how much of the consequence is contained, and Section 4.2 gives the honest form of that argument: the saving is $c(\text{unit})$ times the difference between the window and the closure, not a percentage of the defect rate.

This paper reports the published model faithfully and disagrees with its framing. Compute beta from your own shipping records and bill of materials; it is a better estimate than any percentage a vendor can supply.

6 Discussion

6.1 The binding constraint is capture, not query

Every traceability system on the market can traverse a graph. The differences that decide whether a plant can answer a recall are all upstream of the query, in whether the edges exist.

This reframes the buying decision. The question is not how fast the containment query runs but how many events per shift the floor must contribute, how many of those are automatic, and what happens to the graph when one is missed. A system requiring twelve manual scans per unit will have gaps; one deriving ten of them from equipment will not.

It also reframes the deployment. The expensive part is not configuration — it is instrumenting the consumption points, which is a floor-layout and process question that software cannot answer on its own.

6.2 One graph, five regulatory profiles

The product supports traceability requirements across five regimes. It is worth being precise about what that means, because the sector standards differ in obligation while agreeing almost entirely in structure.

Table 4. What each regime adds on top of the common graph. The structure is shared; the obligations are not.

Regime	What it adds	Where it lands in the model
IATF 16949	Traceability as evidence of process control	Backward closure with quality data attached at each tier
AS9100D	Configuration management, counterfeit-part prevention	Supplier certificate held on the raw-lot node and traversed with it
21 CFR Part 11	Secure, computer-generated, tamper-evident audit trail	Append-only event history; superseding events rather than edits
EU 2017/745	Unique device identifier, economic-operator traceability	Instance-level identity at the finished-unit tier; shipment tier extended past the first consignee
21 CFR Part 1, Subpart S	Critical tracking events, key data elements, 24-hour sortable record	The tier events themselves, with a mandated field set and a response-time obligation

The practical implication is that a manufacturer serving two regulated sectors does not need two traceability systems. It needs one graph, two field profiles and two report formats — which is the whole argument for building on a standard event model rather than a sector-specific one.

6.3 A capability reference framework for traceability systems

Table 5. Capability reference framework. Each dimension separates a traceability system from a lot-number field, and each is answerable by demonstration on a running line rather than by a specification sheet.

Dimension	Question the system must answer by demonstration
D1 Edges, not labels	Given one raw lot number, does the system return finished units — or documents mentioning that lot?
D2 Both closures	Can it answer forward containment and backward genealogy from the same structure?
D3 Instance identity	Does the finished tier carry a serial or vehicle identification number, or only a part number?
D4 Append-only history	Correct a mis-scanned lot. Is the original still visible, superseded rather than overwritten?
D5 Commingling visible	Does a mixing point record an event, so that rho is computable rather than assumed to be 1?
D6 Gap detection	Break one edge deliberately. Does the system report a gap, or silently return a smaller answer?
D7 As-of queries	Ask for the disposition of a lot as it stood ninety days ago. Does it answer?
D8 Time to answer	Run a containment query on live production data, not a demonstration dataset. How long?

D6 is the dimension buyers almost never test and the one that decides whether a recall answer can be trusted. A system that returns a confident wrong set is more dangerous than one that returns nothing.

6.4 Generalisability

The graph model generalises wherever material is transformed and identity is preserved: automotive, aerospace, medical devices, food, pharmaceuticals. The commingling model of Section 4.3 generalises to any process with a mixing vessel and is the limiting factor in exactly those industries where regulation is strictest.

What does not generalise is the blast-radius ratio's magnitude. Beta depends on production rate, lot size and detection delay, so a figure quoted from one plant says nothing about another. It should be computed, not borrowed — including from this paper.

7 Threats to Validity and Limitations

1. The 200-against-20,000 contrast is illustrative. It is published as an illustration of the difference between surgical and catastrophic containment, not as a measured incident, and no sample or site is attached to it.

2. The sub-minute containment figure is vendor-reported. Section 4.1 argues it is architecturally credible, but architectural credibility is not measurement; no dataset size, graph depth or hardware is stated.
3. The audit-preparation result is confounded. Three weeks to one day belongs to a four-module programme, and the largest component of the evidence involved — torque history and process control charts — came from a different module.
4. The 50% scrap and recall reduction is a model assumption. Genealogy contains consequences rather than preventing defects, so this figure is doing work the mechanism does not obviously support; Section 5.3 gives the form of the argument that does.
5. Traceability precision is bounded by process, not software. Wherever lots commingle, ρ exceeds 1 and no system can recover unit-level resolution that the physical process destroyed.
6. Supplier-side genealogy depends on the supplier. Walking backwards past the facility requires the supplier's certificate and lot structure to be in the system; where it is not, the graph terminates at receipt.
7. Quiet capture failures are not fully detectable. A gate reading fourteen of fifteen tags produces a plausible graph. Reconciliation against expected counts limits this, but a system cannot in general prove that an edge which was never captured should have existed.
8. No independent audit of regulatory conformance. Support for five regulatory profiles is a vendor statement; conformance in a given deployment is established by that deployment's audit, not by this paper.

The seventh limitation is the one that should shape a buyer's evaluation. Ask what the system does when an edge is missing, and prefer the answer that reports a gap over the answer that returns a smaller set.

8 Future Work

- Publishing beta distributions rather than a single contrast. A blast-radius ratio measured across a range of real recalls would replace the illustrative 100 with something a buyer could reason about.
- Gap detection as a reported metric. Continuous reconciliation of captured edges against expected counts would let γ be published per tier per shift, turning Section 4.4 from a warning into an instrument.
- Commingling instrumentation. Recording vessel purge and changeover as first-class events would make ρ measurable rather than estimated, and would identify which process changes buy the most resolution.
- Supplier graph federation. Extending the backward closure past receipt requires a shared event model rather than a shared database; ISO/IEC 19987 and 19988 already provide the vocabulary for it.
- Isolating this module's contribution in multi-module programmes, so audit-preparation improvement can be attributed rather than reported as a shared result.

9 Conclusion

A recall is a set-membership problem, and the set is either computable or it is not. What decides that is whether receipt, consumption, assembly and shipment were captured as edges of one graph while production was running, because none of those links can be reconstructed afterwards from records that were never joined.

Three results in this paper are the ones worth carrying away. Containment and audit evidence are the two closures of a single structure, so the cost of an answer scales with the answer rather than the archive. The value of that structure in a given recall is a computable ratio, beta, which a plant can estimate today from records it already holds. And genealogy does not degrade gracefully: one uncaptured tier collapses the answer to that tier's whole population, which is why a system's behaviour on a missing edge matters more than its behaviour on a complete one.

The framework of Section 6.3 is offered as the durable contribution. Its sixth dimension is the one to test first, and testing it takes an afternoon: break an edge on purpose, and see whether the system tells you.

Appendix A Nomenclature

Table 6. Symbols and abbreviations used in this paper.

Symbol / term	Meaning
$G = (V, E)$	The genealogy graph: nodes are lots, sub-assemblies, units and shipments; edges are captured consumption or despatch events
$u \leadsto v$	A directed path exists from u to v
$D(u)$	Forward closure — every node reachable from u ; the containment set
$A(v)$	Backward closure — every node that reaches v ; the audit genealogy
$N(\text{window})$	Units shipped within the exposure window, the fallback containment envelope
beta	Blast-radius ratio: $ N(\text{window}) $ divided by the true affected population
$c(\text{unit})$	Fully loaded cost of containing, recovering and replacing one unit
rho	Commingling resolution: units implicated per unit of suspect lot consumed
$m(k)$	Units drawn from mixing vessel k while the suspect lot was in it
$\gamma(t)$	Capture completeness at tier t : captured edges over true edges
t^*	The earliest tier along the suspect path at which γ is below 1
CTE / KDE	Critical tracking event and key data element, the Food Traceability Rule's terms for a link and its fields
UDI	Unique device identifier, required under EU Regulation 2017/745 Article 27
SSCC / GTIN	Serial Shipping Container Code and Global Trade Item Number

Appendix B Worked Numerical Examples

Appendix B.1 Computing a plant's own blast-radius ratio

A plant builds 900 units in the exposure window. A supplier reports a defect in raw lot L1, delivered on day 3 and consumed over days 4 to 9. The genealogy graph shows L1 was drawn into two sub-assemblies, which were fitted to 120 and 80 finished units respectively.

Applying Equation (closure), the containment set is 200 units, resolving to two customers. Applying Equation (blast), $\beta = 900 / 200 = 4.5$.

At a fully loaded containment cost of 180 currency units per vehicle, the avoided cost is $180 \times (900 - 200) = 126,000$ for this single incident. Note what the calculation does not require: no assumption about defect rates, no percentage reduction, no vendor figure. It uses the plant's own shipping records and its own bill of

materials.

This is the calculation a buyer should run before any procurement conversation. If beta comes out near 1, this class of system is not the plant's largest opportunity — and knowing that is worth the afternoon it takes.

Appendix B.2 What a hopper does to resolution

The same lot L1 is a fastener stock feeding a bulk hopper. 400 fasteners from L1 are tipped into a hopper already holding 600 from lot L0, and the hopper is drawn down over the next 1,000 units without being emptied.

Applying Equation (commingle), $q(L1) = 400$ but the implicated set is the whole vessel population $m(k) = 1,000$, so $\rho = 1,000 / 400 = 2.5$. No software recovers the missing resolution: after mixing, no record distinguishes which unit received an L1 fastener and which an L0.

Two process changes each cut this directly. Emptying the hopper between lots gives $m(k) = 400$ and $\rho = 1.0$. Halving the hopper gives $m(k) = 500$ and $\rho = 1.25$. Both are cheaper than any software, and neither is visible to a plant whose system does not record mixing events.

Appendix B.3 The cost of one missing tier

Take the graph of the first example and remove the sub-assembly tier — the plant records lot receipts and finished units, but never captured which sub-assemblies consumed which lots.

Applying Equation (completeness), gamma at the sub-assembly tier is 0, so t^* is that tier and the computed containment set is its entire population. Every unit built from any sub-assembly of that type in the window is a candidate: 900 rather than 200.

The graph was 75% complete by tier count — three of four tiers captured — and delivered none of the precision. Beta returns to 4.5 despite the substantial capture investment already made, which is the practical meaning of the discontinuity in Figure (blast): partial genealogy buys partial confidence, not partial containment.

Data provenance

Every quantitative claim in this paper resolves to one of the entries below. Figures marked as modelled estimates are not deployment measurements; their assumptions are stated.

Metric	Reported value	Provenance
Containment query time	under one minute	Product specification /products/material-traceability
Containment population with complete versus incomplete	200 units versus 20,000	Product specification /products/material-traceability — Illustrative contrast published in

Metric	Reported value	Provenance
genealogy		the Material Traceability blog post, not a measured incident.
Scrap and recall events avoided per year	3,000 of 6,000 (modelled)	Modelled estimate 50,000 units per month, 600,000 per year; Baseline scrap and recall rate of 1%; 50% reduction attributed to the module; Model and defaults published in src/data/roiModels.ts
IATF 16949 audit preparation time	1 day (from 3 weeks)	Operator-reported — Tier-1 Manufacturers (Pune + Nasik plants, India) /case-studies/industry40
Phased rollout duration	8 weeks (plant 1), 6 weeks (plant 2)	Operator-reported — Tier-1 Manufacturers (Pune + Nasik plants, India) /case-studies/industry40

References

- [1] International Organization for Standardization / International Electrotechnical Commission (2024). ISO/IEC 19987: Information technology - EPC Information Services (EPCIS) Standard, version 2.0. ISO/IEC, Geneva. <https://www.iso.org/standard/85557.html>
- [2] International Organization for Standardization / International Electrotechnical Commission (2024). ISO/IEC 19988:2024 - Information technology - GS1 Core Business Vocabulary (CBV). ISO/IEC, Geneva, edition 3, March 2024; publishes GS1 CBV 2.0, ratified June 2022, which fixes the vocabularies for the EPCIS business step, disposition and business transaction type fields. <https://www.iso.org/standard/85558.html>
- [3] International Organization for Standardization (2007). ISO 22005:2007 - Traceability in the feed and food chain - General principles and basic requirements for system design and implementation. ISO, Geneva; defines traceability as the ability to follow the movement of a feed or food through specified stages of production, processing and distribution, and states the design requirements a traceability system must meet. <https://www.iso.org/standard/36297.html>
- [4] International Automotive Task Force (2016). Quality management system requirements for automotive production and relevant service parts organizations. IATF 16949:2016, first edition, 1 October 2016 (superseding ISO/TS 16949).
- [5] SAE International (2016). AS9100D - Quality Management Systems - Requirements for Aviation, Space, and Defense Organizations. SAE International, Warrendale, PA, September 2016; the fourth revision, which incorporates ISO 9001:2015 and adds configuration management, product safety and counterfeit-part prevention requirements. <https://saemobilus.sae.org/standards/as9100d-quality-management-systems-requirements-aviation-space-defense-organizations>
- [6] United States Food and Drug Administration (1997). 21 CFR Part 11 - Electronic Records; Electronic Signatures. FDA, Silver Spring, MD; §11.10(e) requires secure, computer-generated, time-stamped audit trails that independently record operator entries and actions creating, modifying or deleting electronic records, and that do not obscure previously recorded information. <https://www.ecfr.gov/current/title-21/chapter-I/subchapter-A/part-11>
- [7] European Parliament and Council of the European Union (2017). Regulation (EU) 2017/745 on medical devices (Medical Device Regulation). Official Journal of the European Union, 5 April 2017; Article 25 states the identification and traceability obligations on economic operators and Article 27 with Annex VI establishes the unique device identification system. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32017R0745>

- [8] United States Food and Drug Administration (2022). Requirements for Additional Traceability Records for Certain Foods - 21 CFR Part 1, Subpart S (the Food Traceability Rule, FSMA Section 204). FDA, Silver Spring, MD; final rule published November 2022 at 21 CFR 1.1300-1.1455, requiring key data elements to be recorded at defined critical tracking events and produced as a sortable electronic record within 24 hours of request. <https://www.ecfr.gov/current/title-21/chapter-I/subchapter-A/part-1/subpart-S>
- [9] GS1 (2024). GS1 General Specifications. GS1 AISBL; defines the GTIN, SSCC and GLN identification keys. <https://ref.gs1.org/standards/genspecs/>
- [10] GS1 (2024). Serial Shipping Container Code (SSCC). GS1 identification key series; an 18-digit key for a logistic unit, carried under application identifier 00. <https://www.gs1.org/standards/id-keys/sscc>
- [11] GS1 (2023). GS1 Logistic Label Guideline. GS1 AISBL; the SSCC is the only mandatory field on a GS1 logistic label. <https://www.gs1.org/standards/gs1-logistic-label-guideline/1-3>
- [12] International Organization for Standardization / International Electrotechnical Commission (2014). ISO/IEC 15459-1: Information technology - Automatic identification and data capture techniques - Unique identification - Part 1: Individual transport units. ISO/IEC, Geneva. <https://www.iso.org/standard/54779.html>
- [13] International Electrotechnical Commission / International Society of Automation (2025). Enterprise-control system integration - Part 1: Models and terminology. IEC 62264-1; ANSI/ISA-95.00.01-2025 (IEC 62264-1 Mod).
- [14] International Society of Automation (2013). Enterprise-control system integration - Part 3: Activity models of manufacturing operations management. ANSI/ISA-95.00.03-2013 (IEC 62264-3 Modified).
- [15] International Electrotechnical Commission (2025). IEC 62541-1: OPC unified architecture - Part 1: Overview and concepts. IEC, Geneva. <https://webstore.iec.ch/en/publication/81513>
- [16] International Organization for Standardization (2014). ISO 22400-1: Automation systems and integration - Key performance indicators (KPIs) for manufacturing operations management - Part 1: Overview, concepts and terminology. ISO, Geneva. <https://www.iso.org/standard/56847.html>
- [17] International Organization for Standardization (2014). ISO 22400-2: Automation systems and integration - Key performance indicators (KPIs) for manufacturing operations management - Part 2: Definitions and descriptions. ISO, Geneva; defines OEE as availability x effectiveness x quality ratio. <https://www.iso.org/standard/54497.html>
- [18] International Organization for Standardization (2022). ISO 28000: Security and resilience - Security management systems - Requirements. ISO, Geneva; published 15 March 2022. <https://www.iso.org/standard/79612.html>
- [19] Kagermann, H., Wahlster, W., and Helbig, J. (2013). Securing the future of German manufacturing industry: Recommendations for implementing the strategic initiative INDUSTRIE 4.0. Final report of the Industrie 4.0 Working Group, acatech - National Academy of Science and Engineering. https://www.acatech.de/wp-content/uploads/2018/03/Final_report__Industrie_4.0_accessible.pdf
- [20] Lee, J., Bagheri, B., and Kao, H.-A. (2015). A cyber-physical systems architecture for Industry 4.0-based manufacturing systems. *Manufacturing Letters*, 3, 18-23.