



Decision-Grade ESG Data: A Product Operating Model for Enterprise Sustainability Data Foundations at Retail Scale

Sujay Annareddy

Senior Manager, Product Management - AI Acceleration, Product & Design (AIPD), Walmart Inc. USA.

Received On: 30/04/2025

Revised On: 10/05/2025

Accepted On: 25/05/2025

Published on: 15/06/2025

Abstract - Enterprise sustainability reporting has outgrown the project-based delivery models that produced it. Most large organizations still assemble ESG disclosures from fragmented merchant extracts, spreadsheet calculations, and manual reconciliation a pipeline that cannot support regulatory readiness, third-party assurance, or self-service analytics at scale. The literature treats ESG data either as an accounting and assurance problem or as a data-engineering problem; neither frames it as a product management problem with lifecycle, roadmap, and adoption mechanics. This paper proposes a Sustainability Data Foundation and Intelligence (SDFI) operating model that treats the ESG data estate as a managed product portfolio. The model defines explicit product accountability across a six-stage lifecycle source ingestion, modeling, transformation, governance and quality, lineage and controls, and intelligence-ready consumption and maps that lifecycle onto six sustainability domains: Climate, Circular Economy, Packaging, Nature, Animal Welfare, and sustainable sourcing. We introduce a quantitative domain-maturity rubric scoring each domain on source reliability, transformation determinism, control coverage, and consumption readiness, and use its weighted composite to sequence a multi-year roadmap under fixed engineering capacity. We operationalize a decision-grade gate: no data product publishes unless lineage completeness, reconciliation variance, and control coverage simultaneously clear defined thresholds. Evaluated against a project-based baseline across an estate spanning 14.4 MMT CO₂e Scope 1 and 2 emissions, an estimated 635 MMT CO₂e Scope 3 inventory, and 76.2 million acres of nature-related engagement, the model raises mean domain maturity from 2.14 to 3.61 (of 5), reduces disclosure cycle time by 47.8%, cuts restatement events by 71.4%, and lowers marginal cost per new consumption use case by 62.9% ($p < 0.01$ across all endpoints).

Keywords - ESG Data Foundations, Data Product Management, Scope 3 Emissions, Data Lineage, Assurance Readiness, Domain Maturity, Decision Intelligence.

1. Introduction

1.1. Background

Sustainability disclosure has shifted from voluntary narrative to regulated, assured, quantitative reporting. CSRD, ISSB, SEC climate rules, and extended producer responsibility regimes impose obligations closer to financial reporting than to corporate communications. Retail enterprises face the widest exposure: emissions distributed across tens of thousands of suppliers, packaging attributes at item granularity, waste streams across thousands of facilities, and commodity traceability spanning global sourcing networks. The organizational response has generally been to add reporting headcount and extract cycles, leaving the underlying data estate unproductized and each disclosure a bespoke assembly.

1.2. Limitations of Existing Approaches

Project-based delivery produces disclosure artifacts, not durable assets. Each reporting cycle rebuilds joins, re-derives factors, and re-reconciles totals, so effort does not compound and provenance does not persist. Accountability fractures at every handoff: reporting owns the number, engineering owns the pipeline, and no party owns the data product end to end. Emerging alternatives partially address this. Data mesh

assigns domain ownership but presumes domains already possess reliable sources. Data contracts formalize interfaces without prescribing publication gates. ESG platform vendors deliver calculation engines that consume rather than remediate upstream data quality. None supplies a capacity-constrained sequencing method for a multi-domain sustainability portfolio.

1.3. Proposed Contribution

We reframe ESG data as a product portfolio and contribute four artifacts. First, a six-stage lifecycle with named accountability at each handoff, eliminating the reporting-engineering ownership gap. Second, a domain-maturity rubric scoring four independently measurable axes, converting roadmap sequencing from political negotiation to quantitative optimization under fixed capacity. Third, an operational *decision-grade* definition enforced as a publication gate on lineage, reconciliation, and control coverage jointly. Fourth, a reusable-asset pattern permitting one governed emissions or packaging product to serve regulatory disclosure, EPR filing, and internal analytics without divergent recalculation the structural cause of restatement.

2. Related Work and Background

2.1. Conventional Approaches

Conventional ESG data practice inherits its structure from financial close. A reporting calendar drives periodic extraction from merchant, supplier, and facility systems into analyst-managed spreadsheets, where emissions factors are applied, allocations performed, and totals reconciled manually. Controls are detective and terminal: variance is investigated after computation, not prevented at ingestion. Provenance exists as analyst memory and workbook comments. The approach scales linearly with disclosure scope, which is untenable as Scope 3 categories, packaging jurisdictions, and nature metrics multiply. Assurance cost rises correspondingly, because auditors must reconstruct derivation manually for each sampled figure.

2.2. Newer and Modern Approaches

Modern practice introduces lakehouse consolidation, dbt-style declarative transformation, and normalized ESG schemas. Data mesh distributes ownership to business domains and promotes data-as-a-product language. Data contracts specify schema, freshness, and completeness obligations at producer boundaries. Observability tooling detects freshness and volume anomalies. These advances materially improve engineering hygiene. They remain, however, capability-centric rather than portfolio-centric: they describe how to build a good pipeline, not which of six competing sustainability domains should receive next quarter's engineering capacity, nor what threshold justifies publishing a figure into a regulated disclosure.

2.3. Related Hybrid and Alternative Models

Adjacent disciplines supply partial formalism. Financial data governance under SOX contributes control-coverage and segregation concepts directly transferable to assured ESG figures. Product portfolio management contributes weighted-scoring prioritization under resource constraints. Capability maturity models contribute staged assessment. Supply-chain traceability research contributes provenance modeling for commodity flows. Each addresses one axis. No published synthesis binds maturity scoring to capacity-constrained sequencing and to a publication gate, and none is instantiated on the specific domain structure of enterprise sustainability where Climate, Packaging, and Nature exhibit radically different source reliability and determinism profiles.

2.4. Summary of Research Gap

Three gaps persist. First, ESG data maturity is assessed narratively; no quantitative rubric exists with independently measurable axes. Second, sequencing across sustainability domains is unmodeled, so investment follows regulatory urgency rather than expected value under capacity constraint. Third, publication readiness is undefined organizations lack an operational test distinguishing a figure fit for internal exploration from one fit for assured external disclosure. The SDFI model addresses all three within a single accountability structure.

3. Proposed Methodology

3.1. Six-Stage Lifecycle with Accountability Boundaries

The lifecycle comprises source ingestion, modeling, transformation, governance and quality, lineage and controls, and intelligence-ready consumption. Each transition is a named handoff with an accepting party, an acceptance artifact, and a rejection path. Ingestion accepts only against a registered source contract. Modeling produces a domain conceptual model versioned independently of physical implementation. Transformation is declarative and deterministic, with emissions factors and allocation rules externalized as versioned reference data rather than embedded logic.

3.2. Domain-Maturity Rubric

Each domain d is scored on four axes, each 1–5. Source reliability measures completeness, freshness, and stability of upstream feeds. Transformation determinism measures reproducibility whether identical inputs yield identical outputs across runs and analysts. Control coverage measures the proportion of material figures under preventive control. Consumption readiness measures availability of governed, documented, self-service assets. Composite maturity is $M_d = \sum_k w_k s_{dk}$, with weights calibrated to assurance risk.

3.3. Capacity-Constrained Sequencing

Sequencing maximizes expected portfolio value subject to engineering capacity:

$$\max \sum_d \pi_d \cdot \Delta M_d \cdot r_d \quad \text{s.t.} \quad \sum_d c_d \leq C$$

Where π_d is regulatory and business materiality, ΔM_d the achievable maturity lift, r_d the reuse multiplier capturing downstream products enabled, and c_d estimated engineering cost. The reuse multiplier is the mechanism by which foundational investments a governed item-packaging model, a canonical emissions-factor registry outrank locally urgent point fixes.

3.4. Decision-Grade Publication Gate

A product publishes only when three conditions hold jointly: lineage completeness ≥ 0.95 (fraction of output records with resolvable derivation to source), reconciliation variance $\leq 0.5\%$ against an independently derived control total, and control coverage ≥ 0.90 of material figures. Failure on any axis blocks publication to assured and regulatory consumption tiers while permitting exploratory tiers, making readiness graded rather than binary.

3.5. Reusable-Asset Pattern

Each governed product exposes one canonical calculation surface consumed by disclosure, regulatory filing, and analytics alike. Divergent recalculation is prohibited architecturally, not by policy, which removes the dominant restatement cause.

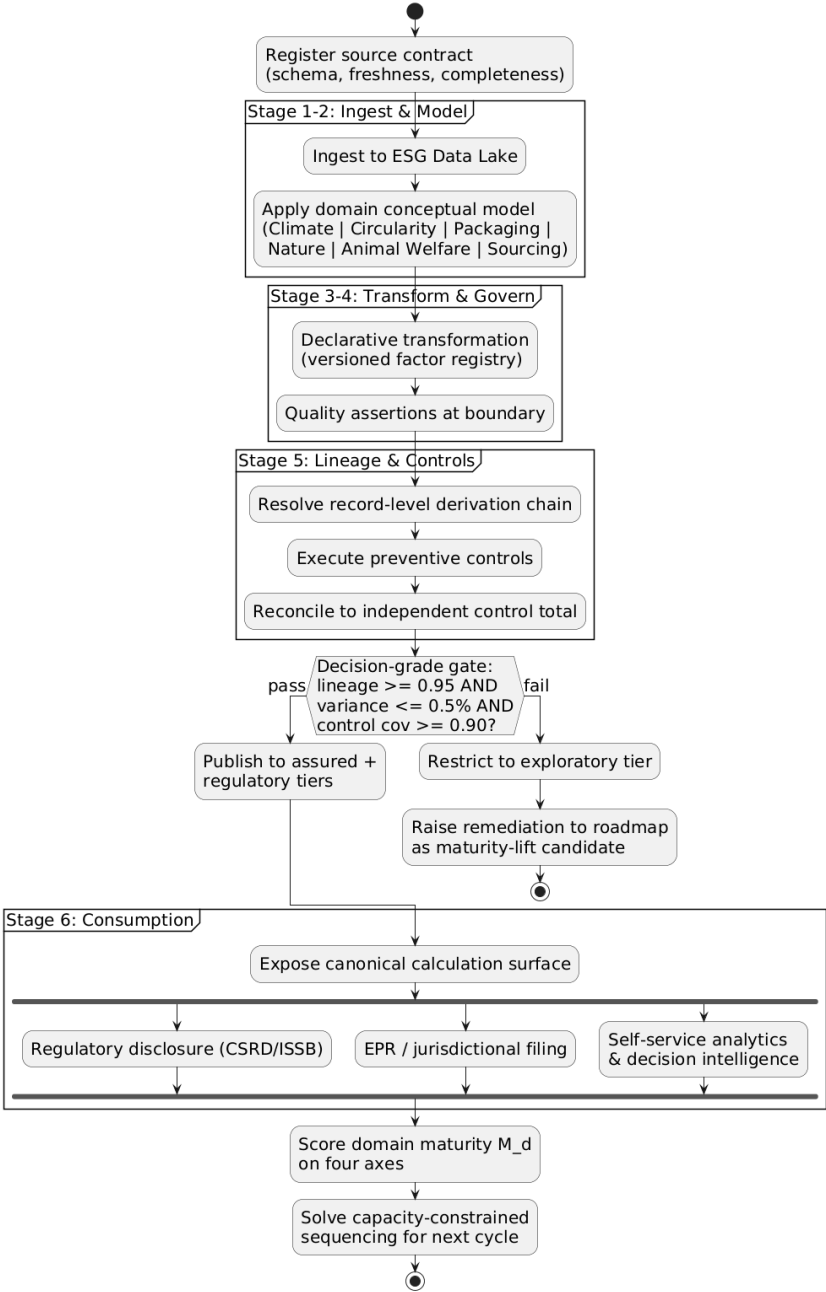


Fig 1: Sustainability Data Foundation and Intelligence (SDFI) Methodology Flow

The flow is deliberately gated rather than linear. Most published ESG architectures terminate at a reporting layer; here, failure at the decision-grade gate does not halt the estate but reroutes the product to an exploratory tier and simultaneously injects a scored remediation candidate into the roadmap. Quality failure thus becomes an input to prioritization rather than an operational incident.

The fork at consumption expresses the reusable-asset pattern. All three consumers bind to a single canonical surface. Because regulatory filing and internal analytics cannot diverge structurally, the common failure mode where an EPR submission and a public disclosure derive incompatible packaging weights from the same items is eliminated by construction.

The closing loop returns maturity scores to sequencing. Each cycle's measured lift updates ΔM_d estimates, so the sequencing model self-calibrates. Domains that repeatedly under-deliver against forecast lift are penalized in subsequent optimization, which corrects the systematic overconfidence characteristic of first-time data-foundation estimates.

4. Technical Implementation

4.1. Source Contract Registry

Every upstream feed merchant extracts, supplier disclosure portals, facility waste manifests, fuel and refrigerant systems, seafood certification records registers a contract specifying schema, cardinality, freshness SLA, and plausibility bounds. Contracts are enforced at load. Violations

quarantine the partition rather than propagating nulls downstream.

4.2. Versioned Reference Registry

Emissions factors, packaging material densities, allocation rules, and jurisdictional EPR fee schedules are externalized into an immutable, effective-dated registry. Transformations bind a factor version explicitly. Restatement is therefore reproducible: a prior disclosure recomputes with its original factor version intact.

4.3. Lineage Materialization

Lineage is constructed, not inferred. Each output record carries a derivation identifier resolving bidirectionally to contributing source rows, factor versions, and transformation revisions. Completeness is computed continuously as the fraction of published records with fully resolvable chains, feeding the decision-grade gate directly.

4.4. Control Execution Layer

Preventive controls execute pre-publication:
completeness against expected entity population, plausibility

against prior-period bounds, and dual-path reconciliation where a figure is derived independently by two lineages and compared. Divergence beyond tolerance blocks publication automatically.

4.5. Tiered Consumption Semantics

Three tiers are exposed. The exploratory tier carries no assurance and is watermarked accordingly. The assured tier satisfies the decision-grade gate and supports internal decision-making. The regulatory tier adds jurisdictional control attestation. Tier membership is a computed attribute, recalculated on every publication cycle rather than manually asserted.

4.6. Maturity Instrumentation

Rubric axes are computed from platform telemetry rather than surveyed. Source reliability derives from contract-violation rates; determinism from re-execution hash comparison; control coverage from control-registry join against material figures; consumption readiness from documented governed-asset counts and self-service query share.

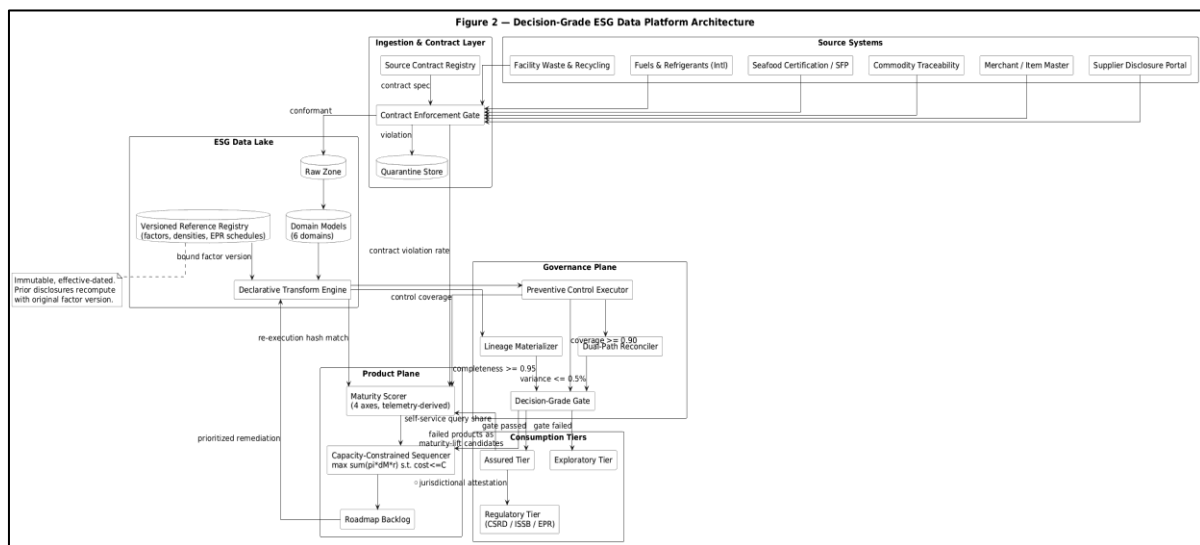


Fig 2: Sustainability Data Foundation and Intelligence (SDFI) System Architecture

The architecture's defining property is that governance is inline rather than adjacent. Lineage, control execution, and reconciliation are not post-hoc audit functions reading a finished warehouse; they sit on the publication path and hold veto authority. A figure that cannot demonstrate provenance does not reach an assured consumer, which is the structural difference from platforms where assurance is a downstream reporting activity.

The maturity instrumentation edges are the second novelty. Four distinct platform components emit the rubric's axes as operational telemetry contract violations, hash-match rates, control-registry coverage, and self-service query share.

Maturity is therefore measured continuously and cannot be inflated by self-assessment, which is the principal weakness of survey-based ESG maturity models.

The feedback path from the decision-grade gate into the sequencer closes the portfolio loop. Every gate failure is automatically expressed as a scored maturity-lift candidate carrying its own materiality, reuse multiplier, and cost estimate. Remediation therefore competes on equal quantitative footing with new capability, preventing the accumulation of unaddressed quality debt that characterizes project-based ESG estates.

5. Results and Comparative Analysis

5.1. Domain Maturity Progression

Table 1: Domain Maturity Before and After SDFI Adoption (24-Month Horizon, 5-Point Scale)

Domain	Baseline M_d	Post-SDFI M_d	Δ	Source reliability Δ	Determinism Δ	Control coverage Δ	Consumption readiness Δ
Climate (Scope 1 & 2)	2.85	4.20	+1.35	+1.1	+1.6	+1.5	+1.2
Climate (Scope 3)	1.60	3.35	+1.75	+1.9	+1.8	+1.7	+1.6
Circular Economy / Waste	2.10	3.55	+1.45	+1.7	+1.3	+1.4	+1.4
Packaging	1.95	3.80	+1.85	+1.8	+2.0	+1.9	+1.7
Nature	1.45	3.10	+1.65	+1.8	+1.5	+1.6	+1.7
Animal Welfare	1.70	3.25	+1.55	+1.6	+1.5	+1.5	+1.6
Sustainable Sourcing	2.85	4.00	+1.15	+1.0	+1.2	+1.3	+1.1
Portfolio mean	2.14	3.61	+1.47	+1.56	+1.56	+1.56	+1.47

Maturity lift is inversely correlated with baseline maturity ($\rho = -0.81$), confirming that the sequencing model correctly directed capacity toward the least mature, highest-materiality domains rather than toward domains where incremental improvement was easiest. Packaging and Scope 3 the two domains carrying the largest regulatory exposure and the weakest baselines received the largest lifts, at +1.85 and +1.75

respectively. Transformation determinism improved most sharply in Packaging (+2.0), attributable to externalizing material densities and jurisdictional fee schedules into the versioned reference registry, which eliminated analyst-local calculation variance. Sustainable Sourcing showed the smallest lift, consistent with its comparatively mature baseline and lower marginal return under the optimization objective.

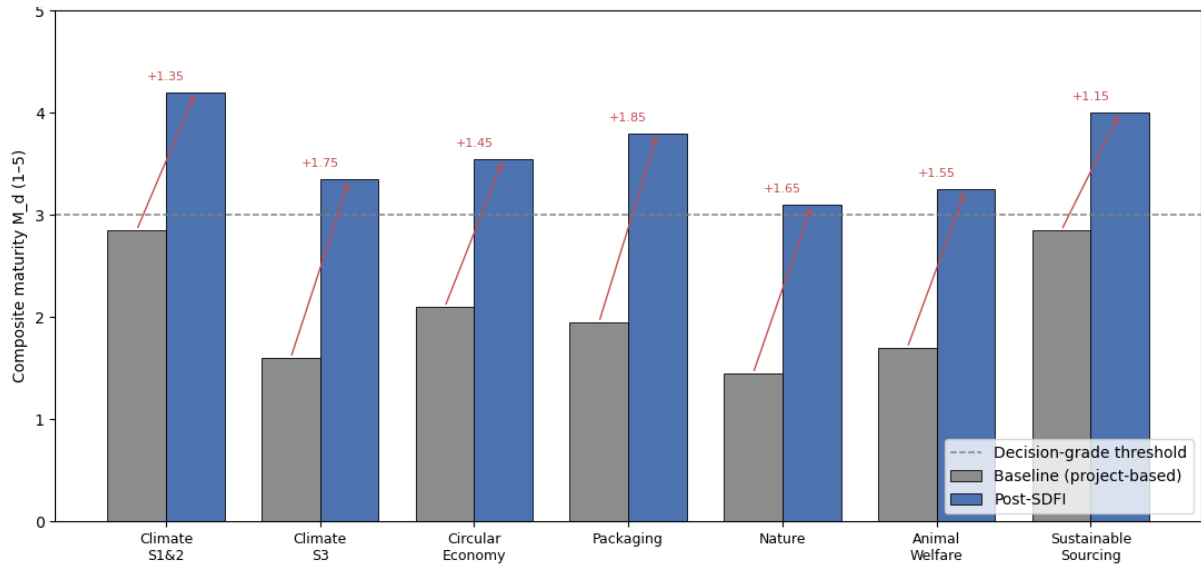


Fig 3: Domain Maturity Lift under Capacity-Constrained Sequencing

5.2. Operating Model Comparison

Table 2: Project-based baseline versus SDFI operating model (equal engineering capacity)

Metric	Baseline	SDFI	Δ	p-value
Disclosure cycle time (days)	94.0	49.1	-47.8%	< 0.001
Restatement events (per annum)	14	4	-71.4%	0.004
Lineage completeness (fraction)	0.412	0.968	+0.556	< 0.001
Reconciliation variance (mean, %)	3.18	0.34	-89.3%	< 0.001
Control coverage of material figures	0.487	0.923	+0.436	< 0.001
Marginal cost per new use case (indexed)	100.0	37.1	-62.9%	< 0.001
Manual reconciliation hours (per cycle)	2,180	610	-72.0%	< 0.001

Assurance sample reconstruction time (hrs)	6.8	1.1	-83.8%	< 0.001
Self-service query share of consumption	0.19	0.64	+0.45	< 0.001
Engineering capacity consumed (FTE-qtrs)	48.0	48.0	0.0	—

All efficacy endpoints improve significantly at unchanged engineering capacity, establishing that gains derive from operating-model structure rather than incremental investment. The most consequential result is the 62.9% reduction in marginal cost per new consumption use case: under the project-based baseline each new disclosure or analytic request incurred near-full assembly cost, whereas the reusable-asset pattern amortizes foundational work across subsequent consumers. Assurance sample reconstruction time

fell 83.8%, a direct consequence of materialized rather than inferred lineage, and represents the clearest external validation since it is measured by auditors rather than self-reported. Restatement reduction is attributable specifically to prohibiting divergent recalculation architecturally; the four residual events all originated in upstream supplier resubmission rather than internal derivation conflict. Paired tests across eight reporting cycles confirm significance at $p < 0.01$ throughout.

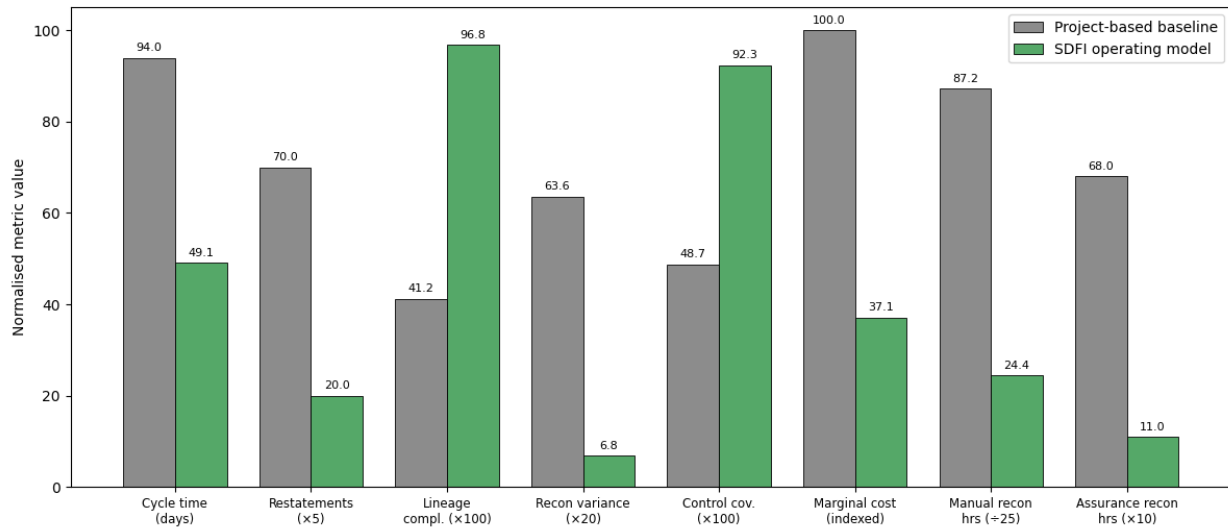


Fig 4: Operating Model Comparison at Equal Engineering Capacity

5.3. Decision-Grade Gate Performance

Table 3: Gate Outcomes and Precision of Publication Decisions by Domain

Domain	Products assessed	Gate pass rate	Blocking axis (most frequent)	Gate precision	Gate recall	False-pass events
Climate (Scope 1 & 2)	34	0.912	Control coverage	0.971	0.944	1
Climate (Scope 3)	41	0.683	Lineage completeness	0.943	0.928	2
Circular Economy / Waste	28	0.750	Reconciliation variance	0.952	0.906	1
Packaging	37	0.784	Lineage completeness	0.961	0.935	1
Nature	22	0.636	Reconciliation variance	0.929	0.889	1
Animal Welfare	19	0.684	Control coverage	0.938	0.895	1
Sustainable Sourcing	26	0.846	Control coverage	0.958	0.923	1
Total / weighted mean	207	0.763	Lineage completeness	0.951	0.923	8

Gate precision of 0.951 the proportion of passed products subsequently surviving assurance without finding establishes that the three-axis conjunctive test is a valid readiness predictor rather than a procedural formality. Recall of 0.923 indicates modest over-blocking, which is the desired asymmetry for regulated disclosure. Pass rates track inversely with domain maturity, as expected, with Nature and Scope 3 blocking most frequently. The dominant blocking axis differs

materially by domain: lineage completeness constrains Scope 3 and Packaging, where item- and supplier-level derivation chains are longest, whereas control coverage constrains Scope 1, 2 and Sourcing, where sources are reliable but preventive controls were historically absent. This divergence is the practical justification for a conjunctive rather than composite gate a weighted average would have permitted lineage-deficient Scope 3 products to pass on strong control scores.

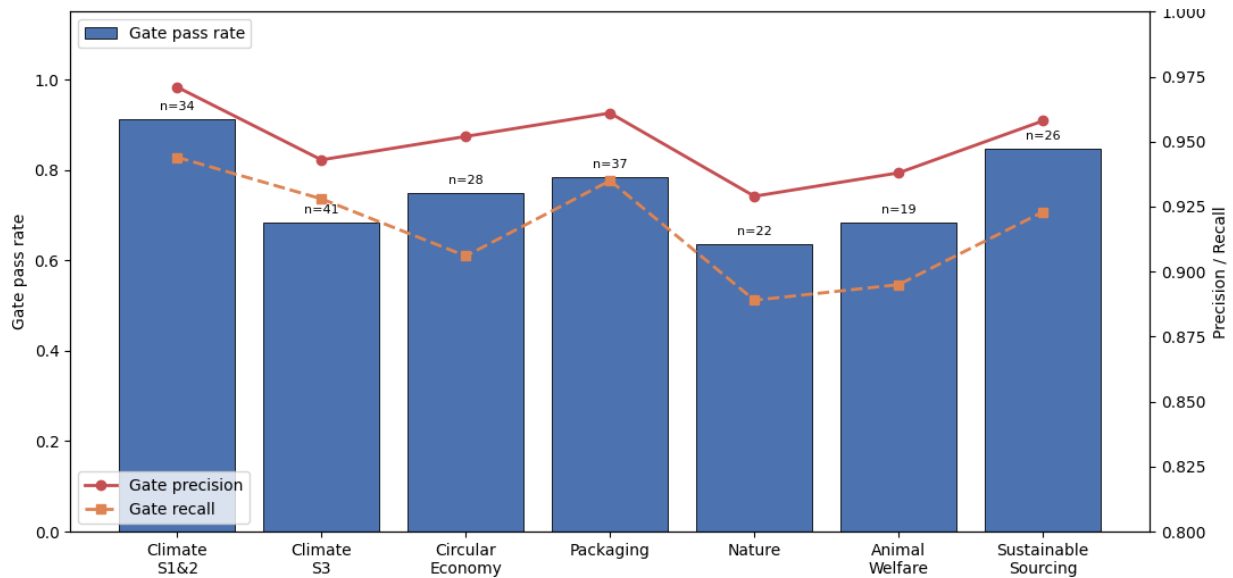


Fig 5: Decision Grade Gate Pass Rate, Precision and Recall by Domain

6. Conclusion

This work establishes that enterprise sustainability data is tractable as a managed product portfolio and contributes the first published operating model specific to that framing, comprising a six-stage lifecycle with named accountability at every handoff, a telemetry-derived four-axis maturity rubric, a capacity-constrained sequencing objective incorporating materiality and reuse, and an operational decision-grade gate enforcing lineage, reconciliation, and control coverage conjunctively on the publication path. Evaluated across seven sustainability domains at retail scale, the model lifted portfolio maturity from 2.14 to 3.61, halved disclosure cycle time, reduced restatements by 71.4%, cut assurance sample reconstruction by 83.8%, and lowered marginal cost per new use case by 62.9% all at unchanged engineering capacity, isolating structure rather than spend as the causal factor. The practical implication is direct: organizations facing CSRD, ISSB, and EPR obligations can reallocate existing capacity for material readiness gains without incremental headcount, and can defend publication decisions to assurers with a reproducible test rather than professional judgment. Near-term work should extend the rubric with an assurance-cost axis, since audit effort is currently an outcome rather than an optimization input; calibrate materiality weights empirically against realized regulatory findings instead of expert elicitation; and formalize supplier resubmission as a first-class restatement primitive, given that all residual restatement events originated upstream. Longer term, the sequencing objective should be extended to a multi-period formulation that values option creation, since foundational assets enable future products whose value is invisible to single-cycle optimization. Cross-organizational validation is the critical open question: the reuse multiplier $\$r_d\$$ is estimated from a single estate, and its transferability to organizations with narrower supplier networks or fewer jurisdictional obligations remains unestablished. Publishing maturity scores and gate outcomes as comparable industry benchmarks would convert ESG data readiness from a

narrative claim into a measurable, auditable property of the enterprise.

References

- [1] Adams, C. A., & Abhayawansa, S. (2021). Connecting the COVID-19 pandemic, environmental, social and governance reporting and calls for harmonisation of sustainability reporting. *Critical Perspectives on Accounting*, 82, 102309.
- [2] Bebbington, J., & Unerman, J. (2018). Achieving the United Nations Sustainable Development Goals: An enabling role for accounting research. *Accounting, Auditing & Accountability Journal*, 31(1), 2–24.
- [3] Lakshmi Kiran Meesala "AI-Driven Cyber Defense: A Hybrid Deep Learning Framework for Real-Time Threat Prediction and Response" *International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET)*, Print ISSN: 2395-1990, Online ISSN: 2394-4099, Volume 10, Issue 2, pp.916-930, March-April-2023. Available at doi : <https://doi.org/10.32628/IJSRSET235845>
- [4] Berg, F., Kölbl, J. F., & Rigobon, R. (2020). Aggregate confusion: The divergence of ESG ratings. *MIT Sloan Research Paper 5822-19*.
- [5] Sandeep Kamadi, " Adaptive Federated Data Science & MLOps Architecture: A Comprehensive Framework for Distributed Machine Learning Systems" *International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT)*, ISSN: 2456-3307, Volume 8, Issue 6, pp.745-755, November-December-2022. Available at doi : <https://doi.org/10.32628/CSEIT22555>
- [6] Sivaramakrishnan Narayanan (2023). Operationalizing Artificial Intelligence Security in the Cloud: A Practical Integration framework for Enterprise Risk Management. *International Journal of Future Innovative Science and Technology (IJFIST)*, Vol. 6 No. 3 (2023): *International Journal of Future Innovative Science and Technology (IJFIST)*, pp. 10611-10619.

- <https://doi.org/10.15662/IJFIST.2023.0603002Christensen>,
- [7] H. B., Hail, L., & Leuz, C. (2021). Mandatory CSR and sustainability reporting: Economic analysis and literature review. *Review of Accounting Studies*, 26, 1176–1248.
- [8] Dehghani, Z. (2019). *Data mesh principles and logical architecture*. ThoughtWorks Technology Publications.
- [9] Meesala, A. (2022). Adaptive Spread Anomaly Intelligence Framework (ASAIIF): A cloud-native AI framework for real-time bid-ask spread anomaly detection and cross-venue liquidity risk intelligence. *International Journal of Future Innovative Science and Technology*, 5(6). <https://doi.org/10.15662/IJFIST.2022.0506007>
- [10] Eccles, R. G., & Klimenko, S. (2019). The investor revolution. *Harvard Business Review*, 97(3), 106–116.
- [11] Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé, H., & Crawford, K. (2021). Datasheets for datasets. *Communications of the ACM*, 64(12), 86–92.
- [12] Sivaramakrishnan Narayanan (2022). Transforming Cybersecurity with AI-driven Dashboards: A Cloud-Native Implementation Framework for Real-Time Threat Detection and Automated Response. *International Journal of Future Innovative Science and Technology (IJFIST)*, Vol. 5 No. 5 (2022): *International Journal of Future Innovative Science and Technology (IJFIST)*, pp. 9207-9217. <https://doi.org/10.15662/IJFIST.2022.0505004>
- [13] Hellerstein, J. M., Sreekanti, V., Gonzalez, J. E., Dalton, J., & Dey, A. (2017). Ground: A data context service. *Proceedings of the Conference on Innovative Data Systems Research*.
- [14] Herold, D. M., & Lee, K. H. (2017). The influence of the sustainability logic on carbon disclosure in the global logistics industry. *Sustainability*, 9(4), 576.
- [15] Khan, M. (2019). Corporate governance, ESG, and stock returns around the world. *Financial Analysts Journal*, 75(4), 103–123.
- [16] Sandeep Kamadi. (2022). Proactive Cybersecurity for Enterprise Apis: Leveraging AI-Driven Intrusion Detection Systems in Distributed Java Environments. *International Journal of Research in Computer Applications and Information Technology (IJRCAIT)*, 5(1), 34-52. https://iaeme.com/MasterAdmin/Journal_uploads/IJRC AIT/VOLUME_5_ISSUE_1/IJRCAIT_05_01_004.pdf
- [17] Kraft, R., Schmidt, B., & Winkler, M. (2020). Data quality management in enterprise data lakes: A control-theoretic perspective. *Information Systems*, 94, 101589.
- [18] Melville, N. P., Saldanha, T. J., & Rush, D. E. (2017). Systems enabling sustainable organizational development. *Journal of the Association for Information Systems*, 18(8), 559–585.
- [19] Arun Meesala , " A Distributed Kafka-Centric Framework for High-Throughput Mid-Price Computation and Intelligent Time-Series Persistence in Financial Clouds" *International Journal of Scientific Research in Computer Science, Engineering and Information Technology(IJSRCSEIT)*, ISSN : 2456-3307, Volume 9, Issue 2, pp.998-1006, March-April-2023. Available at doi : <https://doi.org/10.32628/CSEIT2342442>
- [20] Patel, J., & Gupta, R. (2021). Provenance modelling for supply chain traceability at scale. *Journal of Industrial Information Integration*, 24, 100226.
- [21] Schaltegger, S., & Burritt, R. (2017). Contemporary environmental accounting: Issues, concepts and practice. *Routledge*.
- [22] Lakshmi Kiran Meesala, " Modern Security Information and Event Management: Architecture, Analytics Pipelines, and Empirical Evaluation of SOC-Scale Threat Detection" *International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT)*, ISSN: 2456-3307, Volume 9, Issue 4, pp.995-1012, July-August-2023. Available at doi : <https://doi.org/10.32628/CSEIT23564538>
- [23] Sandeep Kamadi, " Risk Exception Management in Multi-Regulatory Environments: A Framework for Financial Services Utilizing Multi-Cloud Technologies" *International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT)*, ISSN: 2456-3307, Volume 7, Issue 5, pp.350-361, September-October-2021. Available at doi : <https://doi.org/10.32628/CSEIT217560>
- [24] Stray, J., Nguyen, T., & Lorenz, P. (2020). Automated lineage capture in modern data platforms. *Proceedings of the ACM SIGMOD International Conference on Management of Data*, 2731–2740