

Decision-Centered Analytics

Agree the decision, measurement and acceptance test before producing the analysis

An anonymized operating case on an executive performance readout, its rejection at review, and a proposed evidence contract that connects analytical quality to the decision a sponsor must make.

PROJECT CONTEXT / EXECUTIVE ANALYTICS, AI-ASSISTED PRODUCTION & EVIDENCE GOVERNANCE

BLUF / BOTTOM LINE UP FRONT

Agree the decision before building the answer.
Verify the evidence again before release.

ANALYTICAL FINDING

Correct calculations cannot repair a question mismatch. A requested breakdown can also be mathematically invalid: overlapping categories change rank when allocation order changes. Negotiate a coherent answer before producing the chart.

CONTRIBUTION IN THE SOURCE RECORD

Directed the analytical workflow and readout. Provisional contribution; the record reports rejection and an undelivered rebuild.

EVIDENCE STATUS

Original artifacts were not recovered locally. This is a diagnostic and workflow-design case; no causal operating or financial benefit is established.

BLUF: analytical inference. Source contribution: historical record. Client and internal identities anonymized.

STRUCTURED ABSTRACT / THE EXECUTIVE DECISION

Validate the required answer before optimizing its production

[AI] An analytical document has several independent obligations: address the agreed decision, use an admissible population and source version, support its inference, and meet the authorized reviewer’s acceptance criteria. A polished response can fail any one of them.

Element	Decision-relevant finding
Context [HR]	A logistics leadership team questioned whether its input-metric strategy should change. An AI-assisted workflow decomposed the discussion, checked staged evidence, commissioned selected data cuts and produced an executive readout and follow-up answers.
Objective [HR]	Answer the leadership questions within the review cycle and connect the diagnosis to proposed actions. The reviewing sponsor instead required a monthly outcome series and a breakdown in the outcome’s own units.
Approach [HR/RC]	The original work combined source flags, controlled queries, internal verification and consolidated writing. This case audits request fit, category overlap, source completion, ratio validity and the stated rationale for retaining the input strategy.
Findings [HR/AI]	The register records rejection of the main documents and an unfinished rebuild. A later complete export reverses one subgroup’s direction. The requested breakdown mixes overlapping and sibling constructs, while the no-coverage argument conflicts with its own cited predictor evidence.
Decision value [PD]	Adopt a pre-analysis acceptance record: the sponsor agrees the decision and coherent output; the metric owner confirms population and readiness; the analytical lead verifies arithmetic and inference. Test the workflow before claiming reduced rework.
Implications [PD]	Require all release conditions to pass; a stronger writing score cannot offset an invalid denominator or unresolved request. Keep blocked, rejected and unanswered requests in the original evaluation cohort.

ACTION FROM THIS CASE

[PD] Pilot the Decision-Aligned Evidence Contract (DAEC). Confirm question, grain, source readiness and acceptance criteria before analysis; revalidate them at release and after any material source or requirement change.

[HR] Source S1: selected project register; originals remain unrecovered. [RC] recalculated; [AI] analytical inference; [PD] proposed design. Identities and raw source values remain private.

EXHIBIT 1 / SEPARATE CORRECTNESS FROM DECISION FIT

A fact-checked answer can still fail the decision it was commissioned to support

[HR] The recorded workflow included analytical verification and a reviewer agent. The human reviewer rejected the resulting documents for answering an adjacent question. This is a documented failure mode, not an estimate of the workflow’s failure rate.

TWO INDEPENDENT TESTS / NEITHER CAN SUBSTITUTE FOR THE OTHER

Proposed classification / evidence validity includes measurement and inference



A review verdict tests only its declared criteria. It cannot supply missing evidence or decision authority.

Acceptance dimension	Required evidence
Decision fit	The response answers the agreed choice, time horizon, population and requested output. An informative adjacent analysis belongs in supporting material unless the sponsor changes the brief.
Measurement validity	The quantities represent the claimed constructs, use compatible exposure and maturity, and survive source and join checks. A citation to a staged file does not establish those properties.
Inference validity	Descriptions, comparisons, causal effects and financial benefits are distinguished. A lower observed rate in one exposure group does not eliminate its causal effect.
Authorized acceptance	The actual decision owner assesses the agreed criteria. A simulated reviewer’s stylistic approval does not supply that authority, and human acceptance does not cure invalid evidence.

CHANGE THE FIRST REVIEW

[PD] Review an empty output shell and its metric definitions with the sponsor before commissioning the analysis. Require the sponsor and metric owner to resolve contradictions in the requested chart while changes are still inexpensive.

[HR] Source S1: selected project register; originals remain unrecovered. [RC] recalculated; [AI] analytical inference; [PD] proposed design. Identities and raw source values remain private. Diagram is a proposed logical classification, not measured historical performance.

EXHIBIT 2 / REPAIR THE REQUESTED BREAKDOWN

A different allocation order can manufacture a different top driver

[HR/AI] The requested breakdown combined a sibling outcome with overlapping exposure and service-segment views of the target outcome. Copying those labels into one Pareto would not create a valid partition. A grain footnote cannot repair double counting.

SAME RECORDS / DIFFERENT ALLOCATION ORDER / DIFFERENT WINNER

Wholly hypothetical: 100 target-outcome records; inbound tag 60, slot tag 50, both tags 40



Joint cells reconcile to 100. Assigned bars depend on a convention and do not identify causal contribution.

[RC] In the wholly hypothetical cohort, 100 target-outcome records contain 60 inbound-tagged records and 50 slot-tagged records, including 40 carrying both tags. The disjoint cells are inbound only 20, slot only 10, both 40 and neither 30. Adding the two marginal tags yields 110, not 100.

Allocation convention	Result in the same synthetic cohort
Inbound takes precedence	Inbound receives all 60 tagged records; slot receives its remaining 10; neither receives 30. Inbound is the largest assigned category.
Slot takes precedence	Slot receives all 50 tagged records; inbound receives its remaining 20; neither receives 30. Slot is the largest assigned category.
Retain joint membership	Show all four mutually exclusive cells and their definitions. Keep any separate sibling outcome outside the target-outcome total. This preserves accounting without pretending to assign physical cause.

THE NEW FINDING

[AI] Even a known overlap does not produce a unique priority ranking. The allocation convention can determine the winner. Agree the accounting question and show joint membership before treating an assigned bar as a controllable cause.

[HR] Source S1: selected project register; originals remain unrecovered. [RC] recalculated; [AI] analytical inference; [PD] proposed design. Identities and raw source values remain private. All diagram counts are invented; the construction is independently enumerated.

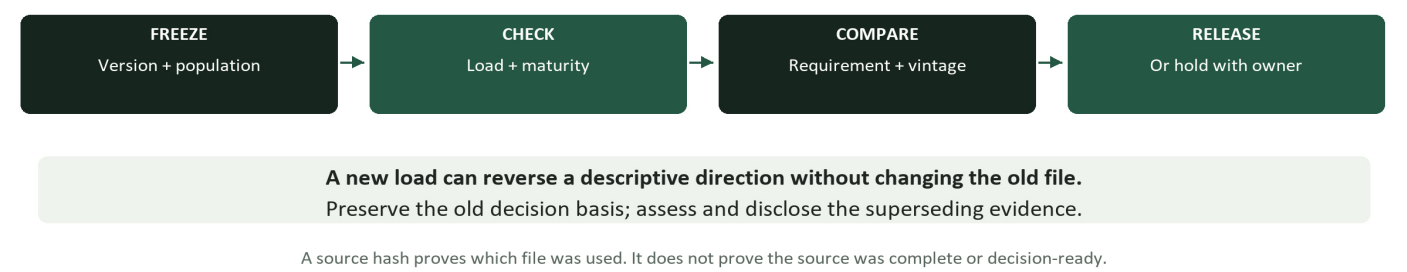
EXHIBIT 3 / MAKE SOURCE READINESS A RELEASE CONDITION

Freezing a snapshot preserves lineage but does not certify completeness

[HR/RC] The staged data warned that a reporting period lacked a final load. The subsequent complete export changes the subgroup’s comparison from a decrease to an increase. Both source versions remain in the private record; the completed version governs the stated full-period question.

AN AUDITABLE SNAPSHOT NEEDS A RELEASE-TIME READINESS CHECK

Proposed control / source completion and outcome maturity are separate requirements



Control	What must be checked
Freeze the version	Retain source identity, extraction timestamp, reporting interval, outcome maturity and a reproducible calculation. This supports auditability, not automatic fitness for every later decision.
Validate completeness	Confirm numerator and exposure cover the same promised period. A full-period denominator paired with an incomplete numerator can manufacture a lower rate.
Reopen on change	A newly loaded period, revised taxonomy, corrected scope or revised sponsor request triggers an impact assessment. Retire affected conclusions explicitly; do not silently replace the old file.
Release the right statement	A complete load can resolve a descriptive comparison while causal impact and process stability remain unresolved. A short ordered series is not a reliable operating-phase baseline.

[AI] The record’s completeness flags were useful but insufficient as release controls. The failure was allowing a known limitation to travel with a decision-ready headline instead of constraining which answer could be released.

DECISION RULE

[PD] If the requested horizon is unavailable, obtain an explicit revised scope or deliver a pending result with an accountable evidence owner. Do not label an incomplete-period comparison as the full-period answer.

[HR] Source S1: selected project register; originals remain unrecovered. [RC] recalculated; [AI] analytical inference; [PD] proposed design. Identities and raw source values remain private. Workflow is proposed; no historical rework saving is estimated.

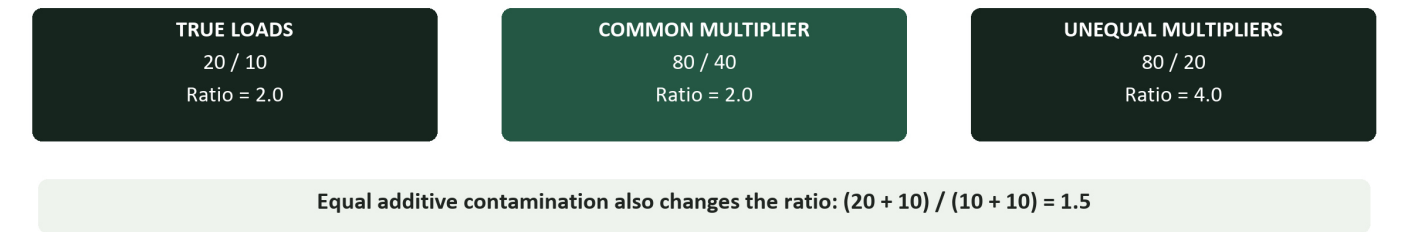
EXHIBIT 4 / TEST THE ASSUMPTION BEHIND A PREFERRED RATIO

A ratio is robust only when the distortions cancel in the required way

[HR/AI] The source favored a relative attach ratio over inflated absolute loads. That preference is not a proof of robustness. Unequal join multiplicity, outcome contamination or selection across the two groups can change the ratio itself.

RATIO ROBUSTNESS REQUIRES A SPECIFIC CANCELLATION CONDITION

Wholly hypothetical loads / equal fixed exposures in both groups



Measurement distortion and selection are assumptions to test; dividing two numbers does not remove them.

[RC] The invented true loads are 20 and 10 units per fixed exposure: ratio 2. Multiplying both loads by four preserves 2. Multiplying the first by four and the second by two produces 4. Common additive contamination also fails to cancel: adding 10 to each produces $30/20 = 1.5$.

Source claim	Evidence required before reuse
Observed attach is causal	Common eligibility, valid linkage and appropriate confounding control are required. A crude group comparison remains descriptive. A unit-per-package ratio is not automatically a probability of any concession.
No same-day spike disproves a link	Link exposure cohorts to mature outcomes before interpreting a temporal comparison. A plausible lag explains a test’s weakness; it does not prove the preferred mechanism.
Recorded owner is the wrong owner	Retrieve the applicable responsibility rule and facts it requires. A record’s ownership field plus a late upstream event does not establish an incorrect charge or a recoverable saving.

RELEASE THE BOUNDARY WITH THE NUMBER

[PD] Validate duplicate handling, outcome inclusion, filters and exposure in both groups. If the required checks are unresolved, keep the ratio as an exploratory data question rather than a released decision fact.

[HR] Source S1: selected project register; originals remain unrecovered. [RC] recalculated; [AI] analytical inference; [PD] proposed design. Identities and raw source values remain private. Numeric constructions are hypothetical and are not transformations of the source ratio.

EXHIBIT 5 / RECONCILE THE DECISION RATIONALE

A claim of no outcome coverage cannot be justified by a predictor of that same outcome

[HR/AI] The no-pivot argument says the largest outcome lies outside the current inputs’ reach. Its own cited evidence also describes a current input as strongest on that outcome. The register therefore does not supply a coherent proof of the claimed coverage gap.

THREE DIFFERENT CLAIMS / THREE DIFFERENT EVIDENCE REQUIREMENTS

Source contradiction: an outcome is described as uncovered and as strongly predicted by a current input

REPRESENTED
Is the outcome in scope?
Definition + ownership

PREDICTED
Can the input forecast it?
Validated predictive model

CHANGEABLE
Does an action affect it?
Credible causal evaluation

A prediction claim cannot establish causal control; a coverage claim cannot erase its own contrary evidence.

Resolve the decision rationale before presenting persistence or pivoting as an evidence-proven strategy.

Question	What the available record can support
Is the outcome represented?	Audit metric definitions and the mechanism map. An outcome can be included in an analytical model while lacking a staffed operational owner; those are different gaps.
Does the input predict it?	Inspect the model, cohort, timing, held-out performance and uncertainty. The cited association is not an independently recovered result in this case.
Can an intervention change it?	Define the actual action package and evaluate mature outcomes against a credible comparison. Predictive usefulness does not establish causal manipulability.
Should the strategy change?	Compare feasible alternatives, costs and decision uncertainty. Neither an aggregate plateau nor one contradictory argument establishes that a pivot or persistence is optimal.

[AI] A leadership team may retain a commitment while commissioning a better test. It should record that as a decision under uncertainty, not as a demonstrated consequence of a complete causal decomposition.

REOPEN THE PREMISE, NOT THE ENTIRE PROGRAM BY DEFAULT

[PD] Assign the analytical lead to reconcile the coverage statement with the cited input evidence. The operating sponsor then evaluates alternatives using the clarified mechanism and decision threshold.

[HR] Source S1: selected project register; originals remain unrecovered. [RC] recalculated; [AI] analytical inference; [PD] proposed design. Identities and raw source values remain private. The case diagnoses an internal contradiction; it does not estimate either strategy’s causal value.

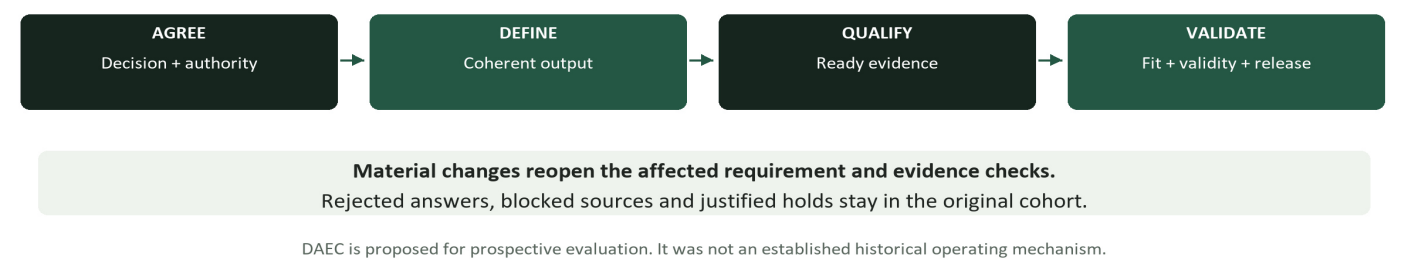
EXHIBIT 6 / DEFINE THE CONTRACT BEFORE THE WORKFLOW

Give every request a coherent output, an evidence owner and a release test

[PD] DAEC is a proposed acceptance record for the analytical request. It connects requirements to evidence without allowing either sponsor preference or a composite quality score to override invalid measurement.

DECISION-ALIGNED EVIDENCE CONTRACT / PROPOSED

An acceptance record from original request to final disposition



Contract field	Acceptance requirement
Decision and authority	State the decision, available alternatives, consequence of delay and person authorized to accept a revised question. Record disagreement rather than treating meeting attendance as agreement.
Output and semantics	Specify outcome, unit, population, horizon and chart structure. Demonstrate that totals, subgroups and sibling outcomes reconcile. Approve any overlap allocation explicitly.
Source readiness	Name evidence owners, permitted access, completeness tests, maturity rules and version cutoffs. A blocked source creates a tracked dependency; it does not authorize invented data or bypassing access controls.
Release and revision	Check fit, validity and authority independently at release. Preserve rejected versions, corrections and the original request clock; reopen affected decisions when a material assumption changes.

TEST THE WORKFLOW AS A WORKFLOW

[PD] Compare pre-analysis contracting and source-readiness checks with usual post-draft review. Count invalid approvals, valid but irrelevant answers, justified holds, effort and elapsed delay. Keep customer-outcome impact as a separate future estimand.

[HR] Source S1: selected project register; originals remain unrecovered. [RC] recalculated; [AI] analytical inference; [PD] proposed design. Identities and raw source values remain private. DAEC was formalized for this case; historical adoption and savings are not claimed.

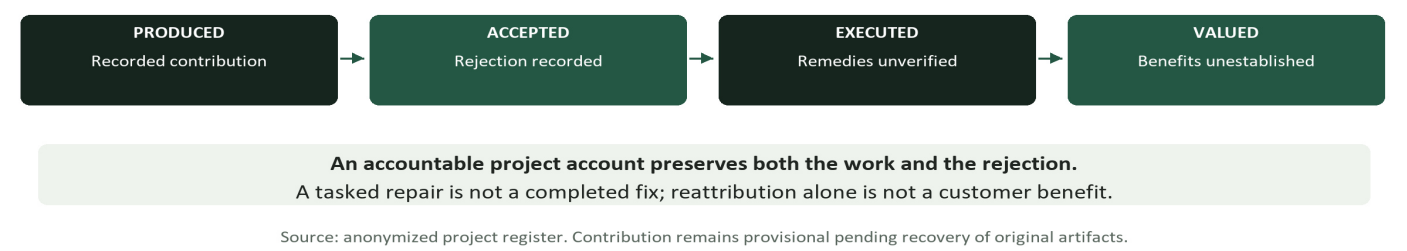
EXHIBIT 7 / ATTRIBUTE THE WORK WITHOUT REWRITING THE OUTCOME

Preserve analytical ownership and the recorded rejection in the same account

[HR/AI] The register attributes the framing, workflow direction, query authorization and editorial decisions to the operating lead. It also records contributions from specialists, a rejected deliverable, a corrected follow-up answer and an unfinished rebuild.

PRODUCTION / ACCEPTANCE / EXECUTION / IMPACT

Four distinct claims; evidence at an earlier stage cannot certify the next



Contribution claim	Process-tracing grade and boundary
Framing and production direction	Provisional hoop support: reported call notes, bylines and workflow records identify the role. Their original files remain unrecovered, so independent authorship authentication is incomplete.
Revising the preferred story	Provisional hoop support: the register describes a changed narrative after selected data cuts. Several headline analyses originate in related work and specialist inputs, not solely in this deliverable.
Accepted leadership answer	Not established for the principal documents: rejection is recorded. A later file timestamp does not show acceptance, presentation or a successful rebuild.
Operating or financial benefit	Not established: proposed targets, routed owners and a tasked query repair do not establish completed remediation, causal customer benefit or released cash.

[AI] The transferable contribution is the observable boundary between producing an answer and delivering one the decision can use. The proposed contract addresses that boundary; its effectiveness must be tested prospectively.

WHAT A COO CAN AUTHORIZE NOW

[PD] Require a decision contract for the next bounded review cycle, resolve the requested partition with the metric owner, and fund calibration of the workflow evaluation. Do not treat this case’s hypothetical economics as realized project value.

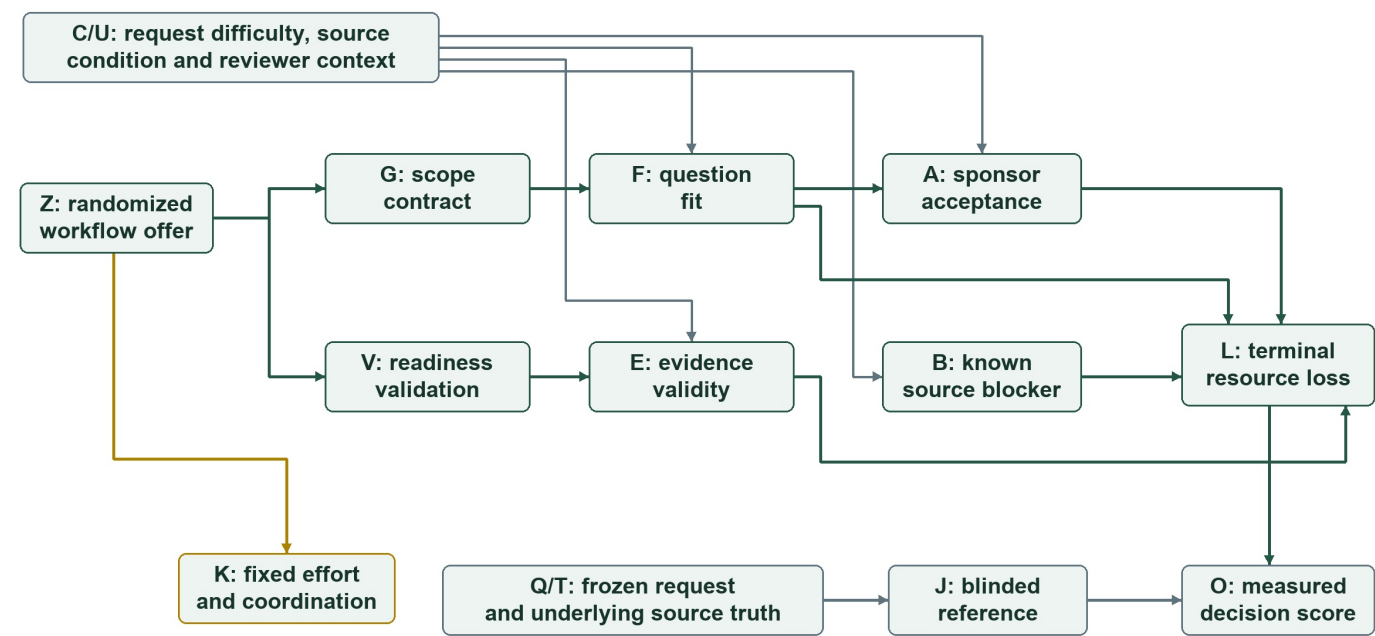
[HR] Source S1: selected project register; originals remain unrecovered. [RC] recalculated; [AI] analytical inference; [PD] proposed design. Identities and raw source values remain private.

TECHNICAL ANALYSIS

Decision fit and evidence validity travel through different paths

[AI/PD] The proposed Decision-Aligned Evidence Contract (DAEC) separates answering the agreed question, establishing a valid evidentiary basis and obtaining authorized acceptance. The source reports that internal verification passed while the requested decision remained unanswered.

PROPOSED BENCHMARK: FIT, VALIDITY AND ACCEPTANCE ARE DISTINCT



Fixed K is charged once outside terminal loss. Known blockers require a hold.
Reference J depends on the frozen request and facts; blinding does not make it unrelated to truth.

[PD] Randomized workflow Z supplies scope clarification G and source-readiness validation V. These affect request fit F and evidence validity E. Sponsor acceptance A responds to fit but is not the definition of validity. A known source blocker B requires a hold. F, E, A and B determine terminal decision-resource loss L; fixed implementation cost K is charged separately once.

[PD] Request/source context C/U can affect fit, validity, acceptance and blockage. Frozen request Q and source truth T inform independent reference J; the observed score O combines the actual response outcome with that reference. Independence means blinded ascertainment, not independence from the facts being judged. The benchmark changes no customer operation or policy.

[AI] Historical selection of polished or accepted documents would condition on downstream outcomes. Randomization instead retains every originally assigned request, including blocked, rejected and unfinished work. A later customer action needs its own treatment, cohort, authority and mature outcome comparison.

Basis: derivative source record [HR]; proposed graph and trial [AI/PD]. Arrows express hypotheses or score dependencies, not identified historical treatment effects.

TECHNICAL ANALYSIS

Agree the decision before freezing its evidence contract

[HR/AI] The register distinguishes a consolidated readout, a companion analysis, later follow-up answers and the rebuilt product the reviewer requested. Delivery and internal review did not establish acceptance. The literal monthly series and reconciled ranking were not completed in the supplied record.

Gate	Required observable acceptance test
Fact correctness	Reproduce each calculation from the authenticated version, with the same grain, filters, exposure and observation age.
Target fit	Identify the decision, alternatives, target outcome, requested period and comparison; show how the answer informs that decision.
Source admissibility	Confirm that categories, joins, clocks and available months can support the specified estimand. Keep unavailable evidence explicit.
Authority and acceptance	Record the authorized reviewer, accepted artifact version and any operating approval separately. An agent review is not that authorization.

[AI] With binary indicators f , t , s and a for these four gates, mandatory release eligibility is $f \times t \times s \times a$. This is a logical conjunction, not a probabilistic independence assumption. A weighted average of gate scores can conceal a failed requirement. The historical failure is not evidence that every calculation was wrong; it demonstrates that local checks did not close the decision contract.

[PD] Before analysis, obtain an agreed example output and rejection criteria. If a requested chart mixes outcome families with overlapping exposure tags, propose a coherent cross-tabulation or an explicitly conventional allocation. Record the agreed revision. Literal compliance with an invalid estimand must not be treated as success.

[HR/AI] The registered instruction to reuse staged data guarded against unsupported invention, but a later completed export superseded that snapshot. Freeze a version, not a permanent truth claim: define expiration, owner notification, dependency revalidation and supersession. The source does not establish exactly what reached the subsequent executive meeting.

[AI] NASA distinguishes verification against specifications from validation for intended use. Gotel and Finkelstein add the missing trace: how the requirement itself was created and revised. These concepts explain a failure mode; neither establishes an effect size for DAEC.

Basis: historical rejection, staged-data rule and blocked rebuild [HR]; four-gate proof [AI]; NASA (2016) and Gotel–Finkelstein (1994). Public examples do not expose source values.

TECHNICAL ANALYSIS

A fully linked overlap table can still reverse the top category

[AI] Consider an independent hypothetical universe of 100 unique Late entities. An inbound tag appears on 60 and a slot tag on 50. Their overlap may range from 10 to 50. Suppose entity linkage establishes that it is exactly 40.

Joint membership	Entities	Inbound-first	Slot-first
Both tags	40	Inbound	Slot
Inbound only	20	Inbound	Inbound
Slot only	10	Slot	Slot
Neither tag	30	Other	Other
Totals: inbound / slot / other	100	60 / 10 / 30	20 / 50 / 30

[AI] Both exclusive summaries are arithmetically correct. Inbound leads under the first convention; slot leads under the second. Let alpha be the fraction of the overlap assigned inbound. Its allocation is $20 + 40 \times \alpha$; the slot allocation is $50 - 40 \times \alpha$. They tie at $\alpha = 3/8$. A convention alone changes the ranking, with no change to events, evidence or causes.

[AI] In general, for sets of sizes a and b in a universe N, $\max(0, a + b - N) \leq \text{overlap} \leq \min(a, b)$. Knowing the overlap solves descriptive double counting, not causal attribution. Here the union is 70; adding both tag counts and the neither group produces 140. Adding a separate Lost metric to a Late-only target would introduce an additional universe error.

[PD] Publish the joint cells when overlap is decision-relevant. If an exclusive reporting allocation is required, declare its precedence, retain the joint table and show ranking sensitivity. Label it a descriptive allocation. A causal contribution requires specified interventions and their interactions; summing isolated prevention opportunities can count the same avoidable outcome twice.

[HR/AI] The source itself flags the requested ranking’s grain problem. A footnote cannot make incompatible rows additive. This constructive proof strengthens the response to that request: negotiate a valid question before producing a persuasive but arbitrary top-driver chart.

Basis: exact independent enumeration in private/counterexample.py [AI]; source category warning [HR]. All numbers on this page are synthetic, not disguised historical values.

TECHNICAL ANALYSIS

A common query does not validate its ratio or its causal story

[HR/AI] The register identifies incompatible extraction bases, a contaminated concession numerator, an incomplete final period and a corrected measurement-grain answer. These are measurement-contract failures to reconcile, not reasons to prefer whichever headline is most persuasive.

Claim to test	Required correction
A ratio survives common contamination	Equal multiplicative distortion cancels; unequal weights or additive contamination generally do not. Audit entity-level multiplicity by comparison group.
Arrival lateness explains customer lateness	Separate arrival after schedule from arrival after cutoff, and each from the original communicated customer promise.
A booked owner is the wrong owner	Authenticate the governing allocation rule. Booking location, physical custody, prevention control and response authority are different constructs.
Current data are complete	Bind event period, extraction time, missing final-day status, family-specific maturity and the latest approved replacement.

[AI] If true group loads are r_1 and r_0 and distortion factors are k_1 and k_0 , the observed ratio is $(k_1/k_0) \times (r_1/r_0)$. The same query text does not establish $k_1 = k_0$. Even a perfectly calculated ratio of concession units to packages need not be a probability. Removing unknown statuses changes the population and requires explicit missing-outcome bounds.

[HR/AI] The source’s crude exposed/unexposed rates do not refute a causal effect: selection, shared causes, cohort differences and uncertainty remain. An imported regression or one contrary week does not identify the effect of changing an operating input. Daily aggregate misses divided into daily concessions are not linked conversion without entity, clock and scope alignment.

[HR/AI] The no-pivot rationale also conflicts with its own cited analysis: one passage says current inputs do not address Late, while another describes a current input as most strongly associated with it. That invalidates the scope argument as stated. It proves neither intervention efficacy nor that the portfolio should pivot. The decision requires admissible causal evidence, costs, authority and alternatives.

Basis: preserved source conflicts [HR]; ratio algebra and causal interpretation [AI]; Wang–Strong1996 and Zave–Jackson1997. Reconciliation is unresolved without original query and entity evidence.

TECHNICAL ANALYSIS

Audit relevance, validity and acceptance as separate measurements

[PD] The unit of validation is an original request paired with its submitted or missing response at the locked deadline. A document count, a clean style review or a named action owner cannot substitute for this unit.

Measurement	Independent evidence and denominator
Fit to the agreed question	Blinded judgment against the frozen decision, scope and output example; all assigned requests remain in the denominator.
Evidence validity	Reference checks of population, overlap, joins, vintage, maturity and supported inference. Known mandatory failures require a hold.
Sponsor acceptance	Version-specific reviewer disposition. Accepted evidence-invalid means acceptance with E = 0; accepted off-target but evidence-valid work is a separate loss category.
Decision-resource burden	Prespecified costs for erroneous acceptance, valid adjacent work, rejection/rework and justified hold; add incremental workflow costs once.

[PD] Independently double-code a random core of 200 request-output judgments, supplemented by deliberately difficult stale, inaccessible, overlap and accepted-invalid examples. Under an iid worst-case agreement approximation, the core gives a 95% screen of about ± 6.9 percentage points. Shared source/request clusters reduce effective sample size; recompute the design effect. Kappa precision requires prevalence-specific pilot simulation, not this agreement interval.

[PD] Report the full confusion matrix, category-specific error and agreement as well as kappa. Oversampled exceptions cannot estimate population accuracy without selection weights. Fit and validity adjudicators must be blind to workflow assignment and sponsor acceptance, using evidence available at the decision deadline. Merge acceptance only after those labels are locked. The hidden reference is reserved for scoring and cannot leak into tuning.

[AI] Operational access has its own test: a successful identity check does not grant resource or device-policy entitlement. A failed headless path is not repaired by assuming a browser session grants equivalent rights. Use an authorized route or return an explicit unavailable-source disposition; do not manufacture missing monthly data.

[PD] Preserve revisions to both the request and its evidence. A later changed question is a new contract version, not a retrospective reason to mark an earlier response correct.

Basis: proposed classification audit [PD], Cohen (1960); source access and version-control limits [HR/AI]. No actual agreement, acceptance probability or data-access reliability has been estimated.

TECHNICAL ANALYSIS

Sparse review-period movements do not certify an operating phase

[HR/AI] The independent source audit retains 57 source and sensitivity screens plus five mechanical proposed-target audits. Historical values and their indexed counterparts remain private. These computations screen arithmetic patterns; they do not establish intervention effects.

[AI] The register's short series mix initial and completed vintages, a holiday period, under-mature endpoints and incompatible exposure definitions. Eight source/sensitivity screens show provisional crossings: the initial and completed topline and Late series, with and without the holiday, contain individual-value or moving-range exceedances. Preserve both the original sequence and any exclusion sensitivity. A selected exclusion cannot manufacture a stable phase.

[AI] Individuals limits use $\text{mean} \pm 2.66 \times \text{mean moving range}$; the moving-range upper limit is $3.267 \times \text{mean moving range}$. A new endpoint may cross limits frozen from earlier observations even when refitting all observations absorbs it. The owner-share endpoint has this problem and a different maturity status. Neither a crossing nor its disappearance identifies the cause. No eligible stable phase or Laney count/exposure panel is established.

[PD] A future process panel should preserve the same question class, source basis, original deadline and independently scored outcome. Track accepted evidence-invalid errors and off-target rework by exposure as well as total effort. Use control limits only after assessing phase consistency, dependence and changing request mix. Deming's warning about reacting to noise applies; declining to react does not excuse a verified source defect.

[AI] No historical worst-N treatment cohort is identified. For a proposed selective rollout, test thresholds, inclusion of all eligible domains, repeated baseline windows and untreated comparators. Under an explicitly assumed linear conditional-mean model, expected untreated follow-up is $m + \rho \times (x - m)$, where m is the population mean and ρ assumed persistence. Values 0.4, 0.7 and 0.9 imply 60%, 30% and 10% reversion of selected excess. These are hypothetical sensitivities, not estimated historical fractions.

[PD] Prespecify source accessibility, overlap complexity, reviewer tenure, request ambiguity and baseline difficulty as moderators. Adjust for pretreatment variables; do not condition the total effect on acceptance, completed delivery or successful validation.

Basis: frozen private source/SPC audit [RC/AI]; NIST, Laney, Barnett and Deming. Sparse or disputed series cannot justify stability, regression-to-the-mean fractions or causal attribution.

TECHNICAL ANALYSIS

Contribution is bounded by rejection, missing originals and unresolved joins

[HR/AI] The contribution grade is a provisional registry-level hoop: the record supports involvement in framing, authorization, consolidation and explicit gap disclosure. It does not authenticate sole execution, final reviewer acceptance, implemented remediation or customer impact.

Contribution claim	Support and limiting evidence
Directed analytical work	The register describes framing and authorization of a workflow with multiple analysts and reviewers. Original notes, run records and deliverables are unrecovered.
Revised the narrative	The record documents revised drafts and disclosed uncertainty. Some revised conclusions still overstate observational evidence; correction is not proof of a causal null.
Delivered review artifacts	Delivery and internal review are reported alongside explicit rejection. The requested rebuilt artifact remained blocked in the supplied record.
Created lasting operating value	A query-fix assignment and follow-up scopes are reported. Receipt, deployment, sustained use and measured outcomes are not established.

[AI] The source is a retrospective registry with repeated quotations and possible shared lineage. Several metrics are imported from sibling analyses with different populations and clocks. Repetition across a narrative, evidence row and memory is not independent corroboration. A byline or first-person agent narrative alone cannot establish who personally executed a query or edit.

[AI] Consolidated limits are material: originals are unrecovered; date/version chronology is incomplete; joins and filters remain disputed; the requested categories are not an identified partition; causal effects and economic values are absent; and no validated post-intervention operating panel exists. Absence of a recovered completion record does not establish whether work occurred elsewhere. The present case’s reconciliation and constructive proofs are new analysis, not retrospective credit for the original project.

[PD] Evidence that would strengthen the grade includes authenticated request revisions, source/query hashes, reviewer disposition on the exact release, independently reconciled entity totals and a deployed change log with mature outcomes. Evidence of rejection or a failed mandatory check must remain visible even when it weakens the preferred story.

Basis: source-derived process tracing [HR/AI], Collier2011. Historical analysis, present reconstruction and prospective design are separate claim classes; no treatment effect is attributed.

TECHNICAL ANALYSIS

Test the workflow on every request before claiming customer value

[PD] The proposed controlled benchmark compares a scope-contract plus source-readiness bundle with usual post-draft quality review. Independent reporting domains receive frozen batches sampled from one eligibility and information law, with equal authority and access; no real customer action is executed.

[PD] Lock each response at original intake + 48 hours and finish hidden independent adjudication by intake + 7 days. Retain every assigned request, including blocked sources, missing responses, rejected outputs and accepted-invalid answers. Known source blockers mandate justified holds. Erroneous release in the planning law represents undetected invalidity, not permission to waive an explicit failure.

Estimand	Interpretation
Primary ITT benefit	Mean terminal resource loss under usual workflow minus that under the offered bundle, over all original requests. Fixed incremental scope, readiness and coordination cost is charged separately once.
Safety gate	Usual minus bundle probability of sponsor acceptance with invalid evidence ($A = 1, E = 0$), on the original denominator. Accepted off-target but evidence-valid work remains in primary loss.
Component exploration	Scope-only, readiness-only and interaction contrasts distinguish fit from validity. Confirmatory claims require their own allocation, multiplicity and power plan.
Later field effect	A separate operating decision and mature customer cohort; benchmark rework savings cannot be relabeled customer savings.

[PD] Prespecify the loss schedule before observing arm results and elicit the willingness to pay for reductions in errors and rework. A detection design must resolve an economically meaningful benefit and meet the accepted-invalid guardrail; full-effect power alone is insufficient. The following model pages quantify an entirely synthetic finite law. Reference error, actual difficulty mix, domain dependence and recruitment feasibility require calibration before a real benchmark.

[PD] If a later operating policy adopts in stages, estimate group-time ATT against not-yet-treated domains, with no anticipation, support and credible conditional parallel trends. Use joint pretrend diagnostics and equivalence checks; address concurrent changes and interference. Mature customer counts and a pilot are required for field power. That extension remains unpowered and unestimated here.

Basis: proposed randomized benchmark [PD]; Callaway–Sant’Anna (2021) for a separate possible field design. Equal assigned information is not a claim that realized request batches are paired.

TECHNICAL ANALYSIS

Four propositions make the contract testable

[AI/PD] DAEC treats a useful answer as a conjunction of decision relevance, admissible evidence and authorized acceptance. It does not reward agreement with a preferred narrative or require a causal ranking when the data support only a descriptive cross-tabulation.

Falsifiable proposition	Prespecified observation that challenges it
P1: Scope clarification reduces adjacent answers	On blinded agreed-question judgments, the bundle fails to reduce valid-but-off-target responses by the minimum material amount.
P2: Readiness checks protect accepted decisions	Risk of acceptance with invalid evidence fails its required improvement or worsens under a prespecified high-risk source stratum.
P3: Joint checks justify their resource cost	A lower confidence bound on all-request benefit does not clear incremental cost, or required reference/feasibility gates fail.
P4: The ranking answers a coherent question	Reasonable declared precedence or unresolved membership changes the top category; a causal-driver claim must then be withdrawn or independently identified.

[AI] Zave and Jackson distinguish a definition from an empirical claim: a precedence rule can define a partition without revealing causes. Wang and Strong explain why accurate values can still lack contextual quality. Cash and colleagues connect useful evidence to relevance, credibility and legitimate participation; an influential reviewer’s preference cannot replace scientific adequacy. Amershi and colleagues motivate explicit capability limits, disambiguation and correction in AI-assisted work, not automated substitution for acceptance.

[AI] The incentive design matters. Campbell and Goodhart warn that a score targeted for reward can displace its purpose. Acceptance and document throughput therefore need independent validity and rework measures. Ittner and colleagues caution against assuming more measures or claimed alignment automatically improve performance. OECD composite-indicator guidance motivates transparent weights and sensitivity; mandatory source or authority failures remain noncompensating gates. No cited framework supplies this project’s missing effect estimate.

[PD] Preregister material thresholds, reference procedures and stopping rules. An imprecise result is inconclusive, not proof that DAEC has no value; a high acceptance rate is insufficient evidence that it works. Retain negative findings and revise the contract if scope alignment merely makes unsupported answers easier to accept.

Basis: new testable synthesis [AI/PD] grounded in primary requirements, data-quality and human–AI research; continuing incentive and composite-measurement literature. DAEC is a proposed framework, not an historically validated intervention.

Reviewer acceptance can rise while accepted errors worsen

[PD] Randomize independent analyst–reviewer domains in a 2 × 2 benchmark: a decision-question contract (A), a source-readiness check (B), both, or usual post-draft review. Full bundle versus usual is the sole confirmatory contrast; components and interaction are exploratory.

[PD] Each domain receives 12 frozen requests from one eligibility law, conditionally independent given shared difficulty. Freeze the original intake, lock the response at 48 hours and finish blinded reference grading by day 7. Retain blocks, unanswered requests and rejected drafts. All cells have identical source access and authority.

Workflow	Decision loss	Sponsor accepts	Accepted invalid
Usual	6.502	53.0%	15.1%
Question contract	5.865	66.3%	19.5%
Readiness check	4.868	53.0%	5.9%
Both	3.752	66.3%	7.7%

[PD] Fit and semantic/source validity are graded separately. A literal but incoherent Pareto can still be invalid. Authors do not observe hidden truth. Accepted-invalid counts sponsor-approved evidence-invalid answers; accepted adjacent answers are separately priced in decision loss. Known source or authority blocks require holds, never unsupported release.

DECISION STANDARD

[PD] The question-only arm increases acceptance while increasing accepted-invalid answers. The full bundle has the same acceptance rate but materially fewer accepted errors. Do not optimize acceptance alone.

[PD] All probabilities and resource weights are invented. Decision loss excludes the separately allocated new protocol work; no customer or historical effect is claimed.

PROPOSED FACTORIAL TEST / 2 OF 4

Size the joint decision gate, not the attractive full-effect forecast

[PD] Exact finite domain-score convolutions calibrate the positive-direction economic test and the accepted-invalid gate. Independent domains and the entire specified probability law are planning assumptions; repeated requests do not replace teams.

[PD] The response family mixes usual and full outcome laws within each difficulty class. Exact rational CDF checks establish marginal ordering for decision loss and accepted-invalid outcomes. This is not observed compliance, direct factor interpolation or a claim of joint pathwise dominance.

Design	Per cell	MDE index	BE detection	Joint floor
Initial	16	33.75	23.2%	4.2%
Base candidate	80	15.27	80.6%	99.3%
Same size / stress	80	15.30	80.4%	65.5%
Conditional redesign	112	12.96	91.6%	84.1%

[PD] Break-even is 15.38 index points. Require 80% positive-direction detection at break-even and 80% joint power to clear cost and reduce accepted-invalid answers. Two-sided zero-test error is at most 5%; economic and safety lower bounds each use 97.5% confidence. A union bound supplies joint power.

[PD] Moderate stress retains 75% of the assumed readiness gains. Halving them leaves only 16.3% power for the accepted-invalid gate at the redesign; its joint lower bound is uninformative. Component-effect significance is not authorized by this sizing.

RECRUITMENT CONDITION

[PD] The redesign needs 448 independent domains and 5,376 original requests. Feasibility, reference quality and reviewer dependence require calibration before recruitment.

[PD] Fixed usual decision loss = 100. MDE uses the correctly directed rejection tail; exact discrete bounds apply only to the disclosed finite response family.

Price better decisions separately from new production work

[PD] Decision-loss units price wrong accepted decisions, rejected rework and deadline holds. They are arbitrary resource equivalents, not historical dollars. Common drafting expense cancels. New work $K = 1.00$ per original request separately includes question negotiation 0.30, readiness checks 0.40 and coordination 0.30.

[PD] Favorable decision-loss reductions total 2.853297; added correctly framed but falsely rejected work costs 0.103224. $Net_raw = a \times h \times 2.853297 - 0.103224 - K$. $Net_index = 100 \times Net_raw / 6.501819$. Avoidability a and non-overlap h discount favorable components only.

Avoidability a	$h = 0.60$	$h = 0.80$	$h = 1.00$
0.50	-3.80	0.59	4.97
0.75	2.78	9.36	15.95
1.00	9.36	18.14	26.92

[PD] Table entries are net index points at $K = 1.00$. With $a = h = 1$, $K = 0.75 / 1.00 / 1.50$ yields 30.76 / 26.92 / 19.23 points. Reprice and re-power other economic scenarios.

[PD] All-hold loss is 6 per request, exceeding the modeled full bundle including K . If at most ϵ of each arm lacks reliable reference truth, conservative loss-contrast uncertainty is 40ϵ and accepted-invalid uncertainty is 2ϵ . Five-percent uncertainty erases the base expected net lower bound; missing labels are not zero errors.

[PD] Future field net = eligible treated volume $\times$ mature customer ATT improvement $\times$ avoidable cost/outcome $\times$ non-overlap – incremental cost. Those inputs are unavailable. The benchmark supplies no customer valuation or field power.

ECONOMIC BOUNDARY

[PD] A signed request contract is necessary. It does not turn an invalid overlap, stale export or unavailable source into usable evidence.

[PD] Synthetic benchmark ITT, fixed original request pool and decision deadline. New protocol work is charged once; research-only hidden adjudication has a separate study budget.

PROPOSED INFORMATION VALUE / 4 OF 4

Information can justify renegotiation or an investment hold

[PD] The evidence-purchase decision is separate from modeled workflow benefit. Its mutually exclusive states are a coherent ready deliverable, a question or partition needing negotiation, or evidence/access unavailable at the deadline.

Action / state	Ready	Negotiate	Unavailable
Bounded readout trial	16	-15	-21
Negotiate scope	3	13	-4
Retain / defer	2	1	9

[PD] Priors are 0.45 / 0.30 / 0.25. A three-category diagnostic reports its true state with accuracy q and each other state with probability $(1 - q)/2$. Downstream action costs are already in utilities; information acquisition is charged separately.

Accuracy	Gross EVSI	Acquisition cost	Net EVSI
0.55	1.2475	0.75	0.4975
0.80	5.6100	1.75	3.8600
0.95	8.2275	3.25	4.9775

[PD] All 27 contingent policies are enumerated. Prior optimal value is 4.2500; EVPI is 9.1000. Gross EVSI is bounded by EVPI. Utilities are arbitrary and are never added to benchmark savings or historical exposure.

AUTHORITY REMAINS EXPLICIT

[PD] Evidence may support a bounded trial, negotiated output or investment hold. It cannot override known access blocks or require presenting overlapping categories as an additive Pareto.

[PD] Signal accuracies, priors and utilities are hypothetical. This model remains meaningful without assuming a positive customer or workflow effect.

Reproduce the requested quantity before judging the usefulness of its answer

[AI] The Decision-Aligned Evidence Contract (DAEC) distinguishes a coherent request, admissible evidence, valid inference and authorized acceptance. Its proposed benefit is lower decision loss across all original requests, not a higher rate of favorable review verdicts alone.

Claim class	Meaning and replication boundary
[HR] Historical record	The supplied register describes the analytical workflow, delivered documents, rejection, revised source values, corrected follow-up and blocked rebuild. These descriptions remain derivative until the original artifacts are recovered.
[RC] Recalculated	Version-specific arithmetic, rounding intervals, short-series process screens, overlap bounds and allocation-order counterexamples. Public numerical constructions are hypothetical; confidential source values remain in the private annex.
[AI] Analytical inference	Internal verification cannot substitute for validating the required decision. A literal chart request can itself be incoherent. A relative ratio does not cancel arbitrary measurement distortion, and a recorded ownership label does not establish blame.
[PD] Proposed design	A decision contract and source-readiness workflow evaluated against usual review. Proposed evaluation, resource values, intervention costs and information values are scenarios, not observed project effects.

Replication sequence

[PD] Freeze the original register and exact source cells. Recover the request, source files, actual deliverable versions, correction record and reviewer disposition. Resolve whether quoted metadata describe the version that was rejected or a later revision before drawing a sequence or authorship conclusion.

[PD] Recalculate each source version in its own population and horizon. Retain the incomplete and completed-period branches, rounding intervals, alternative join cuts and missing category intersections. A missing monthly series remains missing; weekly observations cannot reconstruct the requested monthly history.

[PD] Reproduce the joint table, allocation conventions and declared evaluation law. Bind numerical and visual reviews to the final files; retain superseded versions and their reasons.

THE INTERPRETATION THIS CASE CHANGES

[AI] The response needs an agreed decision, a mathematically coherent output and evidence ready for that output. Increasing verification effort on an adjacent answer leaves those obligations unresolved.

The separate private package contains raw values, identity mappings, source locators and executable checks. The public case discloses methods and protected conclusions.

RESEARCH REFERENCES / DECISION INTENT AND ANALYTICAL QUALITY

Validate the intended use as well as the calculation

Primary research supports the proposed workflow and its interpretation. It does not authenticate unrecovered project artifacts or supply a historical effect, adoption rate or financial return.

NASA (2016). NASA Systems Engineering Handbook, NASA/SP-2016-6105 Rev2. Official handbook; selected verification, validation and configuration-management sections. Compliance with a specification differs from fitness for stakeholder use. This case transfers that distinction to the analytical brief and requires revalidation after a material source change; it does not assert that a NASA standard governed the historical project.

Wang, R. Y., & Strong, D. M. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems*, 12(4), 5–33. DOI 10.1080/07421222.1996.11518099. Original article and publisher record. The consumer perspective separates intrinsic and contextual quality. Copying a source accurately does not establish relevance, timeliness, completeness or appropriateness for the requested decision.

Gotel, O. C. Z., & Finkelstein, A. C. W. (1994). An Analysis of the Requirements Traceability Problem. *Proceedings of the First International Conference on Requirements Engineering*, 94–101. DOI 10.1109/ICRE.1994.292398. Original institutional paper. Trace the origin and refinement of the question as well as the source of each number. A document-to-data map does not recover whose decision was intended or who can approve a change.

Zave, P., & Jackson, M. (1997). Four Dark Corners of Requirements Engineering. *ACM Transactions on Software Engineering and Methodology*, 6(1), 1–30. DOI 10.1145/237432.237434. Original article; selected sections. Ground terms in their external meaning and distinguish definitions from empirical assertions. A precedence rule defines an exclusive allocation; it does not discover which mechanism caused a defect.

Cash, D. W., et al. (2003). Knowledge Systems for Sustainable Development. *Proceedings of the National Academy of Sciences*, 100(14), 8086–8091. DOI 10.1073/pnas.1231332100. Coauthor-hosted original article. Decision relevance, evidential credibility and legitimate participation can require separate management. Transfer to executive analytics is a proposed application; approval by an influential reviewer does not establish factual validity.

Amershi, S., et al. (2019). Guidelines for Human-AI Interaction. *Proceedings of CHI 2019*, Paper 3, 1–13. DOI 10.1145/3290605.3300233. Original author manuscript; guidelines and validation sections. Make system limits visible and support clarification and correction. The study supports interaction-design principles, not the proposition that an AI reviewer can replace the sponsor or an independent evidence reference.

Access scope is recorded privately. Literature applications are methodological or conceptual; outside empirical effects are not attributed to this project.

RESEARCH REFERENCES / EVALUATION AND UNCERTAINTY

Evaluate original requests under their real dependencies

Primary research supports the proposed workflow and its interpretation. It does not authenticate unrecovered project artifacts or supply a historical effect, adoption rate or financial return.

NIST/SEMATECH. e-Handbook of Statistical Methods, section 6.3.2.2: Individuals Control Charts. Official method. The case recomputes tentative XmR bands under complete-series, holiday, maturity and source-version branches. Short, changing-definition series do not identify a stable process phase. Tentative crossings prompt source and process investigation; they are not treatment effects.

Laney, D. B. (2002). Improved Control Charts for Attributes. *Quality Engineering*, 14(4), 531–537. Primary publisher abstract. Eligible attribute charts require actual event counts and exposures across repeated rational subgroups. Rounded rates and fractional concession units do not create a valid Bernoulli panel; no Laney chart is estimated from the missing counts.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20, 37–46. Primary publisher record. Independently judge request fit, source validity and sponsor disposition; report disagreement, prevalence and uncertainty. Agreement with a reviewer is not the same construct as correctness, and a chance-adjusted coefficient cannot replace an external reference.

Barnett, A. G., van der Pols, J. C., & Dobson, A. J. (2005). Regression to the Mean: What It Is and How to Deal with It. *International Journal of Epidemiology*, 34, 215–220. Primary abstract. A peak-selected baseline or a worst-unit cohort can improve through reversion. Fix selection before follow-up and label reliability-based reversion scenarios as assumptions when untreated repeated data are absent.

Hudgens, M. G., & Halloran, M. E. (2008). Toward Causal Inference With Interference. *Journal of the American Statistical Association*, 103, 832–842. Original manuscript. Shared analysts, reviewers, source custodians and deadlines can couple requests. Randomize independent domains and preserve those dependencies in the prospective evaluation rather than treating every answer as independent.

Callaway, B., & Sant’Anna, P. H. C. (2021). Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics*, 225, 200–230. Author manuscript. A later observational operating rollout could estimate group-time effects under explicit comparison, anticipation, overlap and parallel-trend assumptions. The proposed analytical-workflow benchmark does not estimate or power that separate customer-outcome study.

Strong, M., Oakley, J. E., Brennan, A., & Breeze, P. (2015). Estimating the Expected Value of Sample Information Using the Probabilistic Sensitivity Analysis Sample. *Medical Decision Making*, 35, 570–583. Author repository. Value information through the actions its signal changes, after acquisition cost and within the same utility model. Keep this evidence-purchase decision separate from hypothetical workflow savings.

Access scope is recorded privately. Literature applications are methodological or conceptual; outside empirical effects are not attributed to this project.

RESEARCH REFERENCES / CONTRIBUTION AND GOVERNANCE

Keep acceptance, validity and impact as separate claims

Primary research supports the proposed workflow and its interpretation. It does not authenticate unrecovered project artifacts or supply a historical effect, adoption rate or financial return.

Collier, D. (2011). Understanding Process Tracing. *PS: Political Science & Politics*, 44, 823–830. Primary publication. Separate framing, production direction, authorization, acceptance and implementation. Reported bylines and metadata supply provisional role evidence until recovered originals permit authentication. A produced document cannot satisfy the evidential test for an accepted or implemented decision.

Tucker, A. L., & Edmondson, A. C. (2003). Why Hospitals Don't Learn from Failures. *California Management Review*, 45(2), 55–72. Author manuscript. A rejected answer and an unresolved source dependency can expose a recurring delivery failure. Record the requirement and evidence defect instead of repairing only the current document's wording; effectiveness of that workflow change remains a testable proposition.

The W. Edwards Deming Institute. The Funnel Experiment. Official exposition. Reacting to routine variation can worsen performance. An executive interpretation of short-period movement should distinguish source-load changes and special conditions from sustained operating change before it triggers repeated intervention or an input-strategy pivot.

Campbell, D. T. (1976). Assessing the Impact of Planned Social Change. Original institutional record. An acceptance-rate target can change how analysts frame uncertainty, escalate blocked evidence or omit difficult requests. Preserve the original request cohort and measure invalid approvals, unresolved work and decision loss alongside completed outputs.

Goodhart, C. A. E. Problems of Monetary Management: The U.K. Experience. 1975 conference; revised volume listed as 1976 in the RBA archive. A proxy can lose its meaning when targeted. The case's sponsor-acceptance measure is deliberately separated from evidential correctness; neither a favorable verdict nor a high composite review score is treated as sufficient success.

Ittner, C. D., Larcker, D. F., & Randall, T. (2003). Performance Implications of Strategic Performance Measurement in Financial Services Firms. *Accounting, Organizations and Society*, 28, 715–741. Author manuscript. The choice and use of performance measures must connect to the decision mechanism. More tables, verification passes or produced answers do not themselves establish lower operating cost or better customer outcomes.

OECD / European Commission JRC (2008). Handbook on Constructing Composite Indicators. Official handbook. Aggregation requires defensible definitions and treatment of compensability. Keep mandatory request-fit, source and authority conditions separate from an average quality score. A high style or factual-copying score cannot offset a failed requirement that makes the answer unusable.

Access scope is recorded privately. Literature applications are methodological or conceptual; outside empirical effects are not attributed to this project.

FINAL REFEREE REPORT / INTERNAL AGENT-ASSISTED REVIEW

Accept the diagnostic lesson; test the proposed workflow

[AI] This author-led, agent-assisted review is not external peer review or client approval. Its disposition concerns evidence reconciliation, the decision-changing counterexamples and a prospective workflow evaluation.

Major issue	Disposition in the revised case
Verification confused with acceptance	Separated factual and inferential validity, coherent request fit, and authorized acceptance. A simulated reviewer cannot act as the sponsor; a sponsor’s approval cannot validate false evidence.
Literal compliance with an invalid chart	Replaced the mixed-grain Pareto with a joint-membership demonstration. The same records produce different ranked allocations when precedence changes. A footnote is insufficient to repair an invalid sum.
Snapshot presented as full-period truth	Preserved conflicting source versions and their interpretation changes. Completeness, denominator alignment and maturity are release conditions, not descriptive footnotes.
Preferred narratives treated as proven	Bounded crude exposure contrasts, ratio robustness and recorded-owner interpretation. Flagged the contradiction between an uncovered-outcome rationale and a cited predictor of that outcome.
Acceptance rate treated as value	Specified a prospective evaluation that prices invalid approvals, relevant rejections, adjacent answers, blocked sources, holds and intervention work. Operating effects and information value remain separate.

Partner critique: can a COO act on page one?

[AI] Yes: require a coherent decision brief before production, a source-readiness test before release, and a named sponsor and metric owner to resolve conflicts. The immediate commitment is to a bounded workflow test and calibration, not an unsupported claim that the historical input strategy or a proposed remedy succeeded.

Remaining conditions

[AI] Recover originals and version chronology; validate joins and taxonomy; calibrate request mix, reviewer behavior, costs and dependence. Historical acceptance, completed remediation and financial return remain unestablished.

REVIEW DISPOSITION / PASS WITHIN STATED SCOPE

[AI] The diagnostic case and proposed DAEC evaluation pass the stated internal review gates. The verdict does not authenticate unrecovered originals, certify historical causal impact or establish the modeled workflow’s effectiveness in deployment.

Internal evidence, methods, numerical and executive review. Release acceptance is limited to the documented source, model and confidentiality boundaries; no outside endorsement or ranking is implied.