MANUSCRIPT PREPARED FOR HIGHER CATEGORICAL SEMANTICS

Panoptical Observability, Informational Integration, and Organizational Boundaries

Trace semantics, effective descent,
copyable resources, and conditional consolidation

Anonymous

Version 1.0 November 5, 2026

Abstract. This paper develops a conditional mathematical account of organizations that record, integrate, and reuse the traces of their own activity. Operational systems are represented by internal higher categories in an infinity-topos. A chosen localization specifies which differences between executions are irrelevant, while a trace functor specifies what is observed. We distinguish equivalence reflection, mapping-space reconstruction, effective execution, and statistical identification. We prove that full faithfulness of an already defined internal trace functor is local for effective epimorphic covers, and separately explain when compatible local functors descend. Colimits and Kan extensions describe integration relative to specified interfaces; they do not identify the organizations supplying those interfaces.

The economic layer is a finite optimization problem on organizational partitions. Lax value and oplax cost comparisons give sufficient merger inequalities, but their applicability depends on the tensor, unit, permissions, externalities, and information regime. In particular, nonnegative globally superadditive value is incompatible with nontrivial idempotent information resources. We give deterministic, probabilistic, governance, and economic counterexamples to stronger conclusions. We establish exact finite consolidation and termination results, a constrained version, a bottleneck criterion for improving merger paths, and conditions for stable feedback. The resulting theory separates categorical structure from additional computational, empirical, and institutional hypotheses. It supplies no theorem that artificial intelligence necessarily produces a single firm or that complete traces are obtainable from ordinary telemetry.

Keywords: internal infinity-category; effective descent; trace fidelity; Kan extension; resource theory; organizational partition; privacy; feedback.

Contents

Reader's guide and evidential status **4**

1 The architecture of the question **5**

2 Ambient topoi, contexts, and operational categories **7**

3 Operational equivalence, localization, and truncation **9**

4 Trace fidelity and the limits of reconstruction **12**

5 Local completeness and effective descent **15**

6 Observational modalities and authorized knowledge **18**

7 Colimit integration and data migration **19**

8 Governance, interfaces, and the cost of coherence **22**

9 Probabilistic observation and decision sufficiency **23**

10 Symmetric-monoidal resources and the choice of tensor **26**

11 Partitions, merger gains, and exact consolidation criteria **30**

12 Constraints, externalities, and robustness **32**

13 Continuous complementarity and finite integration gains **34**

14 Worked organizational and informational examples **36**

15 Trace-mediated feedback and organizational dynamics **39**

16 Changing theories, learning, and observational scope **42**

17 An empirical program with explicit falsifiers **44**

18 The combined separation theorem **46**

19 Ideal trace embeddings and the role of probes **48**

20 A homogeneous cost model with exact optimal boundaries **51**

A Expanded internal and descent calculations **54**

B Finite observation, reversibility, and information bounds **56**

C Finite partition algorithms and path certificates **59**

D Reusable information and alternative valuation regimes **62**

E Organizational presentations and semantic invariance **65**

F A restricted identification theorem and its boundaries **67**

G Noisy organizational choice and stationary distributions **70**

H Technical concordance and notation **71**

I Prior work and the scope of attribution **73**

J A context-dependent end-to-end countermodel **76**

References **79**

Reader's guide and evidential status

This manuscript expands the problem posed by the web note *Panoptical Observability and Informational Consolidation in Infinity-Topoi* [1]. The source supplies the motivating separation between observation, integration, and ownership. The present treatment develops that separation using independently stated definitions, detailed proofs, new worked examples, and explicit failure conditions. No authorship or publication date is inferred from the source's domain. Its role is conceptual provenance, not independent empirical evidence.

The foundational results about infinity-topoi, localization, and categorical colimits are established mathematics [2, 3, 4]. The local full-faithfulness argument and the finite partition arguments below are applications and elementary deductions from those foundations. Their proofs are included to make the use of hypotheses inspectable. The contribution claimed here is a systematic semantic architecture, several useful compatibility and impossibility observations, and an explicit map of which inferences require extra assumptions. We do not claim to discover effective descent, the universal property of a colimit, superadditivity, or the theory of the firm.

All numerical examples are constructed examples. They are not measurements of an existing company or forecasts of an industry. The stochastic models use finite sets unless stated otherwise. They do not require an internal theory of probability in an arbitrary infinity-topos. The operational semantics are mathematical models rather than a claim that human activity or organizational knowledge is exhaustively computable.

The manuscript was prepared with AI assistance for research, drafting, symbolic reasoning, and document production. The arguments are presented for human mathematical review; no proof assistant certification or independent peer review is claimed. Anonymous presentation does not remove the submitting author's responsibility for the manuscript and its interpretations. Submission, acceptance, and publication are distinct events; this file is a prepared manuscript.

How to read. Sections 1–6 develop observation. Sections 7–9 address integration and statistical information. Sections 10–14 analyze resource aggregation and organizational boundaries. Sections 15–18 treat feedback, evidence, and the combined result. Sections 19–20 supply ideal probe embeddings and an exact cost model. The appendices provide expanded derivations, a comparison with the motivating note, and a reference guide.

1 The architecture of the question

1.1 Four different objects of explanation

A productive organization can observe a transaction, store a representation of it, combine that representation with other information, and use the result to change its future behavior. Each verb introduces a different mathematical question. Observation concerns a map from activity to records. Representation concerns what distinctions that map preserves. Integration concerns coherent relations between records produced in different contexts. Reorganization concerns the distribution of authority, assets, and decision rights. A single diagram can contain all four processes without making them equivalent.

Let $\mathcal{X}$ be an infinity-topos, $\mathcal{O}$ a category of operations internal to $\mathcal{X}$, and S a specified collection of arrows of $\mathcal{O}$. Assuming the indicated localization exists, write

$$\mathcal{O} \xrightarrow{q} \mathcal{K} = \mathcal{O}[S^{-1}] \xrightarrow{F} \mathcal{D}. \tag{1}$$

The category $\mathcal{D}$ describes traces together with the transformations that an observer recognizes between traces. An invariant trace functor $T : \mathcal{O} \to \mathcal{D}$ factors through q. The class S is part of the specification. It may express changes of implementation, translations between interfaces, or transformations certified to preserve an outcome. It does not acquire its meaning merely from being called "irrelevant."

Independently, let $D : I \to \mathcal{X}$ describe information objects and their declared interfaces. Their colimit D_* is an integrated object relative to that diagram. An ownership partition P belongs to a third mathematical setting, the finite poset $\mathrm{Part}(E)$ for a set E of capabilities. To compare partitions one needs an objective $W(P)$ and a family of feasible moves. Neither is supplied by the existence of F or D_*. The distinction is a type distinction before it is an economic disagreement.

Definition 1.1 (Four claims)**.** The following assertions will be kept separate:

(i) a semantic trace functor F is fully faithful;

(ii) a coherent diagram of information objects has an implemented integration;

(iii) a decoder or controller computes useful actions from the traces;

(iv) some organizational partition maximizes a specified objective and is dynamically reachable.

Claim (i) is a statement about mapping spaces. Claim (ii) includes a categorical object and an implementation question. Claim (iii) requires computability and resource bounds. Claim (iv) requires an economic and institutional model. The existence of arrows connecting these settings is useful, but composition alone does not transfer properties between settings.

1.2 What is higher-categorical about the model?

An ordinary event table can record that a transformation occurred. A category additionally records composable transformations. A higher category can record equivalences between transformations, transformations between those equivalences, and the coherent compatibility of all these structures. This matters when two integrations differ by an interface translation, when different translations agree only up to a certified comparison, and when those comparisons must remain compatible across a large network.

Consider a shipment identifier translated successively across a warehouse system, a payment processor, and a delivery service. Pairwise translations can all exist while their composite around a loop changes the identifier. A commuting triangle is already a constraint absent from a list of pairwise connections. If the translations are only equivalent rather than equal, a comparison between the two composite paths must be supplied. For longer paths those comparisons must themselves agree. Higher coherence encodes these requirements; it does not create the missing translations or certify their business meaning.

There are also cases where higher structure contributes nothing essential. A finite partition optimization problem is an ordinary poset problem. A deterministic lookup table may be modeled in sets. We retain those simpler models where appropriate. Putting a finite inequality inside an infinity-topos does not strengthen the inequality. The gain comes from relating it correctly to context-dependent objects and coherent transformations.

1.3 The meaning of "panoptical"

We use *panoptical observability* as a technical name for a strong, explicitly relative representation property. It is always relative to an operational theory, a class of irrelevant distinctions, an observational context, and an authorization regime. It does not mean that an observer sees every physical fact, knows every person's intentions, predicts every future event, or can execute every represented skill. The historical and social uses of panopticism concern asymmetries of visibility and power; our technical definition is not a substitute for those analyses [20, 21].

The relativity is mathematically unavoidable. If two sensor arrangements distinguish different features, they induce different observation maps. If two contracts distinguish different outcomes, they induce different candidate localizations. If permissions change, the set of authorized queries changes. A trace may be complete for delivery status while losing information about the driver's health, or complete for a compiler's input-output behavior while omitting the human reasons for a design decision. Such omissions can be intentional and desirable.

1.4 A claim ledger

The principal logical dependencies can be recorded without a causal forecast:

$$T(S) \subseteq \mathrm{Eq}(\mathcal{D}) \Longrightarrow T \simeq Fq, \tag{2}$$

$$u \text{ effective epi}, \quad u^*F \text{ fully faithful} \Longrightarrow F \text{ fully faithful}, \tag{3}$$

$$D_* \simeq \operatorname*{colim}_I D \Longrightarrow \mathrm{Map}(D_*, Y) \simeq \lim_{I^{\mathrm{op}}} \mathrm{Map}(D_i, Y), \tag{4}$$

$$\Delta W(P \to P') > 0 \text{ on all binary mergers} \Longrightarrow \{E\} \text{ uniquely maximizes } W. \tag{5}$$

The antecedents in these lines have different sources. The first is an invariance condition, the second a descent hypothesis, the third a universal property, and the fourth an assumption about values and feasible organizations. Their separation is the central discipline of the paper.

2 Ambient topoi, contexts, and operational categories

2.1 Size and foundational conventions

Fix a universe of small spaces and a larger universe in which presentable categories are formed. All indexing categories for colimits are small in the first universe; internal categories and localizations are assumed small enough for the constructions being used. This avoids treating a proper class of all possible executions as an object without a size convention. We write $\mathcal{S}$ for spaces and $\mathrm{Cat}_\infty(\mathcal{X})$ for complete Segal objects in $\mathcal{X}$.

An infinity-topos can be presented as an accessible left-exact localization of a presheaf infinity-category. We use its finite limits, small colimits, universal colimits, effective groupoid objects, and descent. These are standard structural properties, not hypotheses about markets [2]. A slice $\mathcal{X}_{/B}$ is again an infinity-topos. For a map $v : A \to B$, pullback $v^* : \mathcal{X}_{/B} \to \mathcal{X}_{/A}$ preserves finite limits and all small colimits.

Definition 2.1 (Context and cover)**.** A context is an object $B \in \mathcal{X}$. A family $\{U_i \to B\}$ covers B when the induced map $u : \coprod_i U_i \to B$ is an effective epimorphism. The overlap of two members is $U_i \times_B U_j$.

One must specify the base B. Arbitrary objects bearing company names need not jointly cover the terminal object. Likewise, a product $U_i \times U_j$ is a meaningful overlap only relative to a chosen common context. We usually work over a fixed B and suppress it by replacing $\mathcal{X}$ with $\mathcal{X}_{/B}$. This makes the locality claims domain-relative rather than universal statements about an economy.

For an ordinary sheaf topos on a space, one can think of U_i as open domains on which a system is specified. In a general infinity-topos the covering map need not be an inclusion. Covers can represent atlases, alternative presentations, or locally available choices. The

vocabulary of organizational context is interpretive; the mathematical definition remains an effective epimorphism.

2.2 Complete Segal objects

A simplicial object $C_\bullet : \Delta^{\mathrm{op}} \to \mathcal{X}$ is Segal when, for $n \geq 2$, the canonical comparison is an equivalence:

$$C_n \simeq C_1 \times_{C_0} \cdots \times_{C_0} C_1. \tag{6}$$

It is complete when its object of identity arrows identifies with its object of equivalences in the usual Rezk sense. Completeness makes the presentation correctly account for equivalence of objects [3]. We use the resulting theory of internal higher categories, including internal functor categories when available [4].

An operational category W has states in W_0, operations in W_1, and composable strings in its higher degrees. The Segal condition says that composition is coherent; it does not say that every formal arrow is physically executable at zero cost. An interpretation may assign runtimes, failure probabilities, or authorization conditions to arrows. These decorations are extra data.

Example 2.2 (A deployment category)**.** Objects can be versioned deployments, arrows can be certified migrations, and equivalences can be migrations with certified inverses. Two migrations from version a to version b might be related by a comparison showing the same observable behavior. A higher comparison can certify that two such certificates compose consistently. If only a preorder of allowed deployments is needed, its nerve gives a simpler operational category.

The internal mapping object between two generalized objects is best described over the parameter context. If $x, y : V \to C_0$, set

$$\underline{\mathrm{Map}}_C(x, y) := V \times_{C_0 \times C_0} C_1 \quad \text{in } \mathcal{X}_{/V}. \tag{7}$$

The underline distinguishes this object from an external mapping space. Checking only global points $1 \to C_0$ is insufficient in an arbitrary topos: an object can have important local sections without global sections. Later full-faithfulness claims are assertions about the internal object, hence about every generalized context.

2.3 Executions and arrow categories

There are two useful conventions for the domain of a trace. In the first, one chooses an execution category $\mathcal{O}$ directly. In the second, one sets

$$\mathcal{O} = \mathrm{Arr}(W) := \mathrm{Fun}([1], W). \tag{8}$$

Objects of $\mathrm{Arr}(W)$ are arrows $f : x \to y$ of W. Its morphisms are coherent squares connecting arrows. A square between f and g is not automatically a proof that f and g perform the same task. It may change both the initial and final states. If task preservation matters, one must select the permitted squares, restrict to suitable fibers, or add a task map.

For example, let $p : \mathcal{O} \to \mathcal{B}$ record task specifications. A vertical transformation in a fiber $\mathcal{O}_b$ compares executions of the same task. A transformation over a nonidentity arrow in $\mathcal{B}$ compares different specifications. Localizing all such arrows would collapse distinctions that the specification was introduced to preserve. We therefore keep $\mathcal{O}$ general and make no claim that the unrestricted arrow category is the unique correct model.

2.4 Ownership is different data

Suppose C is a category of production dependencies. A functor from C into a discrete set of owners cannot send an arrow across distinct owners, since a discrete category has no such arrow. Thus a naive ownership functor into a discrete owner set would forbid the very interfirm dependencies one wants to study.

There are several repairs. One can label the objects of C by a function $\mathrm{Ob}(C) \to F$ without claiming functoriality. One can use an owner category whose arrows represent contracts or interoperable relations. Or one can use a fibration of capabilities over owners, with separately specified cross-owner interfaces. The finite partition model later uses the first repair. An object labeling is enough to specify ownership blocks; it need not encode every dependency as a morphism of firms.

This elementary typing issue illustrates a broader principle. The categories of operations, information, resources, and ownership may be related, but their morphisms mean different things. Treating them as interchangeable can silently introduce strong hypotheses about what agreements, conversions, or observations are possible.

3 Operational equivalence, localization, and truncation

3.1 A relative operational theory

Assumption 3.1 (Existence of a skill localization)**.** The internal category $\mathcal{O}$ and specified class S admit an internal localization $q : \mathcal{O} \to \mathcal{K}$ satisfying the expected functorial universal property. Its construction is compatible with the context changes used below, or this compatibility is supplied as part of the model.

This is automatic in familiar small external settings with suitable universe choices. Internally, one should invoke an appropriate existence theorem or construct the localization in a presentation. The paper does not infer existence for an arbitrary unbounded internally specified class merely from the notation $[S^{-1}]$.

The defining property is that restriction along q induces an equivalence

$$\mathrm{Fun}(\mathcal{K}, \mathcal{D}) \simeq \mathrm{Fun}_S(\mathcal{O}, \mathcal{D}), \tag{9}$$

where the right side is the full subcategory of functors sending S to equivalences. A skill is an object of $\mathcal{K}$ under the chosen interpretation. It is a semantic equivalence class with higher structure, not a claim that all competent human activity has been effectively classified.

Proposition 3.2 (Invariant traces descend through the localization)**.** *Under assumption 3.1, an S-invariant functor $T : \mathcal{O} \to \mathcal{D}$ has a factorization $T \simeq Fq$. The space of factorizations together with the specified comparison is contractible.*

Proof. The object T belongs to the right side of (9). An inverse equivalence supplies an object F and an identification $q^*F \simeq T$. The homotopy fiber of an equivalence over an object is contractible. This proves the asserted uniqueness in its correct higher-categorical form. It does not assert that there is exactly one literal expression for F in a programming language. □

Example 3.3 (Syntax preservation can break invariance)**.** Let two certified implementations be connected by an arrow in S. A logger that stores source-code hashes in a discrete trace category may assign different hashes. The logger then fails S-invariance. One can retain the richer logger while replacing its target by a category containing the certified comparison, or project to a coarser invariant trace. Neither repair is free: the first changes the trace semantics and the second discards information.

3.2 Localization is not a decision procedure

The universal property permits reasoning with semantic equivalences even when no algorithm decides them. Extensional equivalence of arbitrary programs is a standard example. Rice's theorem rules out broad classes of uniform semantic decision procedures [15]. A localization can therefore be mathematically meaningful while its object equivalence relation is not computably decidable.

This distinction affects claims about learning from logs. Storing a program identifier may give a computable injection of syntax into strings. It does not produce a decidable canonical identifier for the program's extensional behavior. Conversely, a semantic construction might identify equivalent programs without supplying a computable normalization. The existence of F in proposition 3.2 does not resolve either issue.

3.3 Why levelwise truncation needs care

One may wish to retain only bounded homotopy information. It is tempting to apply object truncation $\tau_{\leq n}$ to every degree of a complete Segal object and declare the result an internal truncated category. The difficulty is that truncation of objects generally preserves products

but not all pullbacks. The Segal condition (6) involves pullbacks over the object space, so levelwise truncation cannot be used without a preservation argument.

Counterexample 3.4 (Truncation is not generally left exact)**.** In spaces, the homotopy pullback $* \times_{S^1} *$ is $\Omega S^1 \simeq \mathbb{Z}$ as a discrete space. Hence

$$\tau_{\leq 0}(* \times_{S^1} *) \simeq \mathbb{Z}, \qquad * \times_{\tau_{\leq 0} S^1} * \simeq *. \tag{10}$$

These are not equivalent. Thus objectwise 0-truncation fails to preserve this pullback.

The counterexample does not say that every levelwise truncation of every Segal object fails. It shows that an unrestricted finite-limit argument for that construction is invalid. Safe alternatives are available. For external infinity-categories, one can use homotopy categories or a standard categorical homotopy truncation. For an enriched presentation with fixed discrete object set, one can truncate the mapping spaces and verify composition using the preservation of finite products. In the internal setting, one can choose a suitable categorical truncation construction and state its universal property.

Definition 3.5 (Chosen categorical truncation)**.** A categorical n-approximation of $\mathcal{K}$ is a functor $h_n : \mathcal{K} \to \mathcal{K}^{[n]}$ in a specified categorical model, universal for the selected class of targets with bounded mapping types. The indexing convention for n must be declared in that model.

We do not identify $\mathcal{K}^{[n]}$ with levelwise object truncation. If $n = 1$ is used to mean an ordinary homotopy category externally, mapping spaces have become sets of components. The resulting assertion is weaker than full recovery of the original mapping spaces. If instead one truncates mapping spaces to n-types, there is a shift in conventions. Explicit conventions prevent a claim about ordinary arrows from being mistaken for a claim about all higher homotopies.

3.4 Identity, equivalence, and noninvertible maps

An infinity-category has a mapping space $\mathrm{Map}_C(x, y)$, but not every point of that space is invertible. The space of equivalences $\mathrm{Eq}_C(x, y)$ is a union of components of that mapping space. It is the path space from x to y in the maximal subgroupoid $C^{\simeq}$. Consequently,

$$\mathrm{Path}_{C^{\simeq}}(x, y) \simeq \mathrm{Eq}_C(x, y), \tag{11}$$

whereas $\mathrm{Map}_C(x, y)$ also includes directed transformations. Calling the entire mapping space an "identity type" without this qualification conflates directed transformation and equality.

For a groupoid-valued operational theory every morphism is invertible, and the distinction disappears. For a production system with irreversible operations it is essential. A trace

representation may faithfully describe a one-way migration without suggesting that the migration is an equality of states. The formalism should preserve that asymmetry.

4 Trace fidelity and the limits of reconstruction

4.1 A hierarchy of fidelity properties

Definition 4.1 (Internal full faithfulness)**.** A functor $F : C \to D$ of complete Segal objects is fully faithful if the canonical comparison

$$\alpha_F : C_1 \longrightarrow (C_0 \times C_0) \times_{D_0 \times D_0} D_1 \tag{12}$$

is an equivalence in $\mathcal{X}$.

Equivalently, for every generalized pair x, y in every context, the induced mapping-object comparison is an equivalence. This formulation is intrinsic and can be tested after an effective cover. A conservative functor instead reflects equivalences of arrows. A functor can be conservative while losing most morphisms and while identifying inequivalent objects.

Counterexample 4.2 (Conservativity does not reconstruct objects)**.** Let C be the discrete category on two objects and D the terminal category. The unique functor $C \to D$ is conservative: every arrow of C is already an identity. Yet the two objects of C have the same trace and are not equivalent in C. The functor is not fully faithful, since an empty off-diagonal mapping space is sent to a point.

Proposition 4.3 (Reconstruction by a fully faithful trace)**.** *If $F : C \to D$ is fully faithful, then*

$$\underline{\mathrm{Map}}_C(x, y) \simeq \underline{\mathrm{Map}}_D(Fx, Fy), \tag{13}$$

$$\underline{\mathrm{Eq}}_C(x, y) \simeq \underline{\mathrm{Eq}}_D(Fx, Fy). \tag{14}$$

In particular, Fx and Fy are equivalent precisely when x and y are equivalent, in the relevant internal sense. The functor identifies C with the full subcategory spanned by its essential image whenever that subcategory is formed.

Proof. The first assertion is the pullback of (12) along the pair (x, y). A fully faithful functor reflects equivalences: an inverse to Ff lifts to an arrow g by full faithfulness; the two inverse equations lift through the corresponding mapping-space equivalences. Thus the equivalences are exactly the invertible components on both sides, proving the second assertion. Factoring through the essential image gives a fully faithful and essentially surjective functor, hence an equivalence. □

The proof uses mapping-space lifting twice, including the coherence of the inverse equations. It does not require a canonical decoding algorithm on raw records. Essential image is also relative to the target category: if the trace target contains additional arrows between recorded objects, full faithfulness requires those arrows to correspond to operational transformations. A richer target can therefore make full faithfulness harder to satisfy.

4.2 Finite logs and counterfactual behavior

A log of observed executions can identify what occurred on the observed inputs. A skill may instead be a rule for all inputs, including inputs never encountered. Those are different information objects.

Proposition 4.4 (Finite-query nonidentifiability)**.** *Let A be infinite and let the candidate skills include all total functions $A \to \{0,1\}$. No deterministic observation procedure that makes finitely many adaptive input-output queries on every run can uniformly identify the extensional skill from those queries alone.*

Proof. Run the procedure against the constant-zero function f. It makes a finite set Q of queries before stopping. Choose $a_* \notin Q$ and define $g(a_*) = 1$, with $g(a) = 0$ for all other inputs. On g the procedure receives exactly the same responses in the same adaptive order, so it makes the same queries and stops with the same output. But $f \neq g$. Thus the output cannot correctly identify both extensional functions. □

The argument allows arbitrary postprocessing and does not rely on computational limitations. It is a finite-information obstruction. A restricted hypothesis class can avoid it: if candidate functions are known to be affine over a finite field, a suitable finite basis of inputs can identify them. The relevant assumption is then a known restricted class and an identifying experiment, rather than generic completeness of telemetry.

Proposition 4.5 (A decidable full abstraction cannot cover arbitrary program semantics)**.** *Suppose a total computable map c assigns every program a trace in a set with decidable equality, and assume $c(p) = c(q)$ exactly when p and q compute the same partial function. Then extensional program equivalence would be decidable. Such a map does not exist for a standard Turing-complete programming model.*

Proof. Compute $c(p)$ and $c(q)$ and compare them. Under the hypothesis this decides extensional equivalence. Fixing a program computing the nowhere-defined partial function would then decide a nontrivial extensional property of programs, contradicting Rice's theorem [15]. □

This does not prohibit a fully faithful mathematical representation in a target whose equivalence relation is itself undecidable. Nor does it prohibit a syntactic trace carrying

the original program. It rules out the stronger combination of total computable encoding, decidable canonical comparison, and exact extensional classification.

4.3 Representation does not supply execution

To execute a represented skill one needs more structure. Let A and B be input and output objects. An evaluator is a map or functor of a declared type, schematically

$$\mathrm{ev} : \mathcal{K} \times A \longrightarrow B. \tag{15}$$

A decoder on traces would be a map dec from an authorized part of $\mathcal{D}$ to a space of realizers, with a correctness comparison after evaluation. Full faithfulness of F gives neither (15) nor an efficient inverse on presentations. A cryptographic commitment may identify an artifact without revealing the artifact; a semantic embedding may represent a behavior whose efficient realization is unknown.

Even if an inverse exists on the essential image, it may be noncomputable in the chosen representation or require unbounded resources. A finite formal certificate can have an expensive verifier. A reconstructed operation can also be physically infeasible without tools, energy, or access rights. Statements about replacement of skilled labor therefore need assumptions about realization, transfer, control, and cost in addition to trace fidelity.

4.4 What ordinary event records usually guarantee

A practical trace often supplies a typed observation with provenance, timestamps, and a declared loss model. These properties can be valuable without satisfying definition 4.1. They support audit, diagnosis, estimation, and bounded identification. One should state the intended query class $\mathcal{Q}$ and ask whether the trace preserves those queries. If each $q \in \mathcal{Q}$ factors through T, then T is sufficient for that query family:

$$q = \bar{q} \circ T. \tag{16}$$

This is much weaker than preserving every mapping object of an entire operational category. It is often the appropriate engineering specification.

The strongest trace theorem in the paper will therefore be conditional. It tells a designer what follows *if* a high-fidelity representation exists and is locally certified. It does not establish from the presence of logs that the antecedent holds. The finite-query and computability obstructions explain why that distinction matters.

5 Local completeness and effective descent

5.1 Conservativity of pullback along a cover

The topos-theoretic heart of the observation argument is a locality principle. Let $u : U \to B$ be an effective epimorphism. Its Cech nerve has degree n given by the $(n+1)$-fold fiber product $U_n = U \times_B \cdots \times_B U$. The augmented simplicial object realizes to B. Universality of colimits implies that after any base change $A \to B$, the analogous augmented nerve realizes to A.

Lemma 5.1 (Local detection of equivalences)**.** *For an effective epimorphism $u : U \to B$, the pullback functor $u^* : \mathcal{X}_{/B} \to \mathcal{X}_{/U}$ is conservative.*

Proof. Let $f : A \to A'$ be a map over B whose pullback to U is an equivalence. In each simplicial degree n, its pullback to U_n is again an equivalence, since $U_n \to B$ factors through U and pullback preserves equivalences. Thus the map of augmented Cech resolutions of A and A' is a degreewise equivalence. Their geometric realizations are A and A', respectively. The induced map on realizations is an equivalence, and that map is f. □

The proof is not a claim that observations at finitely many sites determine an arbitrary global object. A covering map is a structural condition on the modeled context; a sample is a finite collection of observations. A dataset can leave entire contexts uncovered. Even a mathematically effective cover must be paired with the correct restriction maps and with actual local hypotheses.

Corollary 5.2 (A family of contexts)**.** *If $\{U_i \to B\}$ is an effective cover, then a map over B is an equivalence whenever all its restrictions to the U_i are equivalences.*

Proof. Pull back to $\coprod_i U_i$. In the slice over that coproduct the map is the coproduct of its component restrictions and is therefore an equivalence. Apply lemma 5.1. □

5.2 Full faithfulness is a local property

Theorem 5.3 (Descent detection of full faithfulness)**.** *Let $F : C \to D$ be an already defined internal functor over B. If $u : U \to B$ is an effective epimorphism and u^*F is fully faithful, then F is fully faithful.*

Proof. Form the comparison α_F from (12) in $\mathcal{X}_{/B}$. Pullback preserves the products and fiber product appearing in its target. Consequently $u^*\alpha_F$ is the full-faithfulness comparison for u^*F. The hypothesis says it is an equivalence. By lemma 5.1, α_F is an equivalence. This is precisely definition 4.1. □

Corollary 5.4 (Local trace fidelity)**.** *Suppose $T \simeq Fq$ is a global factorization as in proposition 3.2. If F is fully faithful over every member of an effective cover, then it reconstructs the internal mapping objects and equivalence objects globally.*

Proof. Apply theorem 5.3 to F, and then proposition 4.3. □

Only finite-limit data were needed to express full faithfulness, followed by local detection of an equivalence. Essential surjectivity is a different condition. Full faithfulness identifies the source with a full image, not necessarily with the entire trace category. A target can contain unrelated archived material or hypothetical traces without invalidating full faithfulness on the represented operational theory.

The theorem is also stable under restriction. If F is globally fully faithful, every pullback is fully faithful. Thus, for a fixed effective cover, the property is both preserved and detected locally. This gives a practical pattern: first define the functor and its semantics, then verify its mapping-object comparisons on contexts covering the intended domain.

5.3 Detecting a global functor versus constructing one

The previous theorem starts with a global F. A different problem begins with functors $F_i : C_i \to D_i$ on local contexts and asks for a global functor. Pairwise existence of equivalences between their restrictions is not enough; the equivalences must form descent data.

The slice categories satisfy an equivalence

$$\mathcal{X}_{/B} \simeq \lim_{[n]\in\Delta} \mathcal{X}_{/U_n}. \tag{17}$$

Limits of category-valued diagrams here are infinity-categorical limits, so the right side includes all coherence data. Degreewise descent gives descent for simplicial objects. The Segal comparisons use finite limits. The object of equivalences used for completeness is also compatible with base change: inverse arrows and their coherence can be specified and checked locally. Consequently both conditions descend by local detection of their comparison equivalences. This gives descent for internal categories, as in the internal categorical framework of [4].

Proposition 5.5 (Descent of compatible local categories and functors)**.** *Suppose local internal categories and functors are given as an object of the full totalization of the corresponding Cech diagram, with Segal and completeness conditions in every context. Then they arise from global internal categories and a global functor, uniquely up to the equivalence specified by descent. If the local functors are fully faithful, the descended functor is fully faithful.*

Proof. Equation (17) descends each simplicial degree and each structure map. The simplicial identities descend because the data form an object of the totalization, rather than unrelated

local choices. Segal and completeness comparisons are equivalences locally, hence globally. The same argument descends a natural map between the simplicial objects. Full faithfulness follows from theorem 5.3. □

This statement deliberately includes complete descent data in its antecedent. It does not prove that a particular collection of independently developed systems supplies those data. Finding or implementing them can be the principal integration task.

5.4 A concrete failure of pairwise compatibility

Counterexample 5.6 (An inconsistent triangular interface)**.** Let three contexts have nonempty triple overlap. On each context take a constant two-element set. Let the transition bijections on overlaps 1–2 and 2–3 be identities, and let the transition from 1 to 3 be the transposition. Each pair admits a bijection, but on the triple overlap

$$g_{13} \neq g_{23} g_{12}. \tag{18}$$

These chosen transitions do not define descent data for a set-valued object.

The example separates the existence of local objects from coherence of the chosen integration. Replacing the transitions can repair it; the original transitions cannot simply be called a global object. In a groupoid-valued setting one may replace equality by specified isomorphisms, but then those isomorphisms must satisfy the next coherence condition. Moving up one categorical level relocates the obligation rather than eliminating it.

5.5 Cech descent and hyperdescent

Effective epimorphic Cech descent, which is all that theorem 5.3 uses, holds in every infinity-topos. A claim about reconstruction from all finite Postnikov stages can require hypercompleteness or a separate convergence hypothesis. In a nonhypercomplete infinity-topos, an infinity-connected map need not be an equivalence. Thus one cannot replace a proof of equivalence by "all finite homotopy tests pass" without checking the ambient hypothesis.

The distinction has a useful interpretive counterpart. A system can pass every test in a growing finite family while the asserted limit behavior remains unproved. The mathematical statement about hypercompletion is not an empirical theorem about test suites, but it warns against using a limit conclusion without a limit hypothesis. Our local trace theorem concerns an equivalence after a genuine cover, not asymptotic recovery from incomplete tests.

6 Observational modalities and authorized knowledge

6.1 Reflective observation models

An accessible left-exact reflective localization consists of an adjunction

$$L : \mathcal{X} \rightleftarrows \mathcal{X}_L : i, \qquad L \dashv i, \tag{19}$$

with i fully faithful and L preserving finite limits. The unit $\eta_A : A \to iLA$ compares an object with its localized observation. Such a localization is a useful model for certain stable forms of visibility, but it is a restrictive model. An arbitrary redaction map, stochastic channel, or feature extractor need not be idempotent, reflective, or left exact.

Definition 6.1 (Completeness relative to relevant objects)**.** For a specified full subcategory $\mathcal{R} \subseteq \mathcal{X}$, the localization is complete on $\mathcal{R}$ when every unit η_A for $A \in \mathcal{R}$ is an equivalence. It is conservative on $\mathcal{R}$ when its restriction reflects equivalences between relevant objects.

Completeness on objects is stronger than merely conserving certain arrows. The choice of $\mathcal{R}$ is substantive. If it is only the terminal object, the condition is almost vacuous. If it includes all personal information, it is a much stronger and potentially unauthorized demand. No theorem below supplies a reason to maximize $\mathcal{R}$.

Proposition 6.2 (Local detection of modal completeness)**.** *Let $u : U \to B$ be effective epi. If the pullback $u^*\eta_A$ of the global unit is an equivalence, then η_A is an equivalence.*

Proof. Apply lemma 5.1 directly to the unit map. □

When one describes $u^*\eta_A$ as the unit of a locally defined observation modality, a compatibility between that local modality and the global one is needed. The elementary proof only concerns the pullback of a fixed map. This is another instance of the difference between detecting a given structure and constructing compatible local structures.

6.2 Policy as a separate fibration

A useful general model of authorization is a category $\mathcal{A}$ of authorized information operations with a forgetful functor

$$p : \mathcal{A} \longrightarrow \mathcal{D}. \tag{20}$$

A trace may exist in $\mathcal{D}$ without admitting a lift for a particular observer in $\mathcal{A}$. A map that exists after forgetting permissions is not automatically an authorized operation. If contexts change, $\mathcal{A}$ can be organized as a stack or a fibration whose fibers record purpose, role, expiry, or provenance restrictions.

This formulation follows the insight that information flow depends on context, rather than merely on possession of a bit string [21]. We do not claim that an arbitrary normative

system is automatically a stack. The stack condition is a model assumption: compatible local permissions can be glued, and the resulting permission has the specified global meaning.

Example 6.3 (A query without a joint authorization)**.** Two services each possess a table keyed by a person identifier. Service 1 may answer a delivery query; service 2 may answer a billing query. The underlying tables can be joined on the identifier. If neither authorization permits a combined behavioral profile, that join has no permitted lift of the required type. Mathematical existence of a set-theoretic join does not extend either authorization.

6.3 Losslessness and privacy can conflict

For a deterministic observation $r : A \to O$ in sets, suppose two states a, a' differ only in a protected field and $r(a) = r(a')$. Then r cannot be injective on any relevant class containing both states as distinct elements. If operational equivalence already identifies a and a', however, the observation can remain complete for that coarser operational theory. Thus privacy and completeness are not absolutely incompatible; compatibility depends on which distinctions the task legitimately requires.

A probabilistic privacy mechanism gives a sharper obstruction later. The appropriate design question is not whether every hidden distinction can be recovered. It is whether permitted task distinctions are preserved while protected distinctions remain appropriately obscured. The category of allowed queries, rather than the totality of raw data, often gives the right specification.

7 Colimit integration and data migration

7.1 Integration is relative to a diagram

Let I be small and $D : I \to \mathcal{X}$ a diagram. Write $D_* := \operatorname{colim}_I D$ and $\iota_i : D_i \to D_*$ for the universal cocone. The diagram includes both the data objects and the interfaces along which they are to be identified. Supplying just the objects does not specify a reconciliation of identities, schemas, or meanings.

Theorem 7.1 (Universal compatible integration)**.** *For every $Y \in \mathcal{X}$ there is a natural equivalence*

$$\operatorname{Map}_{\mathcal{X}}(D_*, Y) \simeq \lim_{i \in I^{\mathrm{op}}} \operatorname{Map}_{\mathcal{X}}(D_i, Y). \tag{21}$$

The right side is the space of homotopy-coherent cocones from D to Y.

Proof. Equation (21) is the defining universal property of the colimit. Naturality follows from composition in the target and from the universal cocone. In particular, its homotopy fiber over any compatible cocone is contractible, so the extension is unique with its specified compatibility data up to contractible choice. □

This theorem is a uniqueness theorem for a *compatible* transformation. It does not say every collection of local transformations is compatible. It does not say that their global implementation uses less memory or compute. Nor does it say the colimit maps are embeddings: colimits can identify records. Each of those additional assertions requires a separate property of the diagram or implementation.

7.2 Coproducts, pushouts, and spurious identity merges

If I is the discrete set of two objects, the integrated object is $A \amalg B$. It retains a tag identifying the source. There is no cross-source deduplication. If an interface K maps to both A and B, a pushout

$$A \amalg_K B \tag{22}$$

identifies the images prescribed by that interface. Its maps to Y are pairs of maps from A and B with a specified compatibility on K.

Example 7.2 (A warehouse reconciliation)**.** Take sets $A = \{a_1, a_2\}$ and $B = \{b_1, b_2\}$. Let $K = \{k\}$ map to a_2 and b_1. The pushout in sets has three elements: $[a_1]$, $[a_2 = b_1]$, and $[b_2]$. A pair of price assignments extends precisely when the prices of a_2 and b_1 agree. If the interface linked the wrong records, the universal property would faithfully implement the wrong identification.

For the corresponding diagram of discrete spaces the homotopy pushout is the graph joining a_2 to b_1, with the other two isolated vertices, hence is equivalent to the same three-point space. More complicated diagrams of sets can produce higher homotopy when colimits are taken in spaces. It is therefore necessary to state the target category in which integration occurs. A quotient in sets and a homotopy colimit in spaces need not contain the same information.

7.3 Information integration does not identify owners

Proposition 7.3 (A separation countermodel)**.** *The existence of a colimit of $D : I \to \mathcal{X}$ does not imply $I \simeq *$, nor does it imply that the suppliers of the diagram have a common owner.*

Proof. In $\mathcal{S}$, take I to be a discrete two-object category and take both data objects to be points. The colimit is the discrete two-point space. The indexing category remains nonterminal. More generally, arbitrary fixed ownership labels can be assigned to the suppliers without changing this colimit calculation. Since the ownership labels do not enter its universal property, they cannot be determined by that property. □

Even a terminal integrated object does not force a terminal indexing category. Take an index with a terminal vertex whose data is a point; its colimit is that point although the index can contain many other non-equivalent objects. One should distinguish a terminal *value of a diagram* from a terminal *organization of the index*.

7.4 Left Kan extension along a coarsening

Let $q : I \to J$ be a functor of indexing categories. Since $\mathcal{X}$ is cocomplete, restriction $q^* : \mathrm{Fun}(J, \mathcal{X}) \to \mathrm{Fun}(I, \mathcal{X})$ has a left adjoint $q_!$. Its pointwise value is

$$(q_! D)(j) \simeq \operatorname*{colim}_{(i,\, q(i)\to j)\in(q\downarrow j)} D(i). \tag{23}$$

The comma category records both contributors and the ways they map into the consolidated context. In a discrete grouping of firms, the indexing category over a block is discrete and the value is a coproduct. A genuine reconciliation requires nontrivial interfaces or a later quotient; grouping a list does not itself remove duplicate identities.

Proposition 7.4 (Total integration is compatible with Kan consolidation)**.** *There is a natural equivalence*

$$\operatorname*{colim}_J q_! D \simeq \operatorname*{colim}_I D. \tag{24}$$

Proof. For every Y, use the adjunction $q_! \dashv q^*$:

$$\begin{aligned} \mathrm{Map}(\operatorname*{colim}_J q_! D, Y) &\simeq \mathrm{Map}_{\mathrm{Fun}(J,\mathcal{X})}(q_! D, \mathrm{const}_J\, Y) \\ &\simeq \mathrm{Map}_{\mathrm{Fun}(I,\mathcal{X})}(D, \mathrm{const}_I\, Y) \\ &\simeq \mathrm{Map}(\operatorname*{colim}_I D, Y). \end{aligned}$$

The natural equivalence of representable functors identifies the objects. □

This result says that consolidating the indexing description by its universal left extension does not change the final universal receiving object. It strengthens the separation argument: different organizational presentations can have the same integrated information semantics. Their costs, permissions, latency, and control rights can nevertheless differ.

7.5 Schema migration and its scope

In categorical database theory a schema is a small category and an instance is a set-valued functor. A schema map induces restriction and left and right Kan extensions, giving the familiar adjoint pattern

$$\Sigma_q \dashv \Delta_q \dashv \Pi_q. \tag{25}$$

Spivak develops this framework as functorial data migration [7]. The present use of Kan extension is an application of that general pattern to context-dependent and higher information objects. The analogy is precise at the level of adjunctions; a higher-valued instance retains additional equivalences and coherence.

Migration still needs a semantic interpretation of schema arrows. A functor can preserve a badly chosen equation. If two columns have distinct units, a map identifying them is not

justified merely because a category with that map can be written. Semantic constraints, provenance, and permissions should therefore be included in the schema or in a structured instance category, rather than reconstructed from a bare colimit after information has been lost.

8 Governance, interfaces, and the cost of coherence

8.1 Authorized colimits require an authorized category

Suppose $p : \mathcal{A} \to \mathcal{D}$ forgets authorization. A diagram $\widetilde{D} : I \to \mathcal{A}$ may have a colimit in $\mathcal{A}$, or it may not. Even if it has one, its image under p need not be a colimit in $\mathcal{D}$. Conversely, a colimit in $\mathcal{D}$ may have no lift to $\mathcal{A}$. The three claims are distinct.

Proposition 8.1 (An adequate lifting condition)**.** *If $\widetilde{D}$ has a colimit A_* in $\mathcal{A}$ and p preserves this colimit, then $p(A_*)$ is an authorized realization of the underlying integrated object. If p creates the relevant colimits, an underlying colimit with the required compatible lifted cocone determines such a realization.*

Proof. The first claim is the definition of preservation. For the second, creation supplies a colimit in $\mathcal{A}$ with the given image and specifies its universal property. Both statements depend on the lifted cocone; no lift follows just from the existence of its image. □

The proposition is intentionally formal. It identifies the exact categorical condition a permission-aware integration must satisfy. It does not assert that real permission systems meet it. Time-limited grants, incompatible purposes, or jurisdictional requirements can prevent a common cocone. Institutional constraints are therefore not merely scalar costs appended to an otherwise unrestricted integration.

8.2 Local data objects versus globally chosen identifiers

A nontrivial covering space over a circle can be locally trivial while admitting no global section. For instance, the double cover $S^1 \to S^1$, $z \mapsto z^2$, is locally a two-element choice but has no continuous global choice of a square root. It is a valid global object; what fails is a global section selecting one branch.

This example prevents a subtle error in organizational language. Effective descent can glue local identifier systems into a global object without supplying a globally chosen identifier or credential. The existence of the descended object, the existence of a section, and the existence of a canonical section are separate assertions. An integration can retain a torsor of choices rather than resolve it to a single globally preferred name.

The same issue appears with cryptographic keys and local accounts. Local key material can be compatible under transformations while no global shared key should exist. One can integrate

attestations or proof objects instead of consolidating credentials. The appropriate gluing problem is then typed by the permission model, not by an assumption that all information should become a single plaintext database.

8.3 Composition at boundaries

Open systems can be represented by cospans whose feet are interfaces and whose apex is an implementation. Composition by pushout identifies shared interfaces. Decorated cospans provide a general mathematical framework for adding internal structure to such apices [8]. This supplies a useful model of interfirm interoperability: two owners can compose systems along a negotiated interface while retaining separate apices and separate decision rights.

The compositional advantage is not the elimination of every boundary. It is the possibility of a stable boundary specification. An interface can reduce the need to share internal information. If a query factors through the interface, the receiving system need not reconstruct the entire source system. The existence of good abstractions can therefore reduce the economic incentive to merge even when information processing becomes more capable.

8.4 Coherence is not economically free

Let $K_{\mathrm{coh}}(D)$ be an implementation-dependent cost of supplying and maintaining the comparisons in a diagram. A universal property gives no estimate for this number. A diagram with many overlaps can have expensive identity resolution, schema evolution, and policy reconciliation. One can model a partition-dependent cost

$$K(P) = K_{\mathrm{compute}}(P) + K_{\mathrm{coh}}(P) + K_{\mathrm{govern}}(P) + K_{\mathrm{failure}}(P). \tag{26}$$

The terms need not be additive in reality; (26) is a possible accounting decomposition. Its important feature is that a colimit theorem places no monotonicity constraint on them.

For example, a common schema may remove repeated translation at query time but require costly coordination whenever one participant changes its schema. A centralized model can amortize training while increasing correlated failure. These are possible mechanisms, not measured claims here. Their signs and magnitudes belong to the empirical specification of W, not to the universal property of a pushout.

9 Probabilistic observation and decision sufficiency

9.1 Channels as a different observation model

Semantic full faithfulness is exact. Empirical records are often noisy or partial. For finite sets X and Y, an observation channel is a stochastic matrix $K(y \mid x)$ with nonnegative entries

and row sums one. Composition is matrix multiplication:

$$(LK)(z \mid x) = \sum_{y} L(z \mid y)K(y \mid x). \tag{27}$$

The finite setting avoids measure-theoretic assumptions about regular conditional probabilities. Markov categories provide a broader categorical framework for channels, conditional independence, and sufficient statistics [11]. We use explicit matrices so that every numerical assertion can be checked directly.

Given a prior $\pi(x)$, an action set A, and loss $\ell(x, a)$, the Bayes risk of a channel is

$$R(K) = \inf_{\delta} \sum_{x,y,a} \pi(x)K(y \mid x)\delta(a \mid y)\ell(x, a), \tag{28}$$

where δ ranges over stochastic decision rules. The relevant fidelity is now preservation of achievable decisions for specified losses, not equivalence of all operational mapping spaces.

Proposition 9.1 (Garbling cannot improve optimal information)**.** *If $K' = LK$, then $R(K) \leq R(K')$ for every finite decision problem.*

Proof. Any decision rule δ' based on K' can be simulated from K by the composite rule $\delta' L$. This composite has exactly the same joint distribution of state and action as using K' followed by δ'. Thus the feasible rules for K include simulations of all rules for K', so the infimum risk cannot be larger. □

This is the easy direction of Blackwell's comparison theory [12]; no converse for general measurable spaces is asserted. It supplies a rigorous sense in which a richer signal can support more decisions. It still does not imply greater net value once acquisition, computation, privacy, or coordination costs are included.

9.2 Identifiability is not perfect one-shot recovery

Example 9.2 (A noisy binary trace)**.** Let $X = Y = \{0, 1\}$ and let K be the binary symmetric channel with crossover probability $p \in (0, 1/2)$. The two conditional distributions are distinct, so the latent bit is statistically identifiable from arbitrarily many conditionally independent repetitions. But a single observation has positive probability of error under each state, and no deterministic one-shot decoder recovers the bit almost surely.

The matrix of this channel is invertible as a real linear map for $p \neq 1/2$. Its inverse has negative entries when $0 < p < 1/2$, so that inverse is not a stochastic channel. Algebraic invertibility therefore differs from physical or probabilistic reversibility. This is a simple illustration of why exact reconstruction must be formulated in the correct category.

9.3 Information value need not be superadditive

For a fixed target random variable X, let an information valuation be $V(Y) = I(X;Y)$ under a specified joint law. The chain rule gives

$$I(X;Y,Z) = I(X;Y) + I(X;Z \mid Y). \tag{29}$$

Hence the synergy relative to the sum of separate mutual informations is

$$\sigma = I(X;Y,Z) - I(X;Y) - I(X;Z) = I(X;Z \mid Y) - I(X;Z). \tag{30}$$

Its sign is not fixed.

Example 9.3 (Redundancy and synergy)**.** If $Y = Z = X$ is a fair bit, then $I(X;Y) = I(X;Z) = 1$ bit and $I(X;Y,Z) = 1$ bit, so $\sigma = -1$ bit. If Y, Z are independent fair bits and $X = Y \oplus Z$, both separate informations vanish while the joint information is one bit, so $\sigma = 1$ bit.

Thus a globally lax superadditive value assignment is not a generic consequence of "more data." It selects a particular valuation regime. The tensor must also specify the joint law. Two marginal channels do not determine their joint dependence. Taking them to be conditionally independent is an additional modeling choice, and it can change the sign of the synergy. The information identities used here are standard [13]; the examples are constructed to expose the assumption.

9.4 A privacy obstruction to exact recovery

For adjacent databases $d \sim d'$, an ε-differentially private mechanism M satisfies

$$\Pr[M(d) \in A] \leq e^{\varepsilon} \Pr[M(d') \in A] \tag{31}$$

for every measurable event A, and symmetrically since adjacency is symmetric. This standard definition bounds the effect of one protected change [14].

Proposition 9.4 (Finite privacy and exact distinction)**.** *Let $d \sim d'$ be distinct adjacent databases and let $\varepsilon < \infty$. No ε-differentially private mechanism admits a decoder recovering the exact input almost surely under both d and d'.*

Proof. If a decoder g recovered both inputs, take $A = \{y : g(y) = d\}$. Then $\Pr[M(d) \in A] = 1$ and $\Pr[M(d') \in A] = 0$. Equation (31) would give $1 \leq e^{\varepsilon} \cdot 0 = 0$, a contradiction. □

The obstruction concerns distinctions that adjacency protects. A task quotient that identifies d and d' can still be exactly recoverable. Approximate privacy and approximate decoding require quantitative variants, one of which is derived in the appendix. Privacy is

therefore compatible with task-relative sufficiency, but not with universal exact exposure of every protected difference.

9.5 From semantic theorems to empirical tests

A realistic fidelity claim should declare the latent state or hypothesis class, the channel, the decision family, and the observation horizon. One can then estimate risks, calibrate uncertainty, and test distribution shift. These activities do not establish internal full faithfulness of an unrestricted operational category. Conversely, a formal fully faithful functor says nothing about whether a noisy sensor meets the model's assumptions.

The exact and statistical layers can coexist. A probabilistic program can be represented exactly as a mathematical object, while its finite sampled outputs fail to identify that object. A trace category may contain the program and its likelihood model rather than merely a finite sample. Naming which of these is being observed prevents an ideal semantic theorem from being misreported as a theorem about ordinary telemetry.

10 Symmetric-monoidal resources and the choice of tensor

10.1 Resource theories and operational interpretation

A symmetric-monoidal category $(\mathcal{R}, \otimes, \mathbf{1})$ models jointly available resources and their allowed transformations. It can be internal to $\mathcal{X}$, or external when the economic analysis is performed in ordinary sets. The general resource-theoretic viewpoint is established in [10, 9]. Higher monoidal coherence can be formalized using infinity-operads [5]; our scalar results use only invariance under equivalence and associativity, unit, and symmetry.

The tensor is part of the model. For physical assets it may mean disjoint jointly controlled bundles. For signals it may mean a specified joint experiment. For information already accessible to an observer it may mean the join of information fields. Those tensors have different algebraic properties. A copied file can be a new physical storage token without adding new information. A license can be a resource even if the underlying bits are identical. A permission can be nontransferable while a record is technically copyable.

Let E be a finite set of labeled elementary capabilities and assign a resource r_e to each $e \in E$. For a subset $A \subseteq E$, define

$$r_A := \bigotimes_{e \in A} r_e, \qquad r_\varnothing = \mathbf{1}. \tag{32}$$

An indexed symmetric-monoidal product gives coherent comparisons

$$r_{A \cup B} \simeq r_A \otimes r_B \quad (A \cap B = \varnothing). \tag{33}$$

The restriction to disjoint labeled subsets avoids silently counting a single exclusive asset twice. Equivalences under reordering do not erase the labels of distinct identical assets. If labels should be identified because they refer to the same information, a different tensor or quotient must be specified.

10.2 Value and cost as comparison structures

Write $\mathbb{R}_{\leq}$ for the ordered symmetric-monoidal category with objects real numbers, an arrow $a \to b$ when $a \leq b$, tensor addition, and unit zero. A scalar valuation on the maximal subgroupoid $\mathcal{R}^{\simeq}$ assigns the same number to equivalent resources. A lax symmetric-monoidal structure on V supplies

$$V(x) + V(y) \leq V(x \otimes y), \qquad 0 \leq V(\mathbf{1}). \tag{34}$$

An oplax structure on C supplies

$$C(x \otimes y) \leq C(x) + C(y), \qquad C(\mathbf{1}) \leq 0. \tag{35}$$

We normally impose the explicit normalization

$$V(\mathbf{1}) = C(\mathbf{1}) = 0. \tag{36}$$

For nonnegative cost, the oplax unit inequality already forces $C(\mathbf{1}) = 0$. The normalization makes the empty bundle economically neutral. Unit conditions should not be omitted when presenting lax or oplax monoidal data.

Because the target is a poset, all parallel comparison diagrams commute whenever their arrows exist. Thus the inequalities encode the entire scalar coherence once equivalence invariance and the unit are accounted for. The category-theoretic language organizes the assumptions; it does not empirically establish their signs. Restricting to the core also deliberately avoids assuming anything about monotonicity under arbitrary resource conversions.

Set

$$N(x) := V(x) - C(x), \tag{37}$$
$$\sigma_V(x, y) := V(x \otimes y) - V(x) - V(y), \tag{38}$$
$$\sigma_C(x, y) := C(x) + C(y) - C(x \otimes y), \tag{39}$$
$$\Gamma(x, y) := N(x \otimes y) - N(x) - N(y). \tag{40}$$

Direct expansion gives

$$\Gamma(x, y) = \sigma_V(x, y) + \sigma_C(x, y). \tag{41}$$

The two monoidal inequalities imply $\Gamma \geq 0$. Positive net gain only requires the sum to be

positive; either summand can be negative in a more general model. Hence separate lax-value and oplax-cost assumptions are convenient sufficient conditions, stronger than the merger criterion itself.

Proposition 10.1 (Finite-bundle comparisons)**.** *Under (34)–(36), for every finite family* $(x_i)_{i=1}^m$*,*

$$\sum_{i=1}^{m} V(x_i) \le V(\bigotimes_i x_i), \qquad C(\bigotimes_i x_i) \le \sum_{i=1}^{m} C(x_i). \tag{42}$$

Proof. For $m = 0$ use normalization; for $m = 1$ use equality. The step from m to $m+1$ follows by applying the binary comparison to $(\bigotimes_{i\le m} x_i, x_{m+1})$ and combining with the inductive inequality. Coherence identifies the resulting tensor with the indexed product, independently of the chosen parenthesization. □

10.3 The idempotence obstruction

Nonrivalry of data is central to the economics of information [19]. It does not imply that a globally superadditive nonnegative scalar value is available for every way of combining information.

Theorem 10.2 (Idempotence and global superadditivity)**.** *Suppose* $x \otimes x \simeq x$*, V is invariant under equivalence, and V satisfies the lax inequality on (x, x). Then $V(x) \le 0$. If $V(x) \ge 0$, then $V(x) = 0$. In particular, a nonnegative lax valuation on a tensor-idempotent resource category is identically zero.*

Proof. Equivalence invariance and laxity give $2V(x) \le V(x \otimes x) = V(x)$. Subtracting $V(x)$ yields $V(x) \le 0$. Nonnegativity gives equality. Apply the argument objectwise for the last assertion. □

The theorem is elementary but consequential. If the tensor is the union of already available datasets, then combining a dataset with itself adds nothing, so idempotence is natural. A positive information value cannot obey the global lax inequality for that tensor. If the tensor instead means two separately controlled licenses, two compute allocations, or two physically distinct storage devices, idempotence need not hold and the obstruction need not apply.

Corollary 10.3 (Free copying and monotones)**.** *Suppose arrows are free conversions and V is a monotone in the convention $x \to y \Rightarrow V(x) \ge V(y)$. If there is a free copying arrow $x \to x \otimes x$ and V is lax as in (34), then $V(x) \le 0$.*

Proof. Monotonicity gives $V(x) \ge V(x \otimes x)$, while laxity gives $V(x \otimes x) \ge 2V(x)$. Combine the inequalities. □

Copyability as a conversion is weaker than tensor idempotence as an equivalence, but the additional monotonicity recovers the same obstruction. A valuation defined only on the core avoids this particular argument because it says nothing about the copying arrow. That mathematical escape also means the valuation is not yet a monotone for the full resource theory. Both choices are legitimate if their scope is stated.

10.4 A repair using labeled integration opportunities

For the partition analysis it is enough to define a set function $n(A) = N(r_A)$ on disjoint labeled bundles. One can assume

$$n(A \cup B) > n(A) + n(B) \tag{43}$$

only for specified disjoint coalitions. This does not assert a globally lax valuation on every repeated resource. The labeled model can describe complementarities between different tasks or permissions without assigning extra information value to duplicate copies of the same dataset.

Another repair is to work with a submodular or decision-based information value and a separate physical cost functional. A third is to use a family of task-specific valuations rather than one scalar. These repairs answer different questions. The correct choice depends on what an integration is supposed to make jointly available, and on whose objective is being measured.

10.5 Monoidal trace representations

If $\mathcal{O}$, $\mathcal{K}$, and $\mathcal{D}$ have compatible symmetric-monoidal structures, and S is stable under tensoring, a suitable monoidal localization can induce a monoidal factorization. We assume its universal property rather than infer it from the bare existence of an underlying localization. For a strong symmetric-monoidal fully faithful functor F,

$$F(k \otimes \ell) \simeq F(k) \otimes F(\ell) \tag{44}$$

and all mapping objects are preserved and reflected. Strong monoidality reconstructs parallel composition in the image. A merely lax monoidal trace gives a comparison arrow, not generally an equivalence, so it cannot be used in (44) without strengthening the hypothesis.

If $J : \mathcal{D} \to \mathcal{R}$ maps traces to resources, then JF assigns a resource to a represented operation. The interpretation of J must include any required rights or acquisition costs. It is not a theorem that a record automatically becomes a freely reusable productive asset. Monoidal structure explains how a declared resource assignment composes once that assignment is justified.

11 Partitions, merger gains, and exact consolidation criteria

11.1 The finite organizational space

Let E be a nonempty finite set with m elements. A partition P is a set of nonempty disjoint blocks whose union is E. Order partitions by refinement: $P \leq Q$ means every block of P lies in a block of Q. The bottom element is the singleton partition and the top is $\top = \{E\}$. The poset has Bell number B_m elements.

Assume first that net surplus is block-separable:

$$W(P) = \sum_{A \in P} n(A), \qquad n(\varnothing) = 0. \tag{45}$$

This is a substantial assumption. It excludes cross-firm externalities and partition-dependent costs not already included in the coalition values. We relax it in the next section.

If A, B are distinct blocks, define the binary merger

$$P' = (P \setminus \{A, B\}) \cup \{A \cup B\}. \tag{46}$$

Write $\gamma(A, B) = n(A \cup B) - n(A) - n(B)$.

Lemma 11.1 (Cancellation formula)**.** *For a binary merger in the separable model,*

$$W(P') - W(P) = \gamma(A, B). \tag{47}$$

If $n(A) = N(r_A)$, then $\gamma(A, B) = \Gamma(r_A, r_B)$.

Proof. All blocks except A and B occur in both sums in (45) and cancel. The remaining difference is the definition of γ. The tensor comparison for disjoint bundles and equivalence invariance of N give the last assertion. □

Theorem 11.2 (Strict all-merger sufficiency)**.** *If $\gamma(A, B) > 0$ for every pair of disjoint nonempty subsets of E, then $\top$ is the unique maximizer of W. If all these gains are nonnegative, then $\top$ is a maximizer, possibly among others.*

Proof. Any partition with $k > 1$ blocks can be coarsened to $\top$ by $k - 1$ binary mergers. At each step lemma 11.1 gives a strictly positive, or respectively nonnegative, difference. Summing the differences gives $W(\top) > W(P)$ in the strict case and $W(\top) \geq W(P)$ in the weak case. Since P was arbitrary, the conclusions follow. □

This is a sufficient condition, not a necessary one. A unique global optimum can coexist with negative individual merger gains. One must not call all-merger strict superadditivity "the exact assumption required" for top optimality. The exact condition is simply $W(\top) > W(P)$ for every other partition, and many less restrictive structural conditions can imply it.

Counterexample 11.3 (A top optimum behind negative pair mergers)**.** Take $E = \{1, 2, 3\}$, singleton values zero, all pair values -1, and $n(E) = 1$. The singleton partition has value 0, every two-block partition has value -1, and $\top$ has value 1. Thus $\top$ is the unique global optimum, but no beneficial binary merger is available from the singleton partition.

The example separates global optimality from improving local reachability. A simultaneous three-way merger, an intermediate subsidy, or a temporary loss can cross the barrier. A dynamic model permitting only strictly improving binary moves cannot do so.

11.2 An exact monotone-path condition

Form a directed graph with vertices the partitions and edges the admissible binary mergers having positive gain. A nonempty finite graph ordered by strict coarsening has no directed cycles.

Proposition 11.4 (Reachability criterion)**.** *Every partition has a strictly improving merger path to $\top$ if and only if every partition other than $\top$ has an outgoing positive admissible merger. Under this condition $\top$ is the unique maximizer of W.*

Proof. The forward implication follows by taking the first edge of a path. Conversely, begin at any non-top partition and repeatedly choose a positive outgoing merger. Each step reduces the number of blocks, so the procedure must terminate. The hypothesis excludes termination anywhere other than $\top$. Summing the positive changes along the path proves the value comparison. □

This criterion is weaker than positivity of every merger: it requires only one profitable direction from every non-top partition. It is still stronger than top optimality, as counterexample 11.3 shows. It also specifies feasibility. An attractive but prohibited merger does not supply an outgoing edge.

11.3 Functoriality and what it records

If every binary merger has nonnegative gain, W is an order-preserving functor

$$W : \mathrm{Part}(E) \longrightarrow \mathbb{R}_{\leq}. \tag{48}$$

The content is monotonicity under refinement. In a poset, the functorial coherence adds no numerical restriction beyond the inequalities. Strict positivity on binary mergers makes it strictly increasing on every nonidentity coarsening because every such coarsening decomposes into binary mergers.

If only some moves are admissible, use a specified subcategory or directed graph. A subcategory may omit arrows even between feasible objects, so it is important to distinguish

a family of feasible partitions from a family of feasible transitions. A physical merger process may be irreversible, delayed, or dependent on consents not encoded by the order relation alone.

12 Constraints, externalities, and robustness

12.1 Admissibility need not have a terminal object

Let $\mathcal{P}_{\mathrm{adm}}$ be a finite subcategory of the partition poset. If it has a terminal object P_*, every admissible partition has a permitted arrow to P_*. This is a strong assumption about institutional compatibility.

Theorem 12.1 (Constrained terminal optimum)**.** *Suppose $\mathcal{P}_{\mathrm{adm}}$ has terminal object P_* and $W(Q) > W(P)$ on every nonidentity arrow $P \to Q$ of that category. Then P_* uniquely maximizes W on its objects.*

Proof. For any $P \neq P_*$ terminality gives a nonidentity arrow $P \to P_*$. Its strict inequality is exactly $W(P) < W(P_*)$. □

If the category is generated by admissible binary mergers, it is enough to check positivity on those generators. If it has several maximal objects instead, increasing trajectories may end at different sinks with different values. Finite termination alone gives neither a canonical sink nor a global optimum.

Example 12.2 (Two incompatible endpoints)**.** For $E = \{1, 2, 3\}$, permit only the singleton partition and the partitions $12|3$ and $1|23$, with the two corresponding outgoing mergers. Let their values be $0, 2, 1$. Both sinks are merger-stable, but only $12|3$ is globally optimal. There is no terminal object in this transition category.

It is therefore inaccurate to replace a missing terminal object by "the" merger-stable configuration. Stability is a local property relative to permitted moves. It can be nonunique and suboptimal. The economic analysis must state whether it studies global optimization, Nash-type stability, local improvement, or a particular governance process.

12.2 Partition-dependent costs and externalities

In a broader model let

$$W(P) = \sum_{A \in P} n(A) - H(P), \tag{49}$$

where H records costs or harms depending on the whole partition. Then the exact merger gain is

$$\Delta W = \gamma(A, B) - \big(H(P') - H(P)\big). \tag{50}$$

The same two blocks can have different net merger gains in different surroundings. The scalar $\Gamma(r_A, r_B)$ no longer captures the entire decision unless the surrounding dependence is included in the resource state.

This is especially important when distinguishing private surplus from a broader welfare objective. Write

$$W_{\text{social}}(P) = W_{\text{private}}(P) - H_{\text{external}}(P). \tag{51}$$

A merger can increase the first quantity and decrease the second. Category theory cannot choose whose utility defines V. The theory of organizational boundaries has long examined transaction costs, contractual incompleteness, and ownership incentives [16, 17]. The present model records their possible consequences through W and feasibility; it does not replace those theories with a universal superadditivity axiom.

12.3 A quantitative robustness margin

Economic values are often estimated. A strict theorem with no margin can be fragile under small errors. Assume a nominal coalition function $\hat{n}$ and a bound

$$|n(A) - \hat{n}(A)| \le \varepsilon \quad (A \neq \varnothing). \tag{52}$$

Then the estimated binary gain $\hat{\gamma}$ differs from the true gain by at most 3ε.

Proposition 12.3 (Robust merger positivity)**.** *Under (52), if $\hat{\gamma}(A, B) > 3\varepsilon$ for a merger, its true gain is positive. If this holds for every binary merger, $\top$ remains the unique maximizer of the true separable objective.*

Proof. The gain contains three coalition values with coefficients $1, -1, -1$, so the triangle inequality gives $|\gamma - \hat{\gamma}| \le 3\varepsilon$. Positivity follows from the margin, and theorem 11.2 gives the global conclusion. □

The bound is conservative and transparent. If errors are correlated or coalition-specific intervals are available, one can use the corresponding interval sum. No independence assumption is needed for this deterministic result.

For the general model (49), if $|H(P) - \hat{H}(P)| \le \delta$ as well, the binary margin becomes $3\varepsilon + 2\delta$. Alternatively, a uniform bound $|W(P) - \hat{W}(P)| \le \eta$ preserves a nominal unique optimum whose gap exceeds 2η. These are optimization stability statements; they do not certify that the error bounds are empirically valid.

12.4 Acquisition costs and transfers

The objective W should distinguish production surplus from transfers between owners. A purchase price changes the distribution of surplus and can affect feasibility under budget

constraints, but it is not necessarily a resource cost to the aggregate system. Bargaining failure can block a merger with positive aggregate gain. Conversely, a private acquisition incentive can come from redistributing rents rather than improving production.

A complete equilibrium theory would specify agents, endowments, beliefs, contracts, and strategic behavior. Our finite partition theorems do not contain those ingredients. "Uniquely maximizes W" is not a statement that a decentralized market necessarily reaches that partition. A dynamic or game-theoretic realization must be added and checked separately.

13 Continuous complementarity and finite integration gains

13.1 Cross-partials measure an incremental comparison

Let $F : \mathbb{R}^d_{\geq 0} \to \mathbb{R}$ be C^2 on an open neighborhood of the rectangle under consideration. For distinct coordinates i, j, increments $s, t \geq 0$, and a base point x, define

$$\Delta_{ij} F(x; s, t) = F(x + se_i + te_j) - F(x + se_i) - F(x + te_j) + F(x). \tag{53}$$

Repeated use of the fundamental theorem of calculus yields

$$\Delta_{ij} F(x; s, t) = \int_0^s \int_0^t \partial_i \partial_j F(x + ue_i + ve_j)\, dv\, du. \tag{54}$$

Thus nonnegative cross-partials on the entire rectangle imply nonnegative incremental complementarity on that rectangle. A positive cross-partial at a single point supports only a local conclusion, by continuity on a sufficiently small neighborhood.

This is the smooth relation to increasing differences and supermodularity studied in economic models of complementary decisions [18]. It compares two coordinate increments around a common baseline. The global inequality $F(x + y) \geq F(x) + F(y)$ is a different statement. It compares two bundles that may share coordinates and it depends on normalization, own-coordinate curvature, and fixed terms.

Counterexample 13.1 (Positive cross-partials without global superadditivity)**.** On $\mathbb{R}^2_{\geq 0}$ let

$$F(x, y) = \sqrt{1 + x} - 1 + \sqrt{1 + y} - 1 + \alpha xy, \qquad \alpha > 0. \tag{55}$$

The function is smooth on a neighborhood of the nonnegative quadrant, $F(0, 0) = 0$, and $\partial_x \partial_y F = \alpha > 0$. But

$$F(2, 0) = \sqrt{3} - 1 < 2(\sqrt{2} - 1) = 2F(1, 0). \tag{56}$$

Hence it fails global superadditivity for two bundles in the first coordinate.

13.2 A sufficient Hessian condition

There is a stronger condition that does imply global superadditivity.

Proposition 13.2 (Nonnegative Hessian entries and normalization)**.** *Let F be C^2 on a neighborhood of $\mathbb{R}^d_{\geq 0}$, with $F(0) = 0$. If every Hessian entry $\partial_i \partial_j F$ is nonnegative on the nonnegative orthant, then*

$$F(x + y) \geq F(x) + F(y) \quad (x, y \geq 0). \tag{57}$$

Proof. Define $G(s, t) = F(sx + ty)$ on $[0, 1]^2$. Its mixed derivative is

$$\partial_s \partial_t G(s, t) = \sum_{i,j} x_i y_j \partial_i \partial_j F(sx + ty) \geq 0. \tag{58}$$

Integrating over the square gives $F(x+y) - F(x) - F(y) + F(0) \geq 0$. Apply the normalization. □

If x and y have disjoint coordinate supports, only mixed entries across those supports appear, so nonnegative own-coordinate second derivatives are unnecessary for that restricted comparison. This is a useful bridge to disjoint labeled capability bundles. Strict gain requires the integral to be positive, which follows if an involved Hessian entry is positive on a set of positive area with positive coefficients.

13.3 Scale, scope, and ownership

The function F can measure output from a joint resource vector, but the same vector may be attainable through contracts or shared services without common ownership. To infer a merger advantage one must compare the attainable sets and costs under each ownership arrangement. A technology with strong complementary inputs can be served by a competitive supplier network if interfaces and contracts are adequate.

Conversely, an ownership change can alter incentives or access, so the attainable resource set may depend on P. A richer model is

$$W(P) = \sup_{x \in A(P)} \{F(x) - K(x, P)\}, \tag{59}$$

where $A(P)$ is the feasible set. Cross-partials of F alone do not compare these optimized values. This is the precise place where a technological complementarity argument needs an organizational theory.

13.4 Local gains in a changing environment

If F_t changes as models, demand, or resource prices change, positivity of a cross-partial at time t gives no monotonicity in time. A claim that accumulated data raises integration gain requires, for example, a positive derivative of the relevant gain with respect to a state variable, with other variables controlled. Such a condition is independently testable or refutable.

Even when $\partial_R \Gamma > 0$ in a stylized model, it can fail after saturation, distribution shift, or a change in security cost. We use time-dependent gain functions later only with their stated assumptions. Recursive notation does not turn an unknown derivative into a positive one.

14 Worked organizational and informational examples

14.1 A four-capability model with an intermediate optimum

Let $E = \{1, 2, 3, 4\}$. Define the value of a coalition by

$$n(A) = \sum_{i \in A} a_i + \sum_{\{i,j\} \subseteq A} w_{ij}, \tag{60}$$

where $a_i = 1$, $w_{12} = w_{34} = 3$, and the four cross-pair weights $w_{13}, w_{14}, w_{23}, w_{24}$ equal -2. For any partition,

$$W(P) = 4 + \sum_{\{i,j\} \text{ within a block of } P} w_{ij}. \tag{61}$$

The unique maximizer is $12|34$, with value 10. The singleton partition has value 4, and $\top$ has value $4 + 6 - 8 = 2$. Positive complementarity within two groups coexists with a strong penalty for integrating across them.

To verify uniqueness directly, the two positive edges contribute at most 6. Achieving both requires $1, 2$ to be in the same block and $3, 4$ to be in the same block. The partition $12|34$ achieves both without any negative edge. Any other partition either loses at least one positive edge or adds at least one negative edge. Its value is strictly smaller.

The complete list can be grouped by block sizes:

Partition type	Values occurring	Number
$1+1+1+1$	4	1
$2+1+1$	7 or 2	6
$2+2$	10 or 0	3
$3+1$	3	4
4	2	1

There are $15 = B_4$ partitions. The table makes clear that complete consolidation is neither a generic maximizer nor the only possible effect of strong complementarities.

14.2 All positive pair weights

For the same pair model, a merger of blocks A and B has gain

$$\gamma(A, B) = \sum_{i \in A, j \in B} w_{ij}. \tag{62}$$

If every pair weight is positive, every nonempty disjoint merger is positive and theorem 11.2 applies. With $w_{ij} = 1$ for all six pairs, $W(\top) = 10$ and the singleton value is 4. The proof depends on the sign of the cross-block weights, not on the existence of a common data platform.

If all weights are merely nonnegative, some distinct partitions can tie. For example, when only $w_{12} > 0$ and all other weights vanish, every partition placing 1 and 2 together has the same maximum. Weak monoidal comparisons therefore establish existence of a top optimum, not uniqueness or inevitable convergence to it.

14.3 A constrained integration graph

Suppose only mergers inside $\{1, 2\}$ and inside $\{3, 4\}$ are authorized, and take all weights positive. The admissible partitions refine $12|34$; it is terminal in that restricted poset. If every permitted merger raises value, theorem 12.1 selects $12|34$. The unrestricted technological objective may favor $\top$, but $\top$ is not an admissible object. The conclusion is not an unresolved competition between two theorems; the theorems optimize over different feasible spaces.

One can interpret the restriction as a security domain, a legal prohibition, or a contractual boundary, but none is asserted as an actual legal rule here. The example only shows how an institutional constraint enters mathematically. An economic prediction needs evidence that its assumed feasible set matches the relevant institution.

14.4 A federation with complete information for a query

Let two owners retain datasets A and B, with a shared identifier interface K. They agree on a compatible query to Y and expose only the necessary interface. The pushout property gives a unique compatible global query, while ownership labels remain distinct. If the query is computed through a protocol, the integrated object need not be stored at one owner. The same categorical query can be implemented centrally or through a federation, with different costs and permissions.

This example refutes an inference from universal receiving objects to physical centralization. A colimit is a semantic specification. A particular distributed implementation requires an implementation theorem or protocol. The universal property alone does not choose where the bytes reside or who can inspect them.

14.5 A customer–employee information synergy

Let an employee-side signal Y and a customer-side signal Z be independent fair bits, and let a task target be $X = Y \oplus Z$. The separate signals each have zero mutual information about X, while their combination determines it. This gives a clean example in which joining two different activity sources provides a new task-relevant distinction.

Change the model to $Y = Z = X$. Now each source already determines the task and the second source adds no information. The common architecture of collecting both kinds of trace does not determine which regime applies. Authorization can further prevent their joint use even in the first regime. Thus structural symmetry between employee and customer inputs is not equality of their value, permission, or epistemic status.

14.6 Trace accumulation with diminishing returns

Let each new observation of a fixed bit pass through an independent binary symmetric channel with $p < 1/2$. Additional observations weakly reduce optimal prediction risk, by proposition 9.1, since one can ignore them. The total information about the bit remains at most one bit. It cannot grow at least linearly with a positive fixed value per repeated trace indefinitely.

If each observation costs $c > 0$, the net value of infinitely many observations is not automatically favorable. For a bounded task benefit B, total net surplus is at most $B - nc$ after n observations, regardless of the statistical improvement. This simple bound shows why an endlessly accumulating trace tensor requires a valuation that accounts for task saturation and acquisition cost.

14.7 What the examples establish

The examples are logically complementary. The four-capability model shows a nontrivial intermediate optimum. The all-positive model verifies the sufficient consolidation theorem. The constrained model shows how a different feasible category changes the conclusion. The federation model separates integration from ownership and physical location. The two binary signal models exhibit opposite synergy signs. None is offered as a representative estimate of a real market. Together they show that the formal architecture admits several outcomes and that the outcome is selected by declared assumptions.

15 Trace-mediated feedback and organizational dynamics

15.1 An endofunctor specifies iteration, not improvement

Let $\mathcal{S}$ be a category of system states, $E : \mathcal{S} \to \mathcal{K}$ an execution assignment, $F : \mathcal{K} \to \mathcal{D}$ a trace representation, and $U : \mathcal{S} \times \mathcal{D} \to \mathcal{S}$ an update functor. These types yield an endofunctor

$$\Phi = U \circ \langle \mathrm{id}_{\mathcal{S}}, FE \rangle. \tag{63}$$

For each nonnegative integer n, the iterate Φ^n is defined by composition. Functoriality means that compatible transformations of states induce compatible transformations of updates. It does not make a numerical objective increase, imply a fixed point, or supply convergence.

Proposition 15.1 (Operational invariance of the update)**.** *If $k \simeq k'$ in $\mathcal{K}$, then $U(s, Fk) \simeq U(s, Fk')$ for any state s for which the expression is defined.*

Proof. Apply F to the equivalence and then apply the functor $U(s, -)$. Both preserve equivalences. □

This establishes invariance under the chosen skill semantics. It does not say that the invariant update is beneficial. A constant update, a destructive update, and an oscillatory update can all satisfy it.

Counterexample 15.2 (Perfect observation with perpetual oscillation)**.** Take the discrete state category on $\{0, 1\}$, take the trace representation to be the identity, and set $U(s, d) = 1 - s$. Then Φ alternates the states forever. Observation is exact, but the update has no fixed point. With an objective $J(0) > J(1)$, every second step lowers the objective.

The example is deliberately simple because the invalid inference is simple. Nothing about categorical composition chooses an improvement order. A recursive system becomes an improvement system only after a suitable relation, objective, or contraction property has been supplied and proved compatible with its updates.

15.2 A metric sufficient condition for convergence

Some operational models carry an additional metric or normed realization. That structure is independent of being an infinity-category. Suppose a state set S is a complete metric space and the composite observation $H = FE : S \to D$ is Lipschitz with constant L_H. Suppose the update satisfies

$$d_S(U(s, d), U(s', d')) \leq a\, d_S(s, s') + b\, d_D(d, d') \tag{64}$$

for nonnegative a, b.

Proposition 15.3 (A contraction criterion)**.** *If $q = a + bL_H < 1$, then $\Phi(s) = U(s, H(s))$ has a unique fixed point s_* and*

$$d_S(\Phi^n(s), s_*) \le q^n d_S(s, s_*). \tag{65}$$

Proof. Substitute $d = H(s)$ and $d' = H(s')$ into (64). Then

$$d_S(\Phi(s), \Phi(s')) \le (a + bL_H) d_S(s, s').$$

Thus Φ is a contraction. To recall the proof, consecutive distances in its orbit are at most $q^n d_S(\Phi(s), s)$, so the geometric tail bound makes the orbit Cauchy. Completeness gives a limit, continuity makes it a fixed point, and the contraction inequality forces any two fixed points to coincide. Applying the same inequality against s_* gives (65) by induction. □

The criterion is sufficient and can be checked in a particular metric realization. Full faithfulness of F does not imply any bound on L_H, and a fully faithful categorical functor need not even be presented on metric spaces. A feedback system can become less stable when an observation is made more sensitive. Exact observation and stable control are separate design properties.

15.3 Perturbed updates and a quantitative error floor

Suppose the implemented update $\widetilde{\Phi}_t$ differs from a contraction Φ by at most η_t uniformly along the trajectory:

$$d_S(\widetilde{\Phi}_t(s), \Phi(s)) \le \eta_t. \tag{66}$$

For $s_{t+1} = \widetilde{\Phi}_t(s_t)$, the triangle inequality gives

$$d_S(s_{t+1}, s_*) \le q\, d_S(s_t, s_*) + \eta_t. \tag{67}$$

Induction yields

$$d_S(s_t, s_*) \le q^t d_S(s_0, s_*) + \sum_{j=0}^{t-1} q^{t-1-j} \eta_j. \tag{68}$$

If $\eta_j \le \eta$, the limiting upper bound is $\eta/(1-q)$. If $\eta_j \to 0$, the sum tends to zero: split it into a fixed finite initial part, which decays geometrically, and a tail bounded by an arbitrarily small constant divided by $1 - q$.

This result distinguishes convergence of an ideal update from robustness of an implementation. The perturbations can represent imperfect estimation, stale records, or approximate computation. A semantic equivalence alone does not bound their metric size. An application must specify the metric and explain why the error model is appropriate.

15.4 Finite merger dynamics has a rank function

Let a merger transition be any strict coarsening $P \rightsquigarrow Q$. Define the integer rank

$$\rho(P) = |P| - 1 \in \{0, \dots, m-1\}. \tag{69}$$

Every transition strictly decreases rank. This remains true even if values change over time or the merger is economically harmful.

Proposition 15.4 (Termination by coarsening)**.** *A merger-only trajectory on* $\mathrm{Part}(E)$ *has at most* $|P_0| - 1$ *strict transitions. If every nonterminal state has a permitted positive-gain transition and transitions continue whenever one exists, every maximal trajectory ends at a positive-merger sink. If* $\top$ *is the only such sink, every maximal trajectory ends at* $\top$*.*

Proof. The rank cannot decrease more than $|P_0| - 1$ times. A maximal trajectory cannot stop at a state with an available required continuation. The sink and uniqueness conclusions follow from these two facts. □

The bound counts merger stages, not calendar time. Negotiations may stall indefinitely between stages, and a process may fail to choose an available move. The continued-action hypothesis is essential to translating graph reachability into an actual trajectory. Moreover, stabilization of the partition does not imply that model training, data collection, or organizational state updates have stopped.

If splits are allowed, rank no longer decreases. For a fixed objective on a finite state set, strictly increasing moves still terminate after at most $B_m - 1$ changes, since no state can be revisited. For time-dependent objectives that argument fails. Two states can alternate when each is preferred under a different external regime.

15.5 Transition costs and finite horizons

A static merger gain is not necessarily an investment criterion. Suppose a merger costs $M > 0$ at time zero and yields an incremental flow g at the ends of periods $1, 2, \dots$, discounted by $\beta \in (0, 1)$. Its incremental present value is

$$-M + \sum_{t=1}^{\infty} \beta^t g = -M + \frac{\beta g}{1 - \beta}. \tag{70}$$

Positive static flow g is insufficient when M is too large or the effective horizon is short. For a finite horizon T, replace the fraction by $\beta(1 - \beta^T)g/(1 - \beta)$.

An optimizing dynamic model can use a Bellman equation with state-dependent resources, permissions, and transition costs. The finite static partition theorem becomes a special case in which all changes are available, transition costs are negligible or already capitalized, and

the objective being compared is the relevant long-run value. This identifies an additional bridge required by claims of inevitable consolidation.

16 Changing theories, learning, and observational scope

16.1 A state-indexed family is not one fixed category

An organization can change its operational theory as it acquires tools, introduces new tasks, or changes permissions. Writing W_s, $\mathcal{K}_s$, $\mathcal{D}_s$, and $\mathcal{R}_s$ for each state s is only notation until transport maps are supplied. To compare them functorially one needs, for example, a functor

$$\mathscr{K} : \mathcal{S} \longrightarrow \mathrm{Cat}_\infty \tag{71}$$

or an appropriate cocartesian fibration, together with compatible trace and resource assignments. If transitions are reversible in some directions and contravariant in others, a different variance or a correspondence model may be required.

A sequence of unrelated ambient topoi likewise has no canonical "limiting informational topos" merely because its members are written in a sequence. One needs specified geometric morphisms or other comparison functors and must state the category in which the limit or colimit is taken. Its existence and interpretation then become separate mathematical questions. The fixed-ambient results in this manuscript avoid that unsupported inference.

16.2 State-dependent valuations

Suppose for simplicity that $\mathcal{R}$ and E are fixed while the coalition values vary with a state s. Write $n_s(A)$ and $W_s(P) = \sum_A n_s(A)$. To assert that learning strengthens integration, one might hypothesize

$$\gamma_{s_{t+1}}(A, B) \geq \gamma_{s_t}(A, B). \tag{72}$$

Neither accumulation of records nor functoriality implies this inequality. More accurate learning can reveal that two resources were redundant or that integration carries an underestimated failure cost. The apparent gain can fall even when predictive accuracy improves.

Example 16.1 (Learning reverses a merger ranking)**.** Let two singleton coalitions have net value 1 each. An initial model estimates their union at 3, giving gain 1. After observing correlated failures, the model estimates it at 1.5, giving gain -0.5. The update improves knowledge under the stipulated true model but weakens consolidation pressure. Nothing in the trace architecture excludes this sequence.

If one wants (72) as a theorem, it must follow from a particular model of how the state changes value and cost. For a differentiable scalar accumulation variable R, the sufficient

condition $\partial_R \gamma(A, B; R) \geq 0$ on the entire relevant interval makes the monotonicity transparent. It remains a substantive condition to estimate or justify.

16.3 Choosing an organization from an optimum set

Since $\mathrm{Part}(E)$ is finite, an argmax set exists for any real-valued W_s:

$$\mathcal{M}(s) = \operatorname*{argmax}_{P \in \mathcal{P}_{\mathrm{adm}}} W_s(P). \tag{73}$$

This set can contain several partitions. Turning it into a chosen update requires a selection rule. A fixed total order on the finite feasible set supplies deterministic tie breaking. If each $W_s(P)$ is measurable in s, selecting the first maximizer is measurable, since its selection region is a finite combination of measurable comparisons.

Such a rule is not automatically a functor between state categories, continuous in a metric state space, or institutionally implementable. Small changes in s can switch the maximizing partition discontinuously at a tie. A category-valued argmax correspondence may preserve more structure than an arbitrary single-valued choice, but it does not remove bargaining or authorization requirements.

16.4 Observational improvement and causal effectiveness

Suppose a model predicts an outcome from traces with low error. It does not follow that intervening on a predicted correlate changes the outcome. A trace can record a common cause, a post-treatment variable, or a selection effect. An empirical claim that observation improves production therefore needs a causal comparison or a justified structural model, not just a predictive fit.

This point can be stated without developing a full causal calculus. Let X influence both a logged variable Y and an outcome Z. Prediction from Y can be excellent even if setting Y externally leaves Z unchanged. For example, take $Y = X$ and $Z = X$ under normal operation, but define intervention on Y to leave the structural equation for Z unchanged. The observational equality $Y = Z$ does not identify the effect of intervening on Y.

A higher-categorical trace model can encode interventions and their comparison maps, but it must actually include them. An unadorned functor on observational records does not supply counterfactual semantics. Thus “trace capital” and “productive capability” should be related by a declared realization or intervention model, rather than equated by terminology.

16.5 Distribution shift and contextual restrictions

A trace representation certified over a cover of context B need not remain complete after moving to a larger context B'. One must specify an extension and verify its hypotheses there.

Likewise, statistical sufficiency under one family of distributions does not imply sufficiency under an arbitrary shifted family.

If the model's relevant tasks change, the localization S can change. A trace invariant under the old equivalences may retain distinctions needed in the new task, or it may have irreversibly discarded them. A trace archive is therefore not automatically a universal future skill resource. Its future utility depends on what it recorded, on which models are identifiable from it, and on which uses remain authorized.

17 An empirical program with explicit falsifiers

17.1 What can be measured

The mathematical framework suggests a sequence of empirical questions, each tied to a distinct object. For an operational model one can specify tasks, transformations, and a restricted candidate family. For a trace system one can measure whether selected task distinctions are recoverable, whether certificates are preserved, and which errors occur. For integration one can measure interface compatibility, query equivalence, latency, and maintenance cost. For organizational change one can estimate coalition values or compare partition-dependent objectives.

These measurements are not equivalent to observing the complete internal mapping object of an arbitrary infinity-category. A finite benchmark supports a bounded claim. The paper's exact theorems are useful as specifications and logical constraints, while empirical work concerns how well a concrete system approximates their antecedents.

17.2 Operational tests with declared alternatives

An identification experiment should declare a hypothesis class $\mathcal{H}$, observation procedure, and decision rule. If the class is finite, one can ask which queries separate each pair of hypotheses. A successful finite separating family gives an exact identification theorem for that class. If the class contains all total Boolean functions on an infinite domain, proposition 4.4 forbids the same conclusion.

A meaningful negative result is a pair of operationally distinct candidates that produce the same permitted observations. A meaningful positive result includes both a separating argument and a bound on the cost of the required observations. Merely embedding each candidate as a distinct syntax string does not establish that the operational distinctions can be inferred from observed behavior.

The choice of relevant equivalence is also a potential falsifier. If two executions declared equivalent differ on a task outcome that matters to the application, the localization is misspecified. If a trace distinguishes two operationally equivalent presentations only syntactically,

invariance may fail. Both can be detected before making an economic claim.

17.3 Integration tests and provenance

For a finite interface diagram, one can check whether the declared transition maps compose as required. For systems with higher comparisons, one must test or certify the chosen coherence data at the relevant levels. A discrepancy around a loop is evidence of an integration defect, not evidence that descent itself is false.

Provenance should distinguish observed records, inferred identities, corrected records, and model-generated values. If these are collapsed into one undifferentiated trace object, later queries can treat an inference as an observation. An internal category can represent the transformations between these states, but the type system must retain the distinction. A mathematical colimit can identify entities while a provenance-enriched integration records the evidence and authority for the identification.

One can also compare central and federated implementations of the same semantic query. Equality of outputs on a declared class of instances is an implementation claim. It does not imply equality of operational costs or privacy exposure. Those must be measured under the relevant workload and threat model, rather than assumed from a common universal property.

17.4 Economic tests and sensitivity

The strict all-merger hypothesis quantifies over every disjoint pair of coalitions. For m elementary capabilities, the number of such pairs grows exponentially. Observing a few successful acquisitions is weak evidence for that uniform hypothesis. A parametric model can reduce the burden only by adding structural assumptions, such as the pairwise form (60).

If the pairwise model is justified and every estimated cross weight is positive with a robust margin, one obtains a compact sufficient certificate. If some weights are negative, one must analyze the resulting coalition gains rather than extrapolate from average complementarity. In models with higher-order interactions, pairwise measurements can miss a multiway gain or a multiway harm.

The objective should also specify a time horizon, acquisition and transition costs, and the distribution of gains. A model calibrated to incumbent profits need not predict social welfare. A model of an existing market structure need not identify the value of a counterfactual ownership arrangement without further causal assumptions. These are limitations of the evidence model, not failures of the categorical notation.

17.5 A table of refutable hypotheses

Hypothesis	A possible refutation in a declared model
Trace invariance	An arrow in S maps to a non-equivalence.
Task completeness	Two relevant candidates have identical permitted traces.
Descent compatibility	Chosen transition maps fail a coherence condition.
Authorized integration	The proposed cocone has no permitted lift.
Positive merger gain	The full objective falls after the proposed merger.
Increasing gain over time	A later state has lower integration gain.
Stable feedback	A valid trajectory cycles or diverges under the update.

Each row concerns a hypothesis in an application. The theorems remain conditional implications. A useful formal theory makes such hypotheses inspectable instead of hiding them behind a single assertion that observation necessarily causes consolidation.

18 The combined separation theorem

18.1 Statement with all bridges exposed

Theorem 18.1 (Conditional trace–integration–organization architecture)**.** *Fix the following data:*

(a) an infinity-topos $\mathcal{X}$ with a specified effective epimorphism $u : U \to 1$;

(b) an internal operational category $\mathcal{O}$, an admissible localization $q : \mathcal{O} \to \mathcal{K}$ as in assumption 3.1, and an S-invariant trace functor $T : \mathcal{O} \to \mathcal{D}$;

(c) the induced functor $F : \mathcal{K} \to \mathcal{D}$, whose restriction to U is fully faithful;

(d) a small information diagram $D : I \to \mathcal{X}$ with specified compatible cocones for the downstream operations under consideration;

(e) a finite nonempty capability set E, a declared resource tensor, and a block-separable real objective $W(P) = \sum_{A \in P} n(A)$.

Then the following conclusions hold.

(i) T factors through q uniquely with its comparison up to contractible choice, and F is globally fully faithful.

(ii) The colimit D_* exists and represents compatible downstream cocones by (21).

(iii) Neither conclusion forces common ownership of the information suppliers, a terminal indexing category, an effective decoder, or a positive merger gain.

(iv) If every disjoint binary coalition gain is positive, $\top$ uniquely maximizes W; under the constrained hypotheses of theorem 12.1, the admissible terminal partition uniquely maximizes it instead.

(v) A merger-only trajectory has finitely many strict coarsenings. Reaching the maximizing partition requires the relevant path and continued-action assumptions, such as those of propositions 11.4 and 15.4.

Proof. Factorization is proposition 3.2. The local hypothesis and theorem 5.3 give full faithfulness. Cocompleteness of $\mathcal{X}$ gives D_*, and theorem 7.1 gives its operation space. The ownership countermodel is proposition 7.3; computable decoding is absent from these data and is additionally obstructed in the situations of propositions 4.4 and 4.5. Values can be chosen as in the intermediate-optimum example independently of the categorical data, so no positive merger gain follows. The two optimization conclusions are theorems 11.2 and 12.1. Finally the rank proof in proposition 15.4 gives finite coarsening, with proposition 11.4 supplying an appropriate reachability condition. □

The theorem is an assembly of standard structural results and explicit finite deductions. Its purpose is to prevent an invalid transfer of strength from one clause to another. In particular, clause (iv) depends on the additional economic antecedent; it is not a consequence extracted from clause (i).

18.2 A monoidal sufficient version

Suppose in addition that F and $J : \mathcal{D} \to \mathcal{R}$ are strong symmetric-monoidal and that normalized valuations V, C satisfy (34)–(35) on the declared resource class. Set $n(A) = V(r_A) - C(r_A)$. Then (41) gives nonnegative gain, so $\top$ is a maximizer. To obtain uniqueness one needs strict positive gain for every disjoint merger, or another hypothesis such as proposition 11.4.

The idempotence theorem limits this version. If every resource is tensor-idempotent and V is nonnegative, its lax value part is trivial. This does not invalidate the partition theorem; it invalidates applying that particular global valuation hypothesis to that tensor while expecting positive data value. A labeled disjoint-bundle model or a different information valuation may be appropriate instead.

18.3 Independence countermodels

It is useful to see how little of the architecture forces the other layers. Take $\mathcal{X} = \mathcal{S}$, take $\mathcal{O} = \mathcal{K} = \mathcal{D}$ to be a small discrete category, and let $T = F = \mathrm{id}$. Full faithfulness holds. Assign any finite coalition function n independently. One obtains a top optimum, an intermediate optimum, or many ties by changing n without changing the trace theorem.

Conversely, take a constant trace functor from a two-object discrete operational category to the terminal category. Trace full faithfulness fails by counterexample 4.2. Choose strictly positive pairwise economic weights. The partition model still has a unique top optimum. Thus economic consolidation in this model does not require complete observation either.

For integration, choose any small diagram with a colimit. It can be paired with either trace functor and either valuation. The existence of its universal receiver supplies no missing inequality. These countermodels establish logical independence within the declared framework. They do not claim that the layers are causally independent in an actual industry. Empirical relations between them can be strong, but they are additional relations to model and test.

18.4 The strongest justified interpretation

The model supports a conditional account of organizations that represent their own operations and reuse the resulting information. Effective descent can globalize a locally certified semantic representation. Universal integration can reduce the number of distinct semantic descriptions needed for compatible operations. Resource complementarity can make some ownership changes attractive. Each conclusion is valuable at its own level.

The model does not establish a universal historical direction toward one corporation or one informational topos. Different permissions, interfaces, task boundaries, objective functions, and dynamic constraints can produce federations, partial integration, multiple stable organizations, or cyclical reorganization. The mathematical contribution is to state these alternatives in a common typed language and to identify sufficient conditions for particular outcomes.

The resulting research program is concrete: construct restricted operational theories with meaningful equivalences; design authorized traces with declared fidelity; implement coherent interfaces; estimate the actual gains and costs of alternative partitions; and prove the relevant stability or reachability claims. Higher-categorical structure supplies the language of coherence and descent. The economic and empirical bridges remain visible obligations.

19 Ideal trace embeddings and the role of probes

19.1 A semantic existence construction

The paper has emphasized how strong full faithfulness is. There is nevertheless a standard ideal construction showing that the property is mathematically coherent. Let C be a small external infinity-category and let

$$\mathcal{P}(C) = \mathrm{Fun}(C^{\mathrm{op}}, \mathcal{S}) \tag{74}$$

be its presheaf infinity-category. The Yoneda embedding assigns to an object its entire family of incoming probes:

$$y : C \longrightarrow \mathcal{P}(C), \qquad y(c)(a) = \mathrm{Map}_C(a, c). \tag{75}$$

The Yoneda lemma identifies

$$\mathrm{Map}_{\mathcal{P}(C)}(y(c), y(d)) \simeq \mathrm{Map}_C(c, d). \tag{76}$$

Thus an ideal trace carrying every such probe response is fully faithful. Internal presheaf and Yoneda constructions admit an appropriate internal formulation in the setting developed by Martini and Wolf [4].

The interpretation matters. The presheaf $y(c)$ is not a finite list of events. It packages an entire coherent response theory, including all mapping spaces from all probes. It may be extremely large as a presentation and need not be computably recoverable from finite interactions. Yoneda shows that ideal semantic observability is possible in a suitable mathematical target; it does not contradict the finite-query obstruction.

19.2 Why the Yoneda comparison is fully faithful

For intuition, in an ordinary small category a natural transformation $\alpha : C(-, c) \to C(-, d)$ is determined by

$$f := \alpha_c(\mathrm{id}_c) : c \to d. \tag{77}$$

For any $g : a \to c$, naturality gives $\alpha_a(g) = f \circ g$. Conversely, postcomposition with f defines such a natural transformation. These two assignments are inverse. The infinity-categorical Yoneda lemma upgrades the correspondence from sets to mapping spaces with all coherences.

This proof explains what an ideal trace must preserve. It is not enough to record an answer for each probe independently; the answers must transform naturally when one probe maps to another. The naturality data ensure that the family describes one represented object rather than a collection of incompatible responses. The same coherence principle appears in descent and in colimit integration.

19.3 Restricted probes and a precise completeness condition

Let $j : A \to C$ be a chosen small category of probes. Define the restricted nerve

$$N_A : C \longrightarrow \mathrm{Fun}(A^{\mathrm{op}}, \mathcal{S}), \qquad N_A(c)(a) = \mathrm{Map}_C(j(a), c). \tag{78}$$

Its full faithfulness is an exact formulation of whether the selected probes determine all transformations among the relevant objects. A probe family with this property is commonly described through density. The condition can fail even when the probes distinguish the objects as a set, because they may not determine the mapping spaces or their composition.

In the finite linear example, basis probes identify matrices and therefore their linear action. In an unrestricted function class, finitely many input probes fail to distinguish all objects. In a higher category, probing only object-level invariants can also miss nontrivial automorphisms or homotopies. A completeness specification should therefore state whether it concerns object separation, arrow separation, or reconstruction of the full restricted nerve.

Example 19.1 (Object identification without transformation reconstruction)**.** Let C be a one-object category with endomorphism group $G \neq 1$. A trace that records only the unique object identifies all objects perfectly because there is only one. But the map to the terminal trace category loses the group G. It is fully faithful only when G is trivial. Reconstructing the operational transformation theory is stronger than identifying the object labels.

19.4 Discrete trace targets impose a strong limitation

Proposition 19.2 (Full faithfulness into a discrete category)**.** *If an external infinity-category C admits a fully faithful functor to a discrete category D, then C is equivalent to a discrete category. If the target is an ordinary category, every mapping space of C is* 0*-truncated.*

Proof. For discrete D, each target mapping space is empty or contractible, and every target arrow is an equivalence. Full faithfulness makes each source mapping space empty or contractible and reflects equivalences, so all source arrows are equivalences. Thus C is a groupoid with no nontrivial higher or isotropy information; equivalently, it is a 0-truncated space, hence equivalent to a discrete category. For an ordinary target, its mapping spaces are sets, and full faithfulness transfers that truncation to the source mapping spaces. □

This proposition does not say that a finite file cannot *encode* a rich category. A file can be interpreted as a presentation with generators, relations, and algorithms. The resulting trace category is then the semantic category produced by the interpreter, not the discrete set of raw file strings. The proposition applies only when the chosen target really has the asserted discrete categorical structure.

The distinction is important for claims about machine-readable logs. Raw readability is a property of storage syntax. Full faithfulness concerns a target with a specified interpretation of transformations and equivalences. One cannot infer the latter from the former, nor refute the possibility of a rich interpretation merely because its bytes are finite.

19.5 Avoiding a tautological completeness claim

Given a collection of observed objects, one can define a trace category by transporting exactly the operational mapping spaces onto their names. The resulting embedding can be fully faithful by construction. Such a definition is useful as a semantic specification, but it provides no independent evidence that a practical observer knows those mapping spaces.

An evidentially substantive claim requires an independently described observation process and a correctness argument comparing it with the operational theory. In the restricted linear model, the matrix reconstruction gives that argument. In a real system, one might supply typed certificates, a proven compiler relation, or a bounded experimental identification theorem. Merely defining the target to be the source under another name is not an empirical observability result.

This is the proper reading of the self-representation language. A system can have an internal mathematical representation of its operational theory without possessing an efficient, learned, complete, or authorized practical model of that theory. The extra adjectives each introduce another proof obligation. Keeping them separate makes the ideal construction informative rather than circular.

20 A homogeneous cost model with exact optimal boundaries

20.1 Fixed overhead and convex coordination

The signed-weight example provides one mechanism for partial consolidation. A second uses only homogeneous capabilities and a size-dependent organizational cost. Let there be m identical elementary capabilities. Suppose output is additive and independent of ownership, so it can be omitted when minimizing cost. Let a block of size $k \geq 1$ have cost

$$c(k) = h + bk^2, \qquad h, b > 0. \tag{79}$$

The fixed term represents overhead per organization. The convex term represents a stipulated coordination or correlated-risk cost. It is a constructed model, not an empirical claim that organizational cost is always quadratic.

Merging sizes a and d changes total cost by

$$c(a + d) - c(a) - c(d) = -h + 2bad. \tag{80}$$

The merger saves cost precisely when $2bad < h$. The sign depends on the sizes of both blocks. Positive economies for small mergers can coexist with diseconomies for large ones, so an all-merger oplax-cost axiom need not hold.

20.2 An exact balancing lemma

Lemma 20.1 (Balanced blocks at fixed organization count)**.** *Among partitions into exactly s nonempty blocks, the minimum of $\sum_{A \in P} |A|^2$ is attained precisely when all block sizes differ by at most one.*

Proof. If two sizes satisfy $a \geq d+2$, moving one element from the larger to the smaller changes

their squared sum by

$$(a-1)^2 + (d+1)^2 - a^2 - d^2 = -2(a-d) + 2 < 0. \tag{81}$$

Thus an optimum cannot have such a pair. Repeated balancing terminates because the nonnegative integer squared sum strictly decreases. Every balanced vector has the same sizes up to order, so all such vectors attain the same minimal value. □

Write $m = sq + r$ with $0 \le r < s$. A balanced partition has $s - r$ blocks of size q and r blocks of size $q + 1$. Therefore the exact minimum cost among s organizations is

$$K_s = sh + b\big((s-r)q^2 + r(q+1)^2\big). \tag{82}$$

The global optimum is obtained by comparing K_s for $s = 1, \dots, m$. This reduces the homogeneous instance to m numerical comparisons and identifies all optimal block-size patterns, though many labeled partitions can share the same optimal sizes.

20.3 A twelve-capability instance

Take $m = 12$, $h = 9$, and $b = 1$. Four blocks of size three have cost

$$K_4 = 4 \cdot 9 + 4 \cdot 3^2 = 72. \tag{83}$$

The complete one-dimensional comparison is

s	1	2	3	4	5	6	7	8	9	10	11	12
K_s	153	90	75	72	75	78	85	92	99	106	113	120

The optimal organization count is uniquely four, while the labeled partition is not unique: every division into four triples has the same cost. A single organization is much more expensive than the optimum in this stipulated model, despite the fixed-overhead saving it achieves.

For a divisible continuous approximation with common block size k, total cost is $mh/k + bmk$. Differentiation gives the preferred size $k = \sqrt{h/b}$. The exact formula (82) handles indivisibility and avoids inferring a fractional number of firms. In the chosen instance the approximation yields size three and agrees with the exact optimum.

20.4 Compatibility with complete observation and integration

This cost model can be paired with the identity trace representation and with any fixed integrated information diagram. None of its cost calculations changes. Thus even a perfectly represented, coherently integrated operational system can have a strictly intermediate optimal organization count. The countermodel does not depend on a failure to collect enough data.

One can also let a shared interface reduce h while leaving b fixed. The continuous preferred size then falls, because separate organizations become cheaper to operate. In this model, better interoperability encourages smaller organizations. Alternatively, if shared models reduce b more strongly than h, the preferred size can rise. Both comparative statics are possible; the technological improvement must be related to the actual cost parameters before its organizational direction is determined.

This final example makes the conditional nature of the overall argument concrete. Integration can change the economics of boundaries, but it need not change every cost in the direction favoring larger blocks. A theory of AI-mediated organization should identify which parameters change, why they change, and how robustly the resulting optimizer follows.

A Expanded internal and descent calculations

A.1 Mapping objects in an arbitrary parameter context

The definition of full faithfulness can be unpacked without choosing global objects. Let $V \to B$ be a parameter context and let $(x, y) : V \to C_0 \times_B C_0$ be a generalized pair. Pulling back the right side of (12) along (x, y) gives

$$V \times_{C_0 \times_B C_0} ((C_0 \times_B C_0) \times_{D_0 \times_B D_0} D_1) \simeq V \times_{D_0 \times_B D_0} D_1. \tag{84}$$

The pullback of the left side is $V \times_{C_0 \times_B C_0} C_1$. Thus the global comparison induces exactly the required comparison of internal mapping objects over V. If the global comparison is an equivalence, all these comparisons are equivalences.

Conversely, suppose the comparisons are equivalences for every generalized pair. Take $V = C_0 \times_B C_0$ with its identity pair. The resulting map is α_F itself. Therefore testing all parameter contexts is equivalent to the intrinsic definition. There is no need to assume that the topos has enough global points.

Counterexample A.1 (Why global-point tests are insufficient)**.** In the topos of sheaves on the circle, let P be the sheaf of sections of the nontrivial double cover $z \mapsto z^2$. The sheaf P is locally inhabited but has no global section. The initial sheaf also has no global section. Hence the map $\varnothing \to P$ induces a bijection between empty sets of global sections, although it is not an isomorphism of sheaves. The same example embeds into a space-valued sheaf infinity-topos.

This explains why a statement phrased only for named global skills can miss local distinctions. An internal theorem must quantify over context-dependent skills, not just objects that happen to have globally chosen representatives.

A.2 A fully explicit cover calculation in sets

Take $B = \{0, 1\}$ and $U = \{u_0, u_1, u_1'\}$ with $u_0 \mapsto 0$ and both $u_1, u_1' \mapsto 1$. This is a surjection, hence an effective epimorphism. An object over B is a pair of fibers (A_0, A_1). Pullback to U produces (A_0, A_1, A_1), together with the identifications specified by the Cech nerve.

If $f : A \to A'$ becomes an isomorphism after pullback, its components over u_0 and u_1 show that f_0 and f_1 are bijections. Therefore f is a bijection. The duplicate presentation of the fiber over 1 does not add a new state of the underlying object; it adds descent data ensuring that the two local presentations agree.

Now replace U by $\{u_0\}$, mapping only to 0. The inclusion $A = \{0\} \hookrightarrow B$ pulls back to an isomorphism over U, but is not an isomorphism globally. This map is not a cover of B. The effective epimorphism hypothesis in the locality theorem cannot be replaced by the presence of some observed contexts.

A.3 Descent for the full-faithfulness square

The square whose cartesianness expresses full faithfulness is

$$\begin{array}{ccc} C_1 & \xrightarrow{F_1} & D_1 \\ {\scriptstyle (s,t)}\downarrow & & \downarrow{\scriptstyle (s,t)} \\ C_0 \times C_0 & \xrightarrow[F_0 \times F_0]{} & D_0 \times D_0. \end{array} \tag{85}$$

It is not enough to know that the four objects locally have the correct cardinalities or homotopy groups. The map from the upper-left corner to the pullback of the other three corners must be an equivalence. Local restriction preserves the square and its comparison because it preserves finite limits.

This observation generalizes. Any property of a map expressed as a particular finite-limit comparison being an equivalence can be detected after an effective epimorphism. It does not follow that every property of an internal category is local by the same one-line proof. Properties involving global sections, chosen bases, or a preferred representative can fail even when their local versions hold.

A.4 Equivalence reflection in mapping-space form

For completeness, let $F : C \to D$ be fully faithful and suppose $f : x \to y$ maps to an equivalence. Choose an inverse $h : Fy \to Fx$ and homotopies $hFf \simeq \mathrm{id}_{Fx}$ and $Ffh \simeq \mathrm{id}_{Fy}$. The mapping-space equivalence for (y, x) lifts h to a point $g \in \mathrm{Map}_C(y, x)$. The equivalence for (x, x) identifies the two points gf and id_x through the lifted path. Similarly, the equivalence for (y, y) supplies $fg \simeq \mathrm{id}_y$. These data exhibit f as an equivalence.

The chosen inverses and homotopies are not unique as literal expressions. Their space of coherent choices has the appropriate contractibility for equivalences. This is why the correct conclusion is reconstruction up to equivalence with coherent data, not an equality of concrete log strings or proof terms.

A.5 Truncation as a separate approximation problem

The pullback counterexample involving S^1 has a second lesson. Connectedness of an object does not mean its loop space is connected. Truncating before taking a dependent construction can therefore erase information that survives after the construction. The same issue occurs whenever composability depends on matching endpoints over a nontrivial object space.

For an external category enriched in spaces with a fixed set of objects, define a new enriched category by truncating each mapping space to n-types. Since truncation preserves finite products, the enriched composition maps descend. Associativity and units descend

from the original coherent composition in a suitable enriched model. This is an example where a bounded approximation can be justified, but the justification uses fixed objects and product-based enrichment rather than an unrestricted statement about levelwise complete Segal objects.

When working internally, a reader should either use an established categorical truncation construction or include the corresponding verification in the chosen presentation. The main local full-faithfulness theorem does not need truncation, so there is no reason to make its correctness depend on a casual truncation convention.

B Finite observation, reversibility, and information bounds

B.1 Exactly reversible finite channels

Let $K : Y \leftarrow X$ be a channel between finite sets. A stochastic left inverse is a channel $L : X \leftarrow Y$ with $LK = \mathrm{id}_X$. This notion recovers the input almost surely from one observation while allowing independent noise in the output.

Theorem B.1 (Disjoint-support criterion)**.** *A finite channel K has a stochastic left inverse if and only if the supports of $K(- \mid x)$ and $K(- \mid x')$ are disjoint whenever $x \neq x'$.*

Proof. Suppose $LK = \mathrm{id}_X$. For each x,

$$1 = \sum_y K(y \mid x) L(x \mid y). \tag{86}$$

The right side is a convex combination of numbers in $[0, 1]$, so every y with $K(y \mid x) > 0$ must have $L(x \mid y) = 1$. If the same y were in the support for $x' \neq x$, it would also require $L(x' \mid y) = 1$, contradicting the row-sum condition of a channel.

Conversely, if the supports are disjoint, assign each output in a nonempty support to its unique input. Define L as this deterministic decoder on the union of supports, and arbitrarily on outputs that never occur. Then LK is the identity. □

The criterion allows multiple outputs for one input, so stochastic reversibility does not require K to be a permutation matrix. It requires that the output noise never create ambiguity about the input. A binary symmetric channel with nonzero crossover fails precisely because its two conditional supports coincide.

B.2 Total variation contracts under postprocessing

For probability distributions p, q on a finite set, define

$$d_{\mathrm{TV}}(p, q) = \frac{1}{2} \sum_x |p(x) - q(x)|. \tag{87}$$

For a channel K, the triangle inequality gives

$$2d_{\mathrm{TV}}(Kp, Kq) = \sum_y \left| \sum_x K(y \mid x)(p(x) - q(x)) \right| \tag{88}$$

$$\leq \sum_{y,x} K(y \mid x)|p(x) - q(x)| \tag{89}$$

$$= \sum_x |p(x) - q(x)|. \tag{90}$$

Hence $d_{\mathrm{TV}}(Kp, Kq) \leq d_{\mathrm{TV}}(p, q)$. Postprocessing cannot make two possible worlds more distinguishable in this sense. It can make a relevant distinction more accessible to a restricted computation, but that is a computational qualification rather than an increase in the underlying statistical distance.

For equally likely hypotheses with output distributions p, q, the optimal binary testing error is

$$e_* = \frac{1}{2}(1 - d_{\mathrm{TV}}(p, q)). \tag{91}$$

To see this, choose a decision region A for the first hypothesis. Its success probability is $\frac{1}{2}[p(A) + q(A^c)] = \frac{1}{2}[1 + p(A) - q(A)]$. Maximizing $p(A) - q(A)$ yields total variation. A channel with overlapping supports can therefore be identifiable while still having strictly positive one-shot error.

B.3 An approximate privacy–recovery tradeoff

An (ε, δ)-private mechanism satisfies

$$\Pr[M(d) \in A] \leq e^{\varepsilon} \Pr[M(d') \in A] + \delta \tag{92}$$

for adjacent inputs. Suppose a decoder succeeds with probability at least $1 - \alpha$ on each of two adjacent distinct databases d, d'. With A the event that the decoder outputs d, we have

$$1 - \alpha \leq e^{\varepsilon}\alpha + \delta. \tag{93}$$

Therefore

$$\alpha \geq \frac{1 - \delta}{1 + e^{\varepsilon}}. \tag{94}$$

For a randomized decoder, include its independent randomness in the output space; postprocessing preserves the privacy inequality. Thus the bound is not restricted to deterministic decisions.

The inequality quantifies the earlier exact obstruction. For finite ε and $\delta < 1$, uniformly vanishing reconstruction error for the protected adjacent pair is impossible. It does not rule out exact recovery of a quotient that identifies that pair, nor does it bound every possible task accuracy by the same number. The protected distinction is essential to the statement.

B.4 A finite-hypothesis entropy bound

Let H be a uniformly distributed hypothesis from a set of $M \geq 2$ candidates, let Y be the observed record, and let $\widehat{H}(Y)$ be a decoder with error probability e. Use base-two logarithms. Let B indicate whether the decoder erred. Since B is determined by (H, Y),

$$H(H \mid Y) \leq H(B \mid Y) + H(H \mid B, Y) \tag{95}$$

$$\leq h_2(e) + e \log_2(M-1), \tag{96}$$

where h_2 is the binary entropy. Since $I(H;Y) = \log_2 M - H(H \mid Y)$, we obtain

$$\log_2 M \leq I(H;Y) + h_2(e) + e \log_2(M-1). \tag{97}$$

This familiar finite information bound is derived here to specify exactly what it controls. If the record contains at most b bits, then $I(H;Y) \leq H(Y) \leq b$. Small decoding error therefore requires a record large enough to discriminate the declared candidate family. The bound does not speak about an unspecified totality of human knowledge.

If observations are interactive, let Y include the full transcript of queries and responses. The same inequality applies to that transcript. A query strategy can improve the information acquired per observation, but it cannot evade the entropy requirement for distinguishing a large finite family. For an infinite unrestricted family, proposition 4.4 supplies a more direct nonidentifiability result.

B.5 Repeated noisy binary observations

Let a fixed fair bit X be observed through n independent copies of a binary symmetric channel with crossover $p < 1/2$. For odd n, majority vote has error

$$e_n = \sum_{k=(n+1)/2}^{n} \binom{n}{k} p^k (1-p)^{n-k}. \tag{98}$$

It is positive for every finite n when $p > 0$, although it tends to zero. Exact one-shot recovery and consistent statistical estimation are therefore different properties even in this elementary model.

A direct exponential bound follows from the moment generating function. If Z_i indicates an error, then for any $t > 0$,

$$\Pr\left[\sum_i Z_i \geq n/2\right] \leq e^{-tn/2} \mathbb{E} \exp(t \sum_i Z_i) \tag{99}$$

$$= \left(e^{-t/2}(1 - p + pe^t)\right)^n. \tag{100}$$

Choosing $e^t = (1-p)/p$ gives

$$e_n \le \big(2\sqrt{p(1-p)}\big)^n, \tag{101}$$

whose base is less than one. This is a consistency result under conditional independence and a stationary known noise model. Correlated errors or changing channels need a different analysis.

B.6 Task information as a bounded resource

For a fixed finite target X, every observation Y satisfies $I(X;Y) \le H(X)$. A sequence of increasingly rich observations can approach complete knowledge of that target while its marginal value tends to zero. Assigning a fixed positive additive value to each observation would eventually exceed the target's total entropy if the scalar is meant to measure information about that one target.

Economic value can exceed entropy in any chosen currency, and a record can serve many tasks. The point is not a universal conversion rate between bits and money. It is that a valuation needs a task and use model. A bounded decision benefit, an unbounded portfolio of new tasks, and an expanding set of permissions are different resource regimes. They should not be conflated by a generic statement that trace capital increases forever.

C Finite partition algorithms and path certificates

C.1 Exact optimization in the separable model

For a coalition function n, let $F(A)$ be the maximum partition value of a subset $A \subseteq E$, with $F(\varnothing) = 0$. Fix a total order of E and let a be the least element of a nonempty A. Then

$$F(A) = \max_{\substack{B \subseteq A \\ a \in B}} \big(n(B) + F(A \setminus B)\big). \tag{102}$$

Every partition has a unique block containing a, so this recurrence neither omits an optimum nor needs to count the same partition by different orders of its blocks.

Proposition C.1 (Correctness of subset dynamic programming)**.** *Equation* (102) *computes the global optimum of the block-separable objective. With precomputed coalition values it uses* $O(3^m)$ *arithmetic comparisons and* $O(2^m)$ *stored values, up to polynomial representation costs.*

Proof. Induct on $|A|$. In any partition choose the block B containing a. The remaining blocks form a partition of $A \setminus B$ and contribute at most $F(A \setminus B)$ by induction. This proves the upper bound in the recurrence. Conversely, choose a maximizing B and an optimal partition of its complement; their union is a partition attaining that value. For complexity, each disjoint

pair $(B, A \setminus B)$ corresponds to assigning each elementary item to B, to its complement, or outside A. There are at most 3^m such assignments. There are 2^m states. □

The algorithm is exponential and is intended for small portfolios or for a proof certificate on a restricted model. General partition-dependent externalities $H(P)$ break its simple optimal-substructure argument. One must enrich the state or use a different algorithm. A recurrence that silently drops H would optimize the wrong objective.

C.2 A bottleneck certificate for improving reachability

Let the feasible merger graph be a directed acyclic graph ordered by coarsening, with edge gain $g(P, Q) = W(Q) - W(P)$. Define the bottleneck of a path as its minimum edge gain. Put $b(\top) = +\infty$, and for other partitions define

$$b(P) = \max_{P \to Q} \min\{g(P, Q), b(Q)\}, \tag{103}$$

with maximum of an empty set equal to $-\infty$. Compute states in increasing order of their number of blocks, so each successor value is already known.

Proposition C.2 (Bottleneck characterization)**.** *The value $b(P)$ is the largest minimum gain among admissible paths from P to $\top$, with $-\infty$ if there is no such path. In particular, $b(P) > 0$ if and only if a strictly improving path exists.*

Proof. Every path from P begins with some edge $P \to Q$. Its bottleneck is the smaller of that edge gain and the bottleneck of the remaining path. Maximizing first over continuations and then over first edges yields (103). Induction on the block count proves the characterization. A finite path has all gains positive exactly when its minimum gain is positive. □

This produces a concrete reviewable certificate for a dynamic claim. A unique global optimum can have $b(P) \leq 0$ from some initial states. In counterexample 11.3, the singleton partition has bottleneck -1, although the final objective gap is 1. The difference is not a small technicality: it distinguishes a profitable destination from a viable sequence of individually profitable moves.

C.3 Telescoping and path independence of total gain

For any path $P_0 \to P_1 \to \cdots \to P_k$ under a fixed objective,

$$\sum_{j=0}^{k-1} g(P_j, P_{j+1}) = W(P_k) - W(P_0). \tag{104}$$

Thus total gain is independent of the route, while the minimum intermediate gain and transition costs can depend strongly on it. The bottleneck criterion exploits exactly this

distinction. A static objective comparison determines the sum, not whether every summand is favorable.

If objective values change during the path, the sum no longer telescopes to a difference under one objective. Writing $g_t = W_t(P_{t+1}) - W_t(P_t)$, one must include drift terms comparing $W_{t+1}(P_{t+1})$ with $W_t(P_{t+1})$. A model that treats learning as endogenous should account for this drift rather than apply a stationary potential argument unchanged.

C.4 A less restrictive positive-graph theorem

Consider the pairwise model (60) with all $w_{ij} \geq 0$. Let G_+ be the undirected graph having an edge when $w_{ij} > 0$.

Theorem C.3 (Connected positive interactions)**.** *For $m \geq 2$, the top partition is the unique maximizer of the pairwise objective if and only if G_+ is connected. More generally, every maximizing partition has each connected component of G_+ contained in a block; arbitrary mergers of different components leave the value unchanged.*

Proof. The top value includes all nonnegative edges. A partition has the same value exactly when it loses no positive edge across blocks. This holds exactly when the endpoints of every positive edge lie in one block, equivalently when each connected component is contained in a block. If G_+ is connected, this forces a single block. If it is disconnected, the component partition is a distinct maximizer. □

The theorem gives a useful weakening of strict all-merger positivity. On a path 1–2–3 with positive adjacent weights and $w_{13} = 0$, merging singleton blocks $\{1\}$ and $\{3\}$ has zero gain. Nevertheless the top partition is uniquely optimal because the positive graph is connected. Every non-top partition separates at least one positive edge, so at least one positive merger remains available.

C.5 Recovering clustered organizations with signed weights

Suppose E has a fixed grouping $G_1, \ldots, G_k$. Assume weights inside each G_i are nonnegative with connected positive graph, while every cross-group pair has strictly negative weight. Then the group partition $\{G_1, \ldots, G_k\}$ is the unique maximizer.

To prove this, start from any partition. Split each block along the groups. Every such split removes negative within-block cross-group weights and never removes a positive within-group edge, so it weakly improves value and strictly improves it whenever the original block crossed groups. Next merge all fragments belonging to the same G_i. This loses no value, and connectedness implies that any proper fragmentation loses some positive edge relative to the fully joined group. Hence equality with the group partition is possible only for that partition itself.

The argument generalizes the four-capability example and supplies a clean sufficient condition for durable intermediate organizational boundaries. It is entirely compatible with perfect trace representation. The sign pattern of interaction values, not a failure of observability, determines the boundary.

C.6 Finite computation as illustration rather than universal verification

The companion calculations enumerate all fifteen partitions of four capabilities, evaluate the signed-weight example, and check the three-capability trap. They also compute the binary signal entropies. These checks catch arithmetic and transcription errors in the worked examples. They do not verify the foundational theorems for every infinity-topos, establish empirical values, or replace the proofs.

This distinction matters when presenting executable examples in a mathematical paper. A finite calculation can certify a finite optimization instance. It cannot certify an unrestricted economic prediction. The computational scope is therefore stated alongside the results rather than left to be inferred from the presence of code.

D Reusable information and alternative valuation regimes

D.1 Three tensors that should not be confused

For a fixed universe of records, subsets form a join-semilattice with tensor union and unit the empty set. It is idempotent: $A \cup A = A$. A monotone information value can increase with the accessible subset while assigning no new value to a duplicate. This is the setting in which theorem 10.2 directly rules out a positive globally lax additive valuation.

By contrast, the free commutative monoid of labeled storage tokens has tensor multiset union. Two copies of a token are distinct physical resources, even when their contents match. A cost functional can assign a positive cost to each copy. Here $x \otimes x$ is not equivalent to x as a physical resource bundle. One can have additive storage cost while information content remains idempotent under a separate forgetful map to accessible records.

A third tensor combines experiments under a specified joint distribution. A pair of signals can be identical copies, independent repetitions conditioned on a latent state, or correlated through another source. The marginals alone do not determine the joint experiment. The tensor operation must encode that choice. For the XOR example, conditioning on the target creates dependence between the two signals, and this dependence is exactly part of the synergy mechanism.

These distinctions can be formalized by functors between resource categories. A physical-storage category can map to an information-access poset by forgetting multiplicity. A licensed-resource category can map to a raw-data category by forgetting use restrictions. Such

functors need not preserve or reflect all economic comparisons. Losing multiplicity changes storage cost; forgetting a license changes what can be done lawfully or contractually within the model.

D.2 Cartesian copying is not tensor idempotence

In a cartesian monoidal category every object has a diagonal $x \to x \times x$, but generally $x \times x \not\cong x$. A two-element set has a four-element square. The diagonal lands only on the two matching pairs. Thus the existence of a copying morphism must not be mistaken for an isomorphism with the entire product object.

The free-copying corollary uses a monotonicity assumption to derive its obstruction, whereas the idempotence theorem uses equivalence invariance. Those are distinct proofs with distinct hypotheses. If the arrows of a category are all set maps, interpreting every arrow as a free conversion is already an economic modeling decision. If only certain physical operations are free, use the corresponding subcategory.

Likewise, nonrival use of a dataset by several firms does not mean that every license is transferable or that all simultaneous uses are costless. A digital object can be copyable at the bit level while access, verification, training, storage, and authorized deployment remain costly. The resource theory should represent whichever of these constraints matter to the question.

D.3 Decision value as an idempotent monotone

Fix a finite decision problem with baseline risk R_0 when no signal is available. Let $R(Y)$ be the optimal risk with signal Y, and define

$$V_{\mathrm{dec}}(Y) = R_0 - R(Y). \tag{105}$$

The value is nonnegative because the decision maker can ignore Y. If Y' is a garbling of Y, proposition 9.1 gives $V_{\mathrm{dec}}(Y) \geq V_{\mathrm{dec}}(Y')$. Providing an identical duplicate of the same realized signal does not enlarge the available decision rules, so the value is idempotent under duplicate access.

If the loss is bounded between 0 and L, then $0 \leq V_{\mathrm{dec}}(Y) \leq L$. This valuation is mathematically meaningful and economically interpretable, yet it cannot be a positive globally superadditive valuation for an idempotent tensor. It can exhibit positive synergy for some distinct signals and negative synergy for redundant ones. The alternative valuation regime therefore retains useful complementarity without imposing a false universal sign.

The model can be extended to multiple decision problems by summing their values, but one must state whether a resource use can serve all problems simultaneously and whether their losses are commensurable. A single scalar can conceal distributional differences between stakeholders. A vector-valued objective or an ordered set of feasible tradeoffs may be more

appropriate when no agreed conversion exists.

D.4 Conditional independence yields diminishing information returns

Let $(Y_i)_{i\in E}$ be jointly distributed finite signals and define $v(A) = I(X; Y_A)$, where Y_A is the joint signal from indices in A. Suppose the signals are conditionally independent given X. For $j \notin A$,

$$v(A \cup \{j\}) - v(A) = I(X; Y_j \mid Y_A) = H(Y_j \mid Y_A) - H(Y_j \mid X). \tag{106}$$

The last equality uses conditional independence to replace $H(Y_j \mid X, Y_A)$ by $H(Y_j \mid X)$. If $A \subseteq B$ and $j \notin B$, conditioning reduces entropy, so

$$v(A \cup \{j\}) - v(A) \geq v(B \cup \{j\}) - v(B). \tag{107}$$

Thus v is monotone and submodular under this conditional independence hypothesis.

This derivation explains why repeated observations of one target often have diminishing returns. It does not apply to arbitrary signal systems. In the XOR example the two inputs are not conditionally independent given their parity, and increasing returns are possible. The sign of information synergy depends on the statistical structure, not just on the number of available signals.

Monotonicity says that an unconstrained ideal decision maker can ignore an extra signal. Net value can still decrease when the signal has a mandatory acquisition or compliance cost. A practical model must distinguish optional access from compulsory processing. This is another reason not to interpret information monotonicity as a theorem of economic consolidation.

D.5 A bounded semilattice valuation

For a finite target universe U, let accessible information resources be subsets $A \subseteq U$, with tensor union. Define $v(A) = \sum_{u\in A} w_u$ for nonnegative weights. Then

$$v(A \cup B) = v(A) + v(B) - v(A \cap B). \tag{108}$$

The integration premium relative to the sum is $-v(A \cap B) \leq 0$. Distinct records have additive value; overlapping records create redundancy. There can still be positive productive complementarity after a nonlinear task map $G(A)$ is applied. The raw information valuation and task value are separate functions.

For instance, let $G(A) = 1$ when a required pair of complementary records is present and zero otherwise. Each singleton can have zero task value while their union has value one. The join tensor remains idempotent, so this positive complementarity cannot hold as a strict superadditivity inequality for all pairs, especially repeated identical resources. A local opportunity can be real without supporting a global lax axiom.

D.6 Costs with fixed overheads

Let $C(A) = h + c|A|$ for nonempty A, with $C(\varnothing) = 0$, $h, c > 0$, and disjoint labeled bundles. Then

$$C(A) + C(B) - C(A \cup B) = h \tag{109}$$

for disjoint nonempty A, B. Integrating saves one fixed overhead. If there is also a convex coordination cost $b|A|^2$, the saving becomes

$$\sigma_C(A, B) = h - 2b|A||B|. \tag{110}$$

Small mergers can save cost while sufficiently large mergers increase it. A global oplax-cost hypothesis fails in that regime, although local economies of scope exist. This is a transparent finite mechanism for a nontrivial optimal size.

The formula also shows how a marginal cost estimate can be misleading. Observing savings for singleton mergers does not establish savings for large coalitions, because the interaction term grows with both block sizes. The source of the failure is explicit and can be tested in a specified model rather than described vaguely as bureaucracy.

E Organizational presentations and semantic invariance

E.1 A common integrated object has many presentations

An integrated object is a value of a diagram, not its complete provenance. Consider a fixed object $A \in \mathcal{X}$ and the constant diagram on A indexed by $[n] = \{0 \to 1 \to \cdots \to n\}$, with all structure maps identities. Since $[n]$ has a terminal object, its colimit is A. For every n, therefore, the same information object has a presentation with $n + 1$ labeled contributors.

For $n \geq 1$, the category $[n]$ is not equivalent as an infinity-category to the terminal category: it has noninvertible arrows and multiple inequivalent objects. Its underlying classifying space is contractible, but equivalence of classifying spaces does not imply equivalence of categories. Thus even the homotopy type of an index can hide ownership-relevant categorical structure.

One can assign distinct owner labels to the vertices or the same owner to all of them. The colimit calculation remains A. This gives a nondegenerate family of countermodels in which the integrated information object is fixed while the organizational presentation varies. It is stronger than relying only on the empty object as an example with many trivial coproduct presentations.

E.2 Changing presentations by final functors

If $q : I \to J$ is final and $D : J \to \mathcal{X}$ is a diagram, then

$$\operatorname*{colim}_I Dq \simeq \operatorname*{colim}_J D. \tag{111}$$

This is the standard cofinality principle for colimits [2, 6]. It explains how adding or removing certain redundant presentation data can leave the integrated semantics unchanged. Finality is a categorical condition on the relevant comma categories; it is not a synonym for a merger being economically final.

The equivalence in (111) is about universal receivers. It does not identify the number of servers, legal entities, principals, or failure domains. A presentation can be semantically redundant yet operationally useful for resilience or local control. Conversely, a semantically equivalent presentation can be more expensive to maintain. Economic invariance requires additional assumptions about how the cost functional behaves under the presentation change.

E.3 A forgetful map cannot reconstruct omitted labels

Let OrgInfo be any category of information diagrams decorated with ownership labels, and let U forget the labels and integrate the underlying diagram. If two decorated objects have the same underlying diagram but different ownership labels, U has the same value on them. Hence no procedure depending only on U can recover which labeling was present in every case.

This is a simple nonidentifiability statement. It does not require a deep impossibility theorem: omitted data are not determined by a function that ignores them. To study ownership from information semantics one needs a relation constraining which labels can occur, such as exclusive access, contractual rights, or an empirically estimated cost model. Those relations are the bridges exposed in the main theorem.

The same argument applies to permissions, physical deployment, and control rights. Two implementations can have equivalent query semantics while differing on each of these dimensions. A complete model should retain them as decorations or as separate fibers instead of expecting the bare colimit to reconstruct them later.

E.4 Internal truth and external numerical choice

Suppose an ordered ring object R is given in $\mathcal{X}$, with internal value and cost functions and a proposition $\Gamma(x, y) > 0$. If that proposition holds after an effective epimorphism, it holds globally. Formally, truth of a proposition is equivalence of its corresponding subobject of 1 with 1, and this equivalence is locally detected.

However, an internal real number can vary with context and need not be a single externally

known scalar. A constant-object functor can import the ordinary set of real numbers as a discrete object with algebraic structure; this object should not automatically be identified with every internal construction of real numbers, such as a sheaf of continuous real-valued functions. A model that needs continuity or completeness must specify the relevant internal number object and verify its properties.

The finite economic results in the body are external real-valued theorems. Their internal analogues can be formulated in a suitable internal ordered algebra, but that formulation is not necessary for the external conclusion. Keeping the two versions distinct avoids attaching an unproved internal analytic theorem to a finite combinatorial calculation.

E.5 Universal properties and implementation refinements

To connect semantic integration to a running system, one can define an implementation category $\mathcal{I}$ with an interpretation functor $\mathrm{Sem} : \mathcal{I} \to \mathcal{X}$. A proposed implementation M refines the integrated specification when a comparison

$$\mathrm{Sem}(M) \longrightarrow D_* \tag{112}$$

has the chosen correctness property, such as equivalence on permitted queries. Exact equivalence of entire objects is often stronger than required. A query family $\mathcal{Q}$ can specify a suitable observational equivalence instead.

Runtime, memory, latency, and communication costs are functions on implementations, not generally on semantic equivalence classes. Two programs can compute the same function with very different costs. To assign one cost to a semantic object, one must choose an implementation, take an infimum over implementations, or prove cost invariance under a selected equivalence. The categorical universal property does not make that choice.

This distinction also clarifies the phrase "elimination of duplicated operations." A colimit can represent a family of compatible maps by one semantic map. It does not imply that one physical execution is sufficient for every consumer, that network costs vanish, or that caches should be removed. Replication can be a deliberate implementation strategy even when the semantic transformation is unique.

F A restricted identification theorem and its boundaries

F.1 A finite skill theory that can be identified

The nonidentifiability results should be paired with a positive model. Let $\mathbb{F}_q$ be a finite field and let candidate skills be linear maps $f : \mathbb{F}_q^d \to \mathbb{F}_q^k$. A skill is determined by a $k \times d$ matrix A. Observing its outputs on the standard basis $e_1, \ldots, e_d$ yields all columns of A.

Proposition F.1 (Basis-query identification)**.** *For this candidate class, the trace*

$$T(A) = (Ae_1, \ldots, Ae_d) \tag{113}$$

is a bijection between the set of candidate skills and the set of ordered d-tuples of vectors in $\mathbb{F}_q^k$*. With discrete source and target categories it is fully faithful and essentially surjective.*

Proof. The jth observed vector is exactly the jth column of A, so it reconstructs the matrix uniquely. Conversely every ordered tuple defines a matrix and therefore a linear map. A bijection of discrete categories is an equivalence. □

The result depends on knowing that the unknown skill is linear. If arbitrary functions are admitted, basis observations no longer determine other inputs. A logger that merely samples common operational inputs may also fail to query a basis. This example shows exactly how a restricted operational theory can turn a strong fidelity condition into a valid finite theorem.

F.2 Rank characterizes a general query design

Arrange queries $x_1, \ldots, x_n$ as columns of a matrix $Q \in \mathbb{F}_q^{d \times n}$. The observed outputs form AQ. Then:

$$A \mapsto AQ \text{ is injective for all } A \quad \Longleftrightarrow \quad \operatorname{rank}(Q) = d, \tag{114}$$

provided $k \geq 1$.

If Q has full row rank, it has a right inverse R with $QR = I_d$, so $A = (AQ)R$. If its rank is smaller than d, choose a nonzero row vector v with $vQ = 0$. Let B be a nonzero matrix with one row equal to v and the rest zero. Then $BQ = 0$, so A and $A + B$ have identical observations. This proves both directions.

The proof identifies the unobserved operational subspace explicitly. More data points do not help if they remain in the same deficient span. Coverage of the candidate class is a structural property of the query design, not a consequence of the raw volume of events.

F.3 Noisy observations alter the theorem

If observations are $Y = AQ + Z$ with noise Z, the deterministic inverse formula becomes an estimator. Over real vector spaces one can use least squares when Q has full row rank:

$$\widehat{A} = YQ^\top (QQ^\top)^{-1}. \tag{115}$$

The estimation error is

$$\widehat{A} - A = ZQ^\top (QQ^\top)^{-1}. \tag{116}$$

Thus full rank is necessary for unique noiseless identification, while conditioning controls sensitivity to noise. A nearly singular query design can have exact semantic identifiability and poor practical recovery. This is a concrete instance of the distinction between an equivalence and a resource-bounded robust implementation.

The finite-field model and the real noisy model are different mathematical settings. The latter formula is included to explain sensitivity, not to assert that the finite-field proof automatically gives a real statistical error bound. Such a bound needs a norm and a noise model.

F.4 Composition can be observed separately from object identity

For linear maps with matching dimensions, matrix multiplication models sequential composition. Recording the matrix of each skill determines its composition law effectively. If instead a trace stores only the rank of each matrix, it loses substantial operational information. Two rank-one matrices can have different composites with a third matrix, even when their separate ranks coincide.

For example, over a field take row maps $a = (1, 0)$ and $b = (0, 1)$ from a two-dimensional space to a one-dimensional space. Both have rank one. Composing with the column $c = (1, 0)^\top$ gives $ac = 1$ and $bc = 0$. Rank alone therefore does not support correct compositional prediction. A target category that records only ranks cannot be declared a faithful skill theory for these operations without weakening the relevant notion of behavior.

This example illustrates why a category of traces needs arrows and composition, not only object labels. A feature that identifies a coarse property of each execution can fail to preserve how executions interact. Strong monoidal or categorical reconstruction hypotheses address such interactions, but they remain stronger requirements than storing informative features.

F.5 Task-relative quotients can make coarse traces complete

If the only relevant task is to determine matrix rank, the rank trace is complete for that discrete quotient of the candidate set. It is not complete for composition-sensitive linear algebra. There is no contradiction: the localization or quotient defining relevant behavior has changed.

One should nevertheless verify that the quotient supports the operations being claimed. Rank does not define a well-behaved composition rule on arbitrary linear maps, as the preceding example shows. A quotient adequate for a classification query need not be adequate for a compositional operational category. This is a general reason to define operational equivalence as a congruence or localization compatible with the desired structure, rather than a convenient similarity metric alone.

G Noisy organizational choice and stationary distributions

G.1 A finite stochastic alternative to deterministic consolidation

Let $\mathcal{P}$ be a finite set of feasible partitions with a connected undirected graph of reversible moves. Splits and mergers can both appear. Choose a symmetric proposal matrix $Q(P,R) = Q(R,P)$ with row sums one, and an inverse-noise parameter $\beta \geq 0$. Define the off-diagonal transition probabilities

$$K(P,R) = Q(P,R)\min\{1, e^{\beta(W(R)-W(P))}\}, \tag{117}$$

and assign the remaining row mass to $K(P,P)$. This uses the classical Metropolis acceptance construction [22]. Its application here is a mathematical choice model, not a claim about how actual firms make decisions.

Proposition G.1 (A stationary value-weighted law)**.** *The distribution*

$$\pi_\beta(P) = \frac{e^{\beta W(P)}}{\sum_{R\in\mathcal{P}} e^{\beta W(R)}} \tag{118}$$

is stationary for K. With a connected proposal graph and appropriate self-loop probability, the finite chain converges to it.

Proof. For distinct states,

$$\begin{aligned}\pi_\beta(P)K(P,R) &= Z^{-1}Q(P,R)\min\{e^{\beta W(P)}, e^{\beta W(R)}\}\\ &= \pi_\beta(R)K(R,P).\end{aligned}$$

This detailed balance identity implies stationarity after summing over P. Connectedness gives irreducibility because all proposed moves have positive acceptance at finite β. A positive self-loop makes the chain aperiodic; the standard finite-state convergence then follows. □

At finite β, every feasible partition has positive stationary probability. Even a unique optimum is not an absorbing state. This differs fundamentally from merger-only deterministic dynamics. The dynamic hypothesis, not just the static objective, determines whether consolidation is permanent, fluctuating, or merely more likely.

G.2 A finite concentration bound

Suppose P_* is the unique maximizer and the objective gap is

$$\Delta = \min_{P\neq P_*}\{W(P_*) - W(P)\} > 0. \tag{119}$$

Then

$$\pi_\beta(\mathcal{P}\setminus\{P_*\}) \leq (|\mathcal{P}|-1)e^{-\beta\Delta}. \tag{120}$$

Indeed, divide all weights by $e^{\beta W(P_*)}$. Every nonoptimal weight is at most $e^{-\beta\Delta}$, while the denominator is at least one. Thus as $\beta \to \infty$ the stationary law concentrates on the unique optimum. This is a parameter limit of distributions, not a theorem that every finite-noise trajectory eventually stays at the optimum forever.

If there are several maximizers, the same calculation concentrates mass on their set; with the symmetric construction their limiting weights are equal. Constraints can disconnect the graph, in which case the initial component matters and there need not be a unique global stationary law. Again, feasibility and dynamics cannot be inferred from the objective alone.

G.3 Why this appendix is useful

The stochastic model provides a controlled alternative to the strongest deterministic narrative. It retains a precise objective and a preference for higher value while allowing experimentation, error, and reversal. The comparison with rank-based merger termination shows which conclusions are robust across models and which are not. Finite state space alone does not imply absorption; a unique optimum alone does not imply permanent occupancy.

No claim is made that the parameter β is directly measurable in an organization, or that a Metropolis rule is behaviorally accurate. The model's role is a counterexample and a possible formal starting point. An empirical dynamic theory would need evidence for its transition probabilities and for the stability of the underlying objective.

H Technical concordance and notation

H.1 How the motivating construction is strengthened

The source note [1] is the conceptual starting point. The table records short topic correspondences; the substantive arguments and corrections are developed in the cited sections of this manuscript.

Source topic	Treatment in this manuscript
Organizational contexts	A fixed base and an actual effective cover are required (section 2).
Skill localization	Existence, size, and context compatibility are explicit assumptions (section 3).
Truncated skill theory	Categorical truncation is distinguished from levelwise object truncation (counterexample 3.4).
Operational identity	Equivalence objects are separated from general mapping objects ((11)).
Local reconstruction	Detection of a global functor is separated from construction by descent (section 5).

Source topic	Treatment in this manuscript
Universal integration	Compatibility and the chosen interfaces remain in the antecedent (section 7).
Ownership assignment	Object labels need not be functors into a discrete owner category (section 2).
Monoidal value	Tensor, unit, copyability, and valuation scope are declared (section 10).
Strict merger gains	They are sufficient, rather than necessary, for a unique top optimum (section 11).
Constrained endpoints	Multiple sinks need not be globally optimal (section 12).
Recursive learning	Iteration, improvement, convergence, and reachability are separate properties (section 15).
Limiting integration	A varying family needs specified transport maps and a limit construction (section 16).

The intention is to preserve the source's useful separation of semantic and economic claims while making the bridges precise enough to audit. The result is a conditional theory with stronger explanations of its own limits, rather than a claim that every motivating interpretation has been established.

H.2 Principal notation

Symbol	Meaning
$\mathcal{X}$	A fixed infinity-topos, or a declared slice of one.
B, U, U_n	Base context, effective covering object, and Cech nerve levels.
$W, \mathcal{O}$	State-operation category and chosen execution category.
$S, q, \mathcal{K}$	Inverted arrows, localization, and skill category.
$T, F, \mathcal{D}$	Original trace, factored trace, and trace category.
$\underline{\mathrm{Map}}, \underline{\mathrm{Eq}}$	Internal mapping and equivalence objects.
$D : I \to \mathcal{X},\ D_*$	Information diagram and its colimit.
$q_!$	Left Kan extension along an indexing functor.
$\mathcal{A} \to \mathcal{D}$	Authorized operations and the permission-forgetting map.
$\mathcal{R}, \otimes, \mathbf{1}$	Resource category, tensor, and unit.
V, C, N	Value, cost, and net value of resources.
$E, r_A, n(A)$	Capability set, coalition resource, and coalition net value.
$P, \mathrm{Part}(E), \top$	Ownership partition, partition poset, and one-block partition.
$W(P), H(P)$	Organizational objective and partition-dependent external cost.

Symbol	Meaning
Γ, γ	Resource-level and coalition-level binary integration gains.
Φ, ρ	State update and coarsening rank.
$K(y \mid x), R(K)$	Observation channel and its optimal Bayes risk.
$I(X;Y), H(X)$	Mutual information and entropy under a declared joint law.

The letter q is used both for a localization and, in the Kan-extension section, for an indexing functor. Their domains distinguish them. The scalar contraction constant in proposition 15.3 is also denoted q locally. These notations never occur in the same typed expression. A resource category is denoted $\mathcal{R}$, while $\mathbb{R}$ denotes the real numbers.

H.3 Proof dependency guide

The trace factorization needs only the assumed localization universal property. Global full faithfulness needs a global functor, an effective epimorphism, and local full faithfulness. Construction of that functor from local pieces additionally needs complete descent data. Integration needs a small diagram and cocompleteness. The economic theorem needs a finite nonempty capability set, a declared feasible space, an objective, and the stated gain conditions.

The computability obstructions do not enter the proof of effective descent. The statistical examples do not establish or refute the topos axioms. The idempotence obstruction concerns a chosen tensor and scalar valuation, not every possible economics of data. The dynamic contraction criterion uses metric hypotheses absent from the categorical reconstruction theorem. Keeping this dependency structure visible prevents a valid theorem in one layer from being used as evidence for an unproved antecedent in another.

I Prior work and the scope of attribution

I.1 Higher categories and topoi

Lurie's *Higher Topos Theory* [2] is the foundational reference for infinity-topoi, effective descent, universal colimits, and higher categorical localization and cofinality. The paper uses these established structures. Its organizational interpretation, selected countermodels, and finite economic arguments are not asserted to be claims made in that book. Readers should consult the chapters on presentable categories and infinity-topoi for the ambient machinery, rather than infer that this manuscript proves those foundations.

Rezk's complete Segal spaces [3] supply a model in which object equivalence and coherent composition are correctly represented. We use complete Segal objects internally to a topos. Martini and Wolf [4] develop internal adjunctions, colimits, and Kan extensions in that setting. Their work supports the internal categorical language; our assumption about the particular

operational localization remains explicit. Riehl and Verity [6] provide a complementary model-independent account of higher categorical constructions and Kan extensions.

Lurie's *Higher Algebra* [5] provides the general setting for symmetric-monoidal infinity-categories and lax structures. Our real-valued comparisons use only a small fragment of that machinery. The economic inequalities are assumptions in the chosen ordered target, not consequences of the existence of an infinity-operad. In particular, the distinction between strong and lax monoidal functors must survive the translation into economic language.

I.2 Applied categories and data

Spivak's functorial data migration [7] explains the restriction–left extension–right extension adjunction pattern for categorical database instances. Our treatment of organizational coarsening as Kan extension applies that pattern to a different interpretation and explicitly notes that a discrete grouping only produces a coproduct. The interface and identity-resolution data still have to be supplied.

Fong's decorated cospans [8] provide a compositional treatment of open systems with shared boundaries. Fong and Spivak [9] present broader applied categorical methods, including resource-oriented reasoning. These references motivate the use of stable interfaces and compositional models; they do not establish the specific empirical claim that a federation is cheaper or more expensive than a merger.

Coecke, Fritz, and Spekkens [10] develop a general mathematical theory of resources and convertibility. We borrow that structural viewpoint while distinguishing valuations on the core from monotones on all free conversions. The copyability arguments in this manuscript show why the distinction matters for the proposed combination of positive value and global superadditivity.

I.3 Statistics, information, and privacy

Blackwell's comparison of experiments [12] is the classical source for decision-based comparisons of information. The body proves the elementary garbling direction directly. Fritz [11] places probability and sufficiency in a categorical framework. Our finite channel calculations are a concrete specialization, and no general disintegration theorem in an arbitrary topos is claimed.

Cover and Thomas [13] provide the information identities underlying the redundancy and XOR examples. The paper derives the particular inequalities it uses so that the sign of synergy can be checked independently. Dwork and Roth [14] provide the standard privacy definitions and background. The recovery tradeoff here is a short direct application to an adjacent pair, not a replacement for the broader theory of private data analysis.

Rice [15] supplies the classical obstruction to uniform decidable classification of nontrivial extensional program properties. Our use is narrowly stated: a total computable canonical trace with decidable equality cannot classify arbitrary partial-function semantics exactly. We

do not infer that no useful program analysis or finite restricted identification theorem is possible.

I.4 Economics and organizational boundaries

Coase [16] makes organizational boundaries an economic question rather than an automatic consequence of technical possibility. Grossman and Hart [17] analyze ownership through incomplete contracts and control rights. Our partition objective is far simpler than those theories. It can summarize some of their consequences in a specified instance, but it contains no general contract or bargaining model by itself.

Milgrom and Roberts [18] develop economic reasoning about complementary decisions and organization. Our cross-partial calculation is a smooth local comparison, and the paper explicitly separates it from global bundle superadditivity. Jones and Tonetti [19] analyze the nonrival economics of data. Their work motivates taking reuse seriously; it does not justify assigning positive superadditive value to an idempotent information tensor.

The finite partition and graph results in the manuscript are proved directly. They should be read as transparent sufficient conditions and countermodels, not as a new general theory of industry equilibrium. The exact optimizer, the improving path, the stable sink, and the distribution of gains are different objects of explanation.

I.5 Surveillance and contextual governance

Haggerty and Ericson [20] analyze surveillance through the idea of a surveillant assemblage, which helps situate a distributed observational architecture beyond a single physical watchtower. Nissenbaum [21] makes context and norms of information flow central to privacy analysis. Our policy fibration is a mathematical modeling proposal informed by that distinction; neither source is presented as proving a topos-theoretic result.

The historical vocabulary is therefore used with restraint. A fully faithful trace functor is a technical relative property. It does not settle whether an institution's observation is legitimate, fair, coercive, or socially desirable. Those questions require explicit normative premises and evidence about the institution, rather than an inference from semantic completeness.

I.6 Conceptual provenance and novelty

The web note [1] supplied the initiating architecture and the separation between informational and corporate consolidation. This manuscript does not reproduce it as an independently refereed source. The extended work develops the technical assumptions, adds statistical and computational obstructions, identifies the idempotence incompatibility, and supplies detailed finite optimization and feedback analyses. All named foundational ideas are attributed to the prior literature, while the worked examples and elementary deductions are presented with

their own proofs.

No exhaustive novelty claim is made for every elementary lemma. Many are immediate consequences of definitions and standard tools. Their value here is their combined role in auditing a broad argument about AI organizations. A publication review can assess that synthesis without being asked to accept a claim that category theory alone predicts a particular corporate future.

J A context-dependent end-to-end countermodel

J.1 The ambient informational universe

The following model connects the observational and economic layers in one explicit setting. Let

$$\mathcal{X} = \mathcal{S} \times \mathcal{S}. \tag{121}$$

It is the infinity-topos of space-valued sheaves on a discrete two-point space. Call the two contexts North and South. Its terminal object is $(1, 1)$, and the objects

$$U_N = (1, \varnothing), \qquad U_S = (\varnothing, 1) \tag{122}$$

form an effective cover: their coproduct is the terminal object itself. Limits, colimits, and equivalences are computed componentwise. Although the example has no nonempty cross-context overlap, it makes the distinction between local truth, global internal truth, and a context-independent numerical choice completely explicit.

Let $\mathcal{L}$ be the finite ordinary category whose objects are the sixteen 2×2 matrices over $\mathbb{F}_2$. An arrow $P : A \to B$ is a 2×2 matrix satisfying

$$PA = BP. \tag{123}$$

Identity matrices are identity arrows. If $P : A \to B$ and $Q : B \to C$, then

$$(QP)A = Q(BP) = (QB)P = C(QP), \tag{124}$$

so matrix multiplication defines composition. Use its associated infinity-category with a complete Segal presentation. This is a genuine category with nontrivial transformations, not merely a set of skill names.

J.2 The experimentally described trace category

An object trace records the two basis outputs (Ae_1, Ae_2). A transformation trace records the basis outputs of P, subject to the relation (123) reconstructed from the object traces. Let $\mathcal{T}$

be the resulting finite category of these records, with composition reconstructed as matrix multiplication. The trace functor

$$F_0 : \mathcal{L} \longrightarrow \mathcal{T} \tag{125}$$

is an isomorphism of ordinary categories: every object matrix and every arrow matrix is uniquely recovered from its two columns, and the defining relation and composition are preserved in both directions.

The internal categories $\mathcal{K} = (\mathcal{L}, \mathcal{L})$ and $\mathcal{D} = (\mathcal{T}, \mathcal{T})$ therefore admit a fully faithful internal trace functor $F = (F_0, F_0)$. One can take S to be the already invertible arrows and q to be the identity localization. Alternatively, an execution presentation with redundant names could be localized to this category, but no such extra presentation is needed for the countermodel.

The construction has a substantive identification assumption: candidate operations are known to be linear and the basis experiments are available. If those assumptions are dropped, its fidelity argument does not apply. Within the stipulated class, however, the inverse is explicit and finite. This shows that the countermodel does not depend on an uncomputable ideal trace.

J.3 Integration of a declared interface

In each component take a span of discrete spaces

$$\{a_1, a_2\} \longleftarrow \{k\} \longrightarrow \{b_1, b_2\}, \tag{126}$$

where k maps to a_2 and b_1. Its colimit is the three-point object described in section 7. Taking the pair of these spans gives a diagram internal to $\mathcal{X}$ whose integrated object is (D_*, D_*). Compatible output queries extend componentwise and therefore extend internally.

Assign distinct owners to the two legs if desired. These labels do not change the calculation. Thus exact observation and explicit compatible integration are both present. The economic comparison can now be varied without changing either categorical construction.

J.4 Two locally different organizational optima

Let $E = \{1, 2, 3, 4\}$ and set all singleton values to one. In North give all six pair interactions weight $+1$. In South set $w_{12} = w_{34} = 3$ and every cross-group weight to -2. Let W_N and W_S be the corresponding partition objectives. The previous proofs give

$$\operatorname{argmax} W_N = \{\top\}, \tag{127}$$

$$\operatorname{argmax} W_S = \{12|34\}. \tag{128}$$

North has a unique complete-consolidation optimum; South has a unique two-block optimum. Both contexts have the same trace fidelity and the same integration specification.

The pair (W_N, W_S) is an internal real-valued objective on the internal partition object. Its maximizing global section is the context-dependent pair

$$P_* = (\top, 12|34). \tag{129}$$

There is no single fixed external partition that is optimal in both contexts, since the two unique maxima differ. Internal global optimality is not the demand that one external constant value be chosen everywhere. A section of a constant sheaf on a disconnected space may vary between components.

J.5 What local economic truth actually implies

For a fixed merger, the internal proposition that its gain is positive holds globally exactly when it holds in both components. A North-only positive gain does not establish that proposition globally, because U_N alone is not a cover of $(1, 1)$. Positivity established separately on both members of the effective cover does descend, exactly as the general locality principle states.

This is a concrete check on the phrase "locally favorable integration." One must specify whether all covering contexts verify the same internal inequality, whether different contexts choose different mergers, or whether only an observed subdomain has been checked. Those assertions have different logical forms. The example gives the latter distinctions an elementary componentwise interpretation.

The end-to-end model has exact and computable trace reconstruction, effective local detection, a universal integrated information object, and coherent context-dependent organizational choices. It still has no theorem selecting one common ownership partition from the observational data. The economic values supply that selection, and their contextual variation remains visible in the internal semantics.

References

[1] Anonymous web note. *Panoptical Observability and Informational Consolidation in Infinity-Topoi*. Undated. Accessed November 5, 2026. https://panoptical-observability.grok.me/. Conceptual starting point; no independent review status is asserted.

[2] Jacob Lurie. *Higher Topos Theory*. Annals of Mathematics Studies 170. Princeton University Press, 2009. Open manuscript: https://arxiv.org/abs/math/0608040.

[3] Charles Rezk. A model for the homotopy theory of homotopy theory. *Transactions of the American Mathematical Society* 353 (2001), 973–1007. https://doi.org/10.1090/S0002-9947-00-02653-2. Open version: https://arxiv.org/abs/math/9811037.

[4] Louis Martini and Sebastian Wolf. Colimits and cocompletions in internal higher category theory. *Higher Structures* 8(1) (2024), 97–192. https://doi.org/10.21136/HS.2024.03. Open version: https://arxiv.org/abs/2111.14495.

[5] Jacob Lurie. *Higher Algebra*. Manuscript, September 18, 2017. https://www.math.ias.edu/~lurie/papers/HA.pdf.

[6] Emily Riehl and Dominic Verity. Infinity category theory from scratch. *Higher Structures* 4(1) (2020), 115–167. https://doi.org/10.21136/HS.2020.04. Open version: https://arxiv.org/abs/1608.05314.

[7] David I. Spivak. Functorial data migration. *Information and Computation* 217 (2012), 31–51. https://doi.org/10.1016/j.ic.2012.05.001. Open version: https://arxiv.org/abs/1009.1166.

[8] Brendan Fong. Decorated cospans. *Theory and Applications of Categories* 30(33) (2015), 1096–1120. https://arxiv.org/abs/1502.00872.

[9] Brendan Fong and David I. Spivak. *An Invitation to Applied Category Theory: Seven Sketches in Compositionality*. Cambridge University Press, 2019. Open manuscript: https://arxiv.org/abs/1803.05316.

[10] Bob Coecke, Tobias Fritz, and Robert W. Spekkens. A mathematical theory of resources. *Information and Computation* 250 (2016), 59–86. https://doi.org/10.1016/j.ic.2016.02.008. Open version: https://arxiv.org/abs/1409.5531.

[11] Tobias Fritz. A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics. *Advances in Mathematics* 370 (2020), article 107239.

https://doi.org/10.1016/j.aim.2020.107239. Open version: https://arxiv.org/abs/1908.07021.

[12] David Blackwell. Equivalent comparisons of experiments. *The Annals of Mathematical Statistics* 24(2) (1953), 265–272. https://doi.org/10.1214/aoms/1177729032.

[13] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*, second edition. Wiley, 2006. https://doi.org/10.1002/047174882X.

[14] Cynthia Dwork and Aaron Roth. The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science* 9(3–4) (2014), 211–407. https://doi.org/10.1561/0400000042. Open author copy: https://www.cis.upenn.edu/~aaroth/Papers/privacybook.pdf.

[15] Henry G. Rice. Classes of recursively enumerable sets and their decision problems. *Transactions of the American Mathematical Society* 74(2) (1953), 358–366. https://doi.org/10.1090/S0002-9947-1953-0053041-6.

[16] Ronald H. Coase. The nature of the firm. *Economica*, New Series, 4(16) (1937), 386–405. https://doi.org/10.1111/j.1468-0335.1937.tb00002.x.

[17] Sanford J. Grossman and Oliver D. Hart. The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of Political Economy* 94(4) (1986), 691–719. https://doi.org/10.1086/261404.

[18] Paul Milgrom and John Roberts. The economics of modern manufacturing: Technology, strategy, and organization. *American Economic Review* 80(3) (1990), 511–528. Stanford GSB publication record.

[19] Charles I. Jones and Christopher Tonetti. Nonrivalry and the economics of data. *American Economic Review* 110(9) (2020), 2819–2858. https://doi.org/10.1257/aer.20191330.

[20] Kevin D. Haggerty and Richard V. Ericson. The surveillant assemblage. *The British Journal of Sociology* 51(4) (2000), 605–622. https://doi.org/10.1080/00071310020015280.

[21] Helen Nissenbaum. Privacy as contextual integrity. *Washington Law Review* 79(1) (2004), 119–158. https://digitalcommons.law.uw.edu/wlr/vol79/iss1/10/.

[22] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics* 21(6) (1953), 1087–1092. https://doi.org/10.1063/1.1699114.