Predicate Invention Under Sheaf Constraints

Mathematical Foundations for Compositional Discovery

John Komkov*

March 9, 2026

Abstract

We study the problem of *predicate invention under sheaf constraints* (SCPI): given a Grothendieck site (C, J) equipped with a pseudofunctor of model groupoids $M : C^{\mathrm{op}} \to \mathbf{Grpd}$, when can a new predicate—specified locally on a cover and constrained by data—be extended to a global predicate that is (i) compatible with the local specifications (descent), (ii) conservative over the base signature, and (iii) explicitly definable from the base vocabulary?

We show that the obstruction to predicate invention decomposes into three distinct sequential components: a *topological obstruction* classified by non-abelian Čech cohomology $H^1(C, \mathrm{Equiv}_{\mathrm{ext}})$, a *model-theoretic obstruction* detecting non-conservativity via descent of the expansion property, and a *definability obstruction* measuring the gap between implicit and explicit definability over the site. Each obstruction is detected by different mathematical machinery and provides different diagnostic information. In particular, the overlap topology of information sources is a computable invariant that predicts whether concept alignment across autonomous agents is achievable—connecting classical sheaf theory to problems in AI interpretability and multi-agent coordination, under explicit modeling assumptions (M1–M4, Remark 1.4).

The extension torsor lemma and conservativity descent theorem are proved under Assumption D (effectivity of descent) and stated amalgamation hypotheses; Assumption D is itself proved for finite relational sites, the regime covering all worked examples. A full end-to-end computation of H^1 for a three-source site (Calendar/Email/Slack) demonstrates the complete diagnostic chain from site definition through obstruction identification to architectural prescription. A generalization of Beth's definability theorem to sites is proved unconditionally for geometric theories (assembling results of Kreisel, Johnstone, and the DD1 regime) and conditionally for general first-order theories. A topological invariance theorem shows that the obstruction landscape depends only on the Čech nerve and coefficient group, enabling cross-domain transfer of diagnostics between structurally isomorphic sites. We construct a schema discovery functor compatible with Spivak's functorial data migration adjunctions, with an explicit proof of compatibility for finite sites.

Selected results are formalized in Lean 4 (v4.24.0); see the companion repository for details.

*Selected Lean proofs verified with the assistance of Aristotle (Harmonic).

Contents

1 Introduction

1.1 The lifting problem

Every system that must extract structured knowledge from unstructured sources confronts the same task: a query implies a schema that does not yet exist [3], and the schema must be *discovered* from the source material before it can be populated. The schema is not retrieved. It is not pre-configured. It is *invented*—and then it must be validated against the source under compositional constraints.

We develop three examples at escalating difficulty. These are not toy problems; they recur as running examples throughout the paper, and every theorem is instantiated against at least one of them.

Example 1.1 (Static corpus)**.** "How many types of fruit are mentioned in the Bible and how many times is each type mentioned?" The implied predicates `is_fruit`(x) and `mentioned_at`$(x, \text{book}, \text{chapter}, \text{verse})$ exist in no index. The source decomposes into 66 books, each into chapters, each into verses—a hierarchical cover. The predicate `is_fruit` must be invented locally in each book and then verified for global coherence: "pomegranate" in Song of Solomon and "pomegranate" in Exodus must refer to the same entity, classified the same way.

Example 1.2 (Dynamic corpus)**.** "How many meetings with external parties happened in my organization last week that involved more than 2 internal attendees?" The source material comprises emails, calendar invites, Slack messages, and Zoom logs—overlapping views of the same underlying events. A meeting may appear as a calendar entry, a confirmation email, a Slack thread, and a Zoom recording. The views overlap but do not partition. The predicate `is_meeting`(e) must be reconciled across fundamentally different ontologies, with entity resolution (e.g., `john@co.com` = "John K" = `@jkomkov`) that must compose consistently around the triangle of pairwise overlaps.

Example 1.3 (Open corpus)**.** "Find me hotel options near Berkeley for a Cal home game this fall, subject to constraints $C_1, \ldots, C_n$." The source is the open web, which cannot be scanned exhaustively. The agent must simultaneously choose which sites to crawl (the cover), invent the join schema, perform extraction, and verify coherence—all under a budget that makes the optimal cover problem NP-hard (Theorem 7.2).

1.2 Why this matters beyond logic: interpretability and agent coordination

The SCPI framework addresses a problem of growing interest in AI systems research: *how do autonomous agents, each with partial views of a domain, arrive at shared concepts?*

Interpretability. A neural network trained on medical images may "discover" an internal feature that correlates with disease severity—but this feature exists only within the network's latent space. A second network, trained on patient records, may discover a related but distinct feature. Are these the same concept? Can they be reconciled? This is precisely a predicate invention problem: each network invents a local predicate (its internal feature), and the question is whether these local predicates glue to a coherent global concept. The obstruction theory developed here provides a rigorous diagnostic:

- If the topological obstruction (H^1) is non-trivial, the local features are *fundamentally incompatible*—no choice of alignment can make them agree globally. This is a structural impossibility, not a training failure.
- If the model-theoretic obstruction is non-trivial, the features glue but the combined concept introduces new logical consequences not present in either network's individual "theory." The merged representation is not conservative.
- If the definability obstruction is non-trivial, the concept is "there" (implicit in the combined data) but cannot be expressed in the shared vocabulary. The agents agree on what they see but lack the language to say it.

Each failure mode has a different fix. The framework tells you *which* fix to apply—or whether no fix exists.

Agent–agent coordination. When multiple AI agents share information to complete a task (e.g., a retrieval agent, a reasoning agent, and a verification agent), they must agree on the meaning of intermediate predicates. The structure of their information overlap determines whether agreement is possible. A hierarchical architecture (tree-structured information flow) has contractible nerve—under locally constant coefficients and Assumption D, $H^1 = 0$, so local agreements always globalize (Corollary 3.5). A peer-to-peer architecture (circular information flow) can have non-trivial H^1, creating irreconcilable disagreements that no amount of "negotiation" can resolve without changing the cell structure of the nerve (e.g., adding higher-dimensional overlaps).

A concrete interpretability instance. Consider two models: a vision model V and a text model T, probed on a shared evaluation set E of medical cases. The "cover" is $\{V, T\}$; the overlap $V \cap T$ consists of cases in E where both models produce confident predictions. Each model invents a local predicate ("high-severity") on its domain. On the overlap, the two predicates may agree or disagree. With two contexts and one overlap, the nerve is Δ^1 (contractible)—so $H^1 = 0$ and alignment is always achievable by local adjustment. Now add a third model G (genomic risk). The cover is $\{V, T, G\}$ with pairwise overlaps (cases where two models are confident) but possibly no effective triple overlap (no case where all three are simultaneously confident). The nerve may be S^1, and the SCPI diagnostic applies: compute the cocycle, check if it's a coboundary. If not, no alignment is possible without adding a shared evaluation set that creates a triple overlap. This is a testable prediction.

Diagnostic summary. The overlap topology of your information sources is a *computable invariant* that predicts whether concept alignment is achievable. Compute the Čech nerve of your agent architecture. If it's contractible, the coefficient sheaf is locally constant, and Assumption D holds (as proved for finite relational sites in Lemma 2.19), then Gate 1 cannot fail: $H^1 = 0$ (Corollary 3.5). Any obstruction must be downstream—a conservativity or definability issue, not a topological impossibility. If the nerve is *not* contractible, the non-trivial H^1 class tells you exactly where the obstruction lives and what structural change (adding a shared data source, restructuring the agent graph) would eliminate it.

Empirical instantiation. The companion paper [43] empirically instantiates the three-gate sequence in an abelian linearized regime: three LLM agents operating against three database schemas on a coordination graph with one independent cycle. Gate 1 is measured directly ($\dim H^1 = 2$, two independent blind-spot dimensions invisible to bilateral validation). A closed-loop repair experiment demonstrates a failure straddling Gates 2 and 3: frontier LLMs correctly identify the missing bridge predicates (passing Gate 1) but propose specifications that canonize one agent's existing local policy rather than introducing a witnessable shared predicate—a non-conservative extension (Gate 2) expressed in natural-language ambiguity rather than explicit formulas (Gate 3). The minimum number of 2-cells required to kill H^1—the *coherence fee* of [43]—is a computable topological invariant; in the abelian regime, it equals $\dim H^1$ of the interpretation sheaf on the coordination graph. In the non-abelian regime, the obstruction is a pointed set and the fee is the minimum number of nerve corrections required to trivialize it. The SHEAF protocol [44] builds the diagnostic instrument and economic mechanism for pricing and procuring the minimum repair.

Remark 1.4 (Modeling assumptions for AI applications). The formal framework applies unconditionally to any setting that instantiates a predicate site (C, J, M). The data-integration instances (Examples 1.1–1.3; the Calendar/Email/Slack computation of Section 3.1) are fully formal: finite poset sites with relational signatures. The neural-network interpretability and agent-coordination applications discussed above require four modeling assumptions that are each nontrivial and currently at different levels of empirical support.

(M1) Features as predicates. Internal neural features must approximate logical predicates on a shared domain. Sparse autoencoders [32, 33] find that $\sim$70% of extracted features are monosemantic and behave like unary predicates, but the remaining features exhibit multi-dimensional structure—circular representations of periodic concepts [34]—or context-dependent polysemanticity. Moreover, SAE reconstructions recover a fraction of total model compute [35], and different training runs yield different decompositions. The predicate approximation holds in the *monosemantic regime*; multi-dimensional features are better modeled as sections of fiber bundles, suggesting that sheaf theory may be the right formalism precisely where the predicate assumption is weakest.

(M2) Restriction maps with algebraic structure. Cross-model representation comparison must produce morphisms that compose, not merely scalar similarity scores (CKA, probing accuracy). Model stitching [37] and universal sparse autoencoders [38] provide maps between representation spaces with the right structural properties, but *composition has not been verified* and the sheaf gluing axiom has never been checked for cross-model maps. This is the most demanding assumption. Seely et al. [39] have computed sheaf cohomology *within* single predictive-coding networks, finding that non-vanishing cohomology characterizes irreducible inference errors—establishing operational meaning for sheaf-cohomological obstructions in neural settings—but the cross-model case remains open.

(M3) Identifiable symmetry groups. The coefficient group $\mathrm{Equiv}_{\mathrm{ext}}$ must be computable. Identifiability theory provides precise guarantees: under sparsity conditions, latent variables are recoverable up to permutation and component-wise transformation [40], and neuron permutation symmetries are completely characterized for standard architectures. Feature splitting—where decompositions change qualitatively with dictionary width—requires categorical structure (spans, filtered colimits) beyond standard group actions.

(M4) Site structure for agent systems. Multi-agent coordination must admit a Grothendieck topology. Schmid [41] articulates this research program for multi-agent reinforcement learning; Gavranović et al. [42] provide the categorical vocabulary (representing architectures as monad algebra homomorphisms). No published work has yet computed sheaf-cohomological obstructions to coordination in empirical multi-agent neural systems.

These assumptions should be read as *interface specifications*: the paper's diagnostic machinery applies whenever M1–M4 are instantiated, and the assumptions identify precisely what empirical validation is required. The Platonic Representation Hypothesis [36]—that model representations converge with scale—suggests that M2 may become easier to satisfy as models grow, while the non-trivial cohomology regime that motivates SCPI is precisely the current one where alignment is partial. The contribution of the present paper to the AI-applications context is not to validate M1–M4 but to identify them sharply and to show that *if* they hold, the diagnostic decomposition is a theorem.

1.3 The common structure

These problems share a mathematical structure independent of the source medium:

(1) A *query* implies a *schema* that does not yet exist.

(2) The schema must be *discovered* (predicate invention [9]).

(3) The source decomposes into *overlapping contexts* (a cover of a site).

(4) Discovery is applied *locally* in each context.

(5) Local extractions must *cohere* globally (the sheaf condition).

(6) Extraction has a *cost* that must be bounded.

The thesis formalizes this structure.

1.4 Contributions and honest delineation

We summarize the main results with explicit labels indicating their status:

[**Spine**] Extension Torsor Lemma (Lemma 3.1): the obstruction to gluing local extensions is a class in $H^1(C, \mathrm{Equiv}_{\mathrm{ext}})$.

[**Spine**] No-Go for fixed-topology alignment (Corollary 3.2): when $H^1 \neq 0$, no protocol operating within the fixed overlap structure can achieve global concept alignment.

[**Spine**] Assumption D for finite relational sites (Lemma 2.19): effectivity of descent is proved (not merely assumed) for the regime covering all examples in this paper.

[**Spine**] Conservativity Descent (Theorem 4.3): conservativity is local under a model-amalgamation hypothesis, with a finite counterexample when the hypothesis fails.

[**Spine**] Obstruction Decomposition (Theorem 5.1): in World B, three distinct sequential obstructions exhaust the failure modes.

[**Spine**] — **sketch** Schema Discovery compatibility (Proposition 8.3): Inv extends Spivak's data migration adjunctions, with construction and proof sketch for finite sites.

[**Spine**] Full worked computation (Section 3.1): H^1 computed end-to-end for the Calendar/Email/Slack site, diagnosing both resolvable and irreconcilable cases.

[**Spine**] Beth for Geometric Sites (Theorem 6.2): implicit definability implies explicit definability over sites with geometric theories—proved unconditionally (assembling Kreisel, Johnstone, and DD1).

[**Conditional**] Beth for Sites, general (Theorem 6.4): extends to non-geometric theories under H1/H3.

[**Conjectural**] Schema Discovery universality (Conjecture 8.4): Inv is the universal schema discovery functor.

Hypothesis regime. The following table summarizes which hypotheses each spine result requires:

Result	**Assump. D**	**Thin fibers**	**DD1/DD2**	**Model-cons.**
Torsor Lemma (3.1)	yes	—	yes	—
Assump. D for fin. rel. (2.19)	*proved*	—	—	—
Conserv. Descent (4.3)	yes	yes	—	yes (local)
Obstruction Decomp. (5.1)	yes	—	yes	—
Independence (5.3)	—	—	—	—
Schema Compat. (8.3)	—	—	—	—
Beth-geom (6.2)	—	—	DD1 (auto)	—
Beth-general (6.4)	—	—	H1/H3	—

Reader's guide. Readers primarily interested in the AI applications may begin with the worked example (Section 3.1) and the topological invariance theorem (Theorem 5.6), referring back to Section 2 as needed. Readers interested in the model-theoretic foundations should read linearly from Section 2 through Section 6.

Paper map. Sections 2–5 constitute the mathematical core (definitions, torsor lemma, conservativity descent, diagnostic decomposition). Section 6 extends Beth's theorem to sites. Sections 7–8 develop computational and applied consequences. Section 9 describes formalization.

Positioning. *Not new*: interpreting theories in sheaf toposes (Caramello [2]; Johnstone [7]; Makkai–Reyes [8]), functorial data migration (Spivak [10]), effective descent for pretoposes via definability (Makkai [17]; Zawadowski [18]; Ballard–Boshuck [16]). *New*: treating predicate invention as a descent problem for extensions of a model stack *over a specific Grothendieck site*, with a computable diagnostic decomposition (the three-obstruction theorem), a bridge to schema discovery, and applications to AI interpretability and multi-agent coordination (Section 1.2). The architecture of the paper—descent = definability + covering—is validated by Ballard–Boshuck [16], who prove that categorical descent theorems for pretoposes decompose into a Beth/Tarski definability component plus a covering theorem. Our contribution is to instantiate this decomposition on concrete Grothendieck sites with a computational diagnostic.

2 Objects: Sites, Model Stacks, Extension Stacks

2.1 Predicate sites

Definition 2.1 (Theory)**.** A *first-order theory* $T = (\Sigma, \mathrm{Ax})$ consists of a signature $\Sigma = (S, F, R)$—sorts, function symbols, relation symbols—and a set of axioms consisting of sentences over Σ.

Definition 2.2 (Model stack — Frozen Definition 1)**.** Let (C, J) be a Grothendieck site. A *model stack* over (C, J) [6] is a pseudofunctor

$$M : C^{\mathrm{op}} \to \mathbf{Grpd}$$

assigning to each object $U \in C$ a groupoid $M(U)$ of Σ-structures (models), with restriction functors $M(f)\colon M(U) \to M(V)$ for each morphism $f\colon V \to U$, satisfying pseudofunctor coherence: $M(g \circ f) \cong M(g) \circ M(f)$ via coherent natural isomorphisms.

The *stack condition* (effective descent for models) is an explicit hypothesis when needed, not assumed globally.

Definition 2.3 (Predicate site)**.** A *predicate site* is a triple (C, J, M) where (C, J) is a Grothendieck site and $M\colon C^{\mathrm{op}} \to \mathbf{Grpd}$ is a model stack.

Remark 2.4 (Why models, not theories)*.* The primitive is M (models/structures), not a presheaf of theories. Sheaf conditions are about sections; in logic, the natural "sections" are models. Descent for models (amalgamation) is well-posed; descent for axiom sets is not without extra machinery. The syntactic counterpart $\mathrm{Th}(U)$ = common theory of $M(U)$ is derived.

Remark 2.5 (Strictification)*.* M is a *pseudofunctor*: restrictions compose up to coherent natural isomorphism, $M(g \circ f) \cong M(g) \circ M(f)$. In the Lean formalization, we work with a strict functor (equality, not isomorphism).

What is strictified: the composition law for restriction functors ($M(g \circ f) = M(g) \circ M(f)$, on the nose) and the identity law ($M(\mathrm{id}_U) = \mathrm{id}_{M(U)}$). This is justified by Mac Lane's coherence theorem for bicategories: every pseudofunctor from a small category to **Grpd** is equivalent to a strict 2-functor.

What is not claimed: we do not claim that the strictification preserves all 2-categorical structure. In particular, we do not strictify the coefficient groupoid $\mathrm{Equiv}_{\mathrm{ext}}$ (which remains a 1-group in the spine), nor do we strictify the descent data (which is isolated in Assumption D). The theorems are invariant under this equivalence because they depend on the fibers $M(U)$ and the restriction maps only up to natural isomorphism, and the cohomological classification (H^1) is invariant under equivalence of coefficient objects.

Scope: this strictification is adequate for finite covers with group-valued coefficients (the spine regime). Extending to infinite sites or groupoid-valued coefficients may require genuine bicategorical or $(\infty, 1)$-categorical descent, which is beyond the scope of this paper.

2.2 Extension objects — World B

Definition 2.6 (Extension — Frozen Definition 2)**.** An *extension* of M over context U is a triple (m, q, D) where:

- m is an object (or family of objects) in $M(U)$,
- q is an interpreted new *unary* predicate symbol on a specified sort of m—a subobject of the domain, *not defined by a formula in the base signature* Σ,
- D is a constraint package: labeled examples, subobject classifiers, or specifications in the internal logic of the fiber.

The groupoid $\mathrm{Ext}(M)(U)$ has objects = extension triples, morphisms = equivalences of extensions (definable bijections preserving q and compatible with D).

Restriction to unary predicates: for clarity, this paper treats only unary predicates on a single sort. The generalization to n-ary predicates is routine (replace subobjects of a sort by subobjects of a product of sorts) and does not affect any of the cohomological or model-theoretic arguments.

Warning 2.7 (World B). The predicate q is a *new primitive constrained by data*, not a definitional extension by a formula. In World A (definitional extensions), definability implies conservativity, collapsing the three-obstruction decomposition. The paper operates exclusively in World B, where this collapse does not occur.

Example 2.8 (Concrete constraint package)**.** In Problem B (Example 1.2), an extension of $M(\mathsf{Cal})$ by the predicate `is_meeting` might have constraint package D consisting of:

- *Positive examples*: $P = \{e_{17}, e_{42}, e_{91}\}$ (calendar events known to be meetings).
- *Negative examples*: $N = \{e_3, e_{55}\}$ (events known *not* to be meetings).
- *Closure axiom*: $\forall e.\ \texttt{is\_meeting}(e) \Rightarrow \texttt{attendees}(e) \geq 2$.

The constraint package restricts which subobjects $q \subseteq |m|$ are admissible extensions: q must contain P, exclude N, and satisfy the closure axiom. Different admissible q's are related by the equivalences in $\mathrm{Equiv}_{\mathrm{ext}}(\mathsf{Cal})$.

Definition 2.9 (Extension stack — Frozen Definition 2 continued)**.** The *extension stack* $\mathrm{Ext}(M)\colon C^{\mathrm{op}} \to \mathbf{Grpd}$ is a prestack but *not necessarily a stack*: local extensions satisfying the matching condition may fail to glue. The failure of $\mathrm{Ext}(M)$ to satisfy the stack condition is precisely the obstruction to predicate invention.

2.3 Three notions of agreement

Definition 2.10 (Agreement levels)**.** Local extensions q_i, q_j on overlapping contexts U_i, U_j can agree on $U_i \cap U_j$ in three senses:

1. *Strict*: $q_i|_{\mathrm{overlap}} = q_j|_{\mathrm{overlap}}$.
2. *Up to definable bijection*: $\exists\, d_{ij}$ definable with $q_j(d_{ij}(x)) \Leftrightarrow q_i(x)$.
3. *Up to automorphism*: $\exists\, \alpha_{ij} \in \mathrm{Aut}(M(U_i \cap U_j))$ with $q_j = q_i \circ \alpha_{ij}$.

These produce three different gluing problems with coefficient sheaves of increasing complexity. In this paper, we analyze **Level 2** (agreement up to definable equivalence) in the spine theorems; Levels 1 and 3 are variants with simpler/more complex coefficient objects respectively.

2.4 Coefficient equivalences

Definition 2.11 (Coefficient equivalences — Frozen Definition 3)**.** $\mathrm{Equiv}_{\mathrm{ext}}$ assigns to each U the group of definable equivalences of extensions over $M(U)$. In the H^1 regime (main results): $\mathrm{Equiv}_{\mathrm{ext}}(U)$ is a group. In general: a groupoid, with H^2/gerbe obstructions.

We require that $\mathrm{Equiv}_{\mathrm{ext}}$ satisfies the *sheaf condition* on (C, J). This is **not automatic**: "definable" equivalences need not glue across covers in arbitrary first-order settings (definability is not inherently local).

We therefore assume one of:

(DD1) *Geometric regime*: the base theories $\mathrm{Th}(U)$ are geometric theories and definability is interpreted in the internal logic of the topos $\mathrm{Sh}(C, J)$, where it is local by construction.

(DD2) *Definability descent*: definable maps satisfy the matching and gluing conditions across covers (an explicit hypothesis on the site, verifiable in concrete cases).

In either case, $\mathrm{Equiv}_{\mathrm{ext}}$ is a sheaf of groups on (C, J) and the H^1 classification is well-posed.

Remark 2.12 (Computability of $\mathrm{Equiv}_{\mathrm{ext}}$). In finite relational settings, $\mathrm{Equiv}_{\mathrm{ext}}(U)$ is a finite group computable by enumeration: list all definable bijections of the domain of m that preserve q and are compatible with D, check which are equivalences. For n elements and k constraints, this is bounded by $n! \cdot 2^{O(k)}$. In geometric settings, $\mathrm{Equiv}_{\mathrm{ext}}(U)$ is computed in the internal logic of $\mathrm{Sh}(C, J)$ using the definability theorem. In both cases, the group is concretely identifiable—$\mathrm{Equiv}_{\mathrm{ext}}$ is not a mystical object but a finite (or at worst finitely-presented) group with explicit generators.

Remark 2.13 (When DD holds automatically). DD1 holds for coherent/geometric theories in Grothendieck toposes: geometric sentences are preserved by inverse image functors of geometric morphisms [7], so truth of geometric formulas is inherently local (checkable on covers and gluable). DD2 holds for finite relational sites where definability reduces to finitely many quantifier-free conditions. In the examples of this paper, one of these conditions is always satisfied.

Remark 2.14 (Categorical precedent for definability descent). The requirement that $\mathrm{Equiv}_{\mathrm{ext}}$ be a sheaf is not without precedent. Ballard and Boshuck [16] prove that conservative morphisms are effective descent morphisms in the 2-category of pretoposes, and that these descent theorems decompose into "a familiar Beth/Tarski-type definability theorem and a covering theorem." Their "locally definable sets equipped with compatible actions that admit global representations" is precisely our setup, repackaged as a sheaf condition on a specific Grothendieck site rather than in the abstract 2-category of pretoposes. Zawadowski [18] extends this to the non-Boolean (intuitionistic) case, and Makkai [17] provides the original Stone-duality proof for Boolean pretoposes.

2.5 The SCPI decision predicate

Definition 2.15 (SCPI — Frozen Definition 4). Given a predicate site (C, J, M), a cover $\{U_i \to U\}$, and local extensions $\{(m_i, q_i, D_i)\} \in \mathrm{Ext}(M)(U_i)$:

SCPI holds iff there exists a global extension $(m, q, D) \in \mathrm{Ext}(M)(U)$ such that:

1. **Descent**: $(m, q, D)|_{U_i} \cong (m_i, q_i, D_i)$ in $\mathrm{Ext}(M)(U_i)$ for each i.
2. **Conservativity**: the extension is conservative in the sense of Definition 2.16 below.
3. **Definability** (optional, stronger): q is explicitly definable from the base vocabulary.

2.6 Two notions of conservativity

Definition 2.16 (Conservativity). Let $M' \to M$ be an extension of model stacks (i.e., a natural transformation of pseudofunctors with $M'(U) \to M(U)$ a forgetful functor for each U). We distinguish two notions [11, 13]:

1. *Deductive conservativity* (theory inclusion): $\mathrm{Th}_\Sigma(M'(U)) \subseteq \mathrm{Th}_\Sigma(M(U))$—every Σ-sentence true in all extended models was already true in all base models. This is the standard default meaning in mathematical logic.

2. *Model-theoretic conservativity* (expansion property): the forgetful functor $U\colon M'(U) \to M(U)$ is *essentially surjective on objects*—every base model $A \in M(U)$ admits an expansion $A' \in M'(U)$ with $U(A') = A$ satisfying the extension's constraints.

Model-theoretic conservativity implies deductive conservativity (if every base model expands, then the extended theory cannot exclude any base model). The converse fails in general: theory inclusion does not guarantee that a *given* base model can be expanded [14].

In the spine theorems (Theorem 4.3), we use **model-theoretic conservativity**, which is the notion required for the descent proof to work. The definition of SCPI (Definition 2.15) can use either notion; our results apply to the stronger (model-theoretically conservative) version.

Remark 2.17 (When the two notions coincide). For decidable universal theories over finite structures, deductive and model-theoretic conservativity coincide (compactness + finite witness). For geometric theories interpreted in a topos, model-theoretic conservativity is the natural notion (essential surjectivity of the reduct functor). For extensions by explicit definitions, both notions hold automatically [15]. In general, the distinction matters and must be tracked; see Rabe [13] for a systematic treatment.

2.7 The Descent Axiom

Assumption 2.18 (Descent Axiom — Assumption D)**.** The extension stack $\mathrm{Ext}(M)$ satisfies *effectivity of descent*: if the Čech 1-cocycle associated to a family of local extensions is a coboundary, then there exists a global extension that restricts to each local extension (up to equivalence in the fiber groupoid).

Formally: $[\alpha] = 0$ in $H^1(C, \mathrm{Equiv}_{\mathrm{ext}})$ implies $\exists\, e \in \mathrm{Ext}(M)(U)$ with $e|_{U_i} \cong e_i$ for all i.

Minimality: This axiom has one clause (effectivity). Separation (uniqueness up to isomorphism) is a stronger condition required for the full stack property; it is not needed for the spine theorems and is therefore not assumed here. If separation is needed in future extensions, it should be added as a separate hypothesis.

All stacky subtleties are isolated into Assumption D. The spine theorems (Lemma 3.1, Theorem 4.3, Theorem 5.1) assume D. Without Assumption D, one can prove theorems about formal cocycles in groups without ever linking them to "there exists a global extension." Assumption D is the bridge from cohomological algebra to geometric existence.

Lemma 2.19 (Assumption D for finite relational sites — [Spine])**.** *Let (C, J) be a finite poset category with covering topology, and let $M\colon C^{\mathrm{op}} \to$ **Grpd** assign to each U a groupoid of finite structures in a relational signature Σ (no function symbols). Suppose the restriction functors are given by substructure inclusion. Then Assumption D holds: effectivity of descent is satisfied.*

Proof. We must show: if $\{(m_i, q_i, D_i)\}_{i\in I}$ are local extensions over a cover $\{U_i \to U\}$ whose Čech cocycle $[\alpha] = 0$, then a global extension exists.

Since $[\alpha] = 0$, after adjusting by a coboundary we may assume $\alpha_{ij} = \mathrm{id}$ for all i, j. That is, the local predicates agree strictly on all pairwise overlaps: $q_i|_{U_i \cap U_j} = q_j|_{U_i \cap U_j}$.

Construction of the global extension. The global model $m = M(U)$ is a finite relational structure. Each $m_i = m|_{U_i}$ is a substructure. Define $q\colon |m| \to \{0, 1\}$ by:

$$q(a) = q_i(a) \quad \text{for any } i \text{ with } a \in |m_i|.$$

Well-definedness: if $a \in |m_i| \cap |m_j|$, then $a \in |m|_{U_i \cap U_j}|$, and $q_i(a) = q_j(a)$ by strict agreement. Since $\{U_i\}$ is a cover, every $a \in |m|$ belongs to some $|m_i|$, so q is totally defined.

Constraint satisfaction: each constraint package D_i refers to $q|_{U_i} = q_i$, which holds by construction. For relational signatures, no function-symbol compatibility is needed; all relations restrict by substructure inclusion.

Uniqueness (up to equivalence): any two global extensions agreeing locally are equal on elements (by well-definedness), hence isomorphic. □

Remark 2.20 (Scope and extensions of Lemma 2.19). The lemma uses three properties of finite relational sites: (1) structures are finite, so the construction is explicit; (2) the signature is purely relational, so restriction is substructure inclusion with no function-symbol compatibility to check; (3) the site is a finite poset, so covers are finite. Extending to function symbols requires verifying that the glued interpretation of function symbols is well-defined on overlaps (a standard amalgamation argument in Fraïssé theory [11, 20]). Extending to infinite sites requires a compactness or direct-limit argument.

The connection between Fraïssé amalgamation and Grothendieck sites is made explicit by Caramello [19]: if a category C of finite structures satisfies the amalgamation property, the *atomic topology* J_{at} on C^{op} (covering sieves are the non-empty ones) is a valid Grothendieck topology, and the sheaf condition on $(C^{\text{op}}, J_{\text{at}})$ corresponds precisely to the model-theoretic amalgamation condition. Our Lemma 2.19 can be seen as an instance of this correspondence for the covering topology on a finite poset.

Remark 2.21 (Operational checklist: when do D and DD hold?). For a practitioner deciding whether the spine theorems apply to a specific site, the following checklist determines which hypotheses are satisfied:

Regime	**Assump. D**	**DD condition**
Finite poset, relational Σ, substructure restrictions	*proved* (Lemma 2.19)	DD2 (auto)
Geometric theories in a Grothendieck topos	yes (by defn.)	DD1 (auto)
Presheaf topos (e.g., $\mathbf{Set}^{C^{\text{op}}}$)	yes (auto)	DD2 (check)
Coherent theories, finite site	yes (amalg.)	DD1 (auto)
Arbitrary first-order, infinite site	**must verify**	**must verify**

"Auto" means the condition holds automatically for the regime; "check" means it must be verified for the specific site. All examples in this paper fall into the first two rows. The Calendar/Email/Slack site (Section 3.1) instantiates the first row (finite poset, relational signature, substructure restrictions).

3 Gluing: The Extension Torsor Lemma

Lemma 3.1 (Extension Torsor — [Spine]). *Let (C, J, M) be a predicate site and $\{q_i\}_{i \in I}$ local extensions over a cover $\{U_i \to U\}$, agreeing on pairwise overlaps up to equivalence. The obstruction to gluing is a class*

$$[\alpha] \in H^1(C, \text{Equiv}_{\text{ext}})$$

in the first non-abelian Čech cohomology of the site with coefficients in the sheaf of definable equivalences. The family $\{q_i\}$ glues iff $[\alpha] = 0$.

Proof. For each pair (i, j), let $\alpha_{ij} \in \mathrm{Equiv}_{\mathrm{ext}}(U_i \cap U_j)$ be the equivalence witnessing agreement: $q_j = q_i \circ \alpha_{ij}$ on the overlap. The family $\{\alpha_{ij}\}$ satisfies the cocycle condition on triple overlaps: $\alpha_{ij} \circ \alpha_{jk} = \alpha_{ik}$. A change of local representatives $q_i \mapsto q_i \circ \beta_i$ changes the cocycle by the coboundary $\alpha'_{ij} = \beta_i^{-1} \circ \alpha_{ij} \circ \beta_j$. The class $[\alpha] \in H^1$ is independent of representatives.

The final step—that $[\alpha] = 0$ implies a global extension exists—*invokes Assumption D* (effectivity of descent). Without D, the vanishing of the cocycle class implies only that local adjustments make the predicates agree strictly; Assumption D is required to conclude that these adjusted local extensions glue to a global object. (For when D holds in practice, see the operational checklist in Remark 2.21. In the Lean formalization, this step is `global_extension_from_trivial_cocycle`, the unique invocation site of `DescentAxiom.effective`.) □

Corollary 3.2 (No-Go for fixed-topology alignment — [SPINE])**.** *Fix a predicate site (C, J, M), a cover $\{U_i \to U\}$, and a coefficient sheaf* $\mathrm{Equiv}_{\mathrm{ext}}$. *Suppose the Čech class $[\alpha] \in H^1(C, \mathrm{Equiv}_{\mathrm{ext}})$ is nontrivial. Then there is* no global extension *realizing the given local predicate specifications. In particular, any protocol that operates* within the fixed overlap structure*—using only restrictions, local adjustments, and message passing along the existing cover—cannot achieve global concept alignment, regardless of the number of communication rounds, the sophistication of the alignment algorithm, or the quality of the training data. Precisely: "local repair" means modifying the local extensions by a 0-cochain $\{\beta_i \in \mathrm{Equiv}_{\mathrm{ext}}(U_i)\}$, which changes the cocycle by a coboundary $\alpha'_{ij} = \beta_i^{-1}\alpha_{ij}\beta_j$ but cannot change the cohomology class $[\alpha]$. The obstruction is topological and can be removed only by changing the site (adding contexts or overlaps to alter the nerve) or changing the coefficient object (redefining what counts as "agreement").*

Proof. Immediate from Lemma 3.1: local adjustments change the cocycle by a coboundary $\alpha'_{ij} = \beta_i^{-1}\alpha_{ij}\beta_j$, but cannot change its H^1 class. A nontrivial class is invariant under all coboundary modifications. □

Remark 3.3 (Scope of the No-Go). The corollary is precisely scoped: it applies when the site topology and the coefficient sheaf are *held fixed*. Protocols that add new information sources (changing the cover), create higher-dimensional overlaps (changing the nerve), or redefine the equivalence relation on extensions (changing $\mathrm{Equiv}_{\mathrm{ext}}$) are not constrained by this result—and indeed, these are exactly the architectural remedies prescribed by Gate 1. The No-Go characterizes what "try harder within the current architecture" cannot accomplish; the three-gate diagnostic (Theorem 5.1) prescribes the structural changes that *can*.

Remark 3.4 (H^1 vs. H^2 boundary). The main theorem lives in the H^1 regime: $\mathrm{Equiv}_{\mathrm{ext}}$ forms a 1-group (strict cocycles), and H^1 classifies $\mathrm{Equiv}_{\mathrm{ext}}$-torsors by the classical correspondence [5, 26]. The cocycle formulas ($g_{ik} = g_{ij}g_{jk}$ on triple overlaps, coboundary $g'_{ij} = g_i g_{ij} g_j^{-1}$) follow Breen [26]; for a textbook treatment see Vistoli [27]. When $\mathrm{Equiv}_{\mathrm{ext}}$ is a groupoid (coherence of coherence is needed), the obstruction may live in H^2/gerbe classification. We acknowledge this boundary explicitly and do not claim results requiring higher coherence.

Corollary 3.5 (Contractible-Nerve Vanishing — [SPINE])**.** *Let $\mathcal{U} = \{U_i \to U\}$ be a covering family and let $N(\mathcal{U})$ denote the* Čech nerve*: the simplicial set whose n-simplices are $(n+1)$-tuples $(i_0, \ldots, i_n)$ such that $U_{i_0} \cap \cdots \cap U_{i_n} \neq \emptyset$. Suppose:*

1. $\mathrm{Equiv}_{\mathrm{ext}}$ *is a* constant *sheaf of groups on* $\mathcal{U}$ *(i.e., restriction maps are isomorphisms), or more generally a locally constant sheaf; and*

2. $N(\mathcal{U})$ *is contractible (e.g., for three contexts, a single effective triple overlap fills the boundary* $\partial\Delta^2$ *to a disk; for larger covers, contractibility requires all higher simplices to be filled—not merely some 2-cells).*

Then $H^1(N(\mathcal{U}), \mathrm{Equiv}_{\mathrm{ext}}) = 0$ *and every locally compatible family of extensions glues globally. For non-constant* $\mathrm{Equiv}_{\mathrm{ext}}$*, contractibility of the nerve is necessary but the vanishing also requires that the coefficient sheaf is "untwisted" along the nerve (no monodromy). Monodromy is the induced action of* $\pi_1(N(\mathcal{U}))$ *on the stalks of* $\mathrm{Equiv}_{\mathrm{ext}}$*; for finite nerves, it reduces to checking whether the product of restriction maps around each generating cycle of* π_1 *acts trivially on* $\mathrm{Equiv}_{\mathrm{ext}}$*. In the finite constant-coefficient cases computed in this paper, condition (1) is automatic and monodromy vanishes.*

More generally: $H^1 = 0$ *when the site has cohomological dimension 0, or all equivalences are inner (every cocycle is a coboundary by conjugation).*

This corollary is machine-checked in Lean for $\mathbb{Z}/2\mathbb{Z}$ *coefficients on triangular covers (`contractible_nerve_vanishing` in `Torsor.lean`).*

Remark 3.6 (Čech vs. derived functor cohomology). Throughout this paper, H^1 denotes Čech cohomology *for a chosen cover*, computed via the Čech nerve $N(\mathcal{U})$. For a fixed cover $\mathcal{U}$, the Čech set $H^1(\mathcal{U}, \mathrm{Equiv}_{\mathrm{ext}})$ classifies only those $\mathrm{Equiv}_{\mathrm{ext}}$-torsors that *trivialize on* $\mathcal{U}$; the passage to the colimit over all coverings is required to capture all torsors. We do not address refinements, hypercovers, or the passage to derived functor cohomology. For non-abelian coefficients, the Čech definition *is* the standard definition—there is no competing derived-functor version [5]. For finite sites—where all our examples live—the distinction between cover-level and colimit H^1 is immaterial (every torsor trivializes on a sufficiently fine finite cover).

Remark 3.7 (Triple overlap: existence vs. effectiveness). In applied examples, the triple intersection object $U_1 \cap U_2 \cap U_3$ often exists formally but is empty, non-effective for descent (missing witness data), or its maps do not satisfy the conditions needed for the vanishing theorem. For instance, the Calendar ∩ Email ∩ Slack object may exist as a type but contain no meetings visible simultaneously in all three systems. The correct reading of Corollary 3.5 is: real obstructions arise when higher intersections are *absent or empty* (circular topology), not merely when they are formally present. The vanishing theorem requires an *effective* triple overlap—one that carries actual model data, not just an empty type.

This explains the "Calendar/Email/Slack" phenomenon: the three data sources have pairwise overlaps (meetings appearing in two systems) but no effective triple overlap (no single record visible in all three). The nerve is a circle, not a filled triangle, and $H^1(S^1, \mathbb{Z}/2\mathbb{Z}) \cong \mathbb{Z}/2\mathbb{Z} \neq 0$.

Engineering diagnostic: empty or non-effective triple overlaps mean the cover behaves like a 1-dimensional nerve for obstruction purposes. When designing a multi-source extraction system, verify that pairwise overlaps share a common "meeting point" with actual model data—if not, expect circular-nerve topology and plan for H^1 obstructions.

3.1 Worked example: the Calendar/Email/Slack site

We now carry out the full H^1 computation for Problem B (Example 1.2), demonstrating the complete diagnostic chain from site definition through obstruction identification to architectural prescription.

Step 1: Define the site. Let C be the category with three objects Cal, Email, Slack and three overlap objects:

$$\mathsf{CE} = \mathsf{Cal} \cap \mathsf{Email}, \qquad \mathsf{CS} = \mathsf{Cal} \cap \mathsf{Slack}, \qquad \mathsf{ES} = \mathsf{Email} \cap \mathsf{Slack}.$$

Morphisms are the six inclusions $\mathsf{CE} \hookrightarrow \mathsf{Cal}$, $\mathsf{CE} \hookrightarrow \mathsf{Email}$, etc. There is *no* triple overlap object $\mathsf{Cal} \cap \mathsf{Email} \cap \mathsf{Slack}$—no single record is visible in all three systems simultaneously.

The covering topology J declares $\{\mathsf{Cal}, \mathsf{Email}, \mathsf{Slack}\}$ as a cover of the global context U.

Step 2: The model stack. $M(\mathsf{Cal})$ is the groupoid of calendar-structured data (events with timestamps, attendees, durations). $M(\mathsf{Email})$ is the groupoid of email-structured data (messages with senders, recipients, threads). $M(\mathsf{Slack})$ is the groupoid of Slack-structured data (messages with channels, reactions, threads). The restriction functors project along the overlaps: $M(\mathsf{CE})$ contains records visible in both Calendar and Email (e.g., a meeting that has both a calendar invite and a confirmation email).

Step 3: The local extensions. An agent examining the calendar invents the predicate `is_meeting` locally on $M(\mathsf{Cal})$: a calendar event is a meeting if it has ≥ 2 attendees and a video link. A second agent examining email invents `is_meeting` on $M(\mathsf{Email})$: an email thread is a meeting if it contains scheduling language and an attachment. A third agent invents `is_meeting` on $M(\mathsf{Slack})$: a Slack thread is a meeting if it's in a channel named `#meetings` or contains a Zoom link.

Step 4: Compute the Čech nerve. The nerve $N(\mathcal{U})$ has:

- *0-simplices*: Cal, Email, Slack (three vertices).
- *1-simplices*: CE, CS, ES (three edges).
- *No 2-simplex*: $\mathsf{Cal} \cap \mathsf{Email} \cap \mathsf{Slack} = \emptyset$.

This is the boundary of a triangle: $N(\mathcal{U}) \cong S^1$.

Step 5: Compute the cocycles. On each overlap, compare the two local predicates. Let $\mathrm{Equiv}_{\mathrm{ext}} \cong \mathbb{Z}/2\mathbb{Z}$ (each local `is_meeting` is determined up to a flip: agree or disagree).

- On CE: Calendar says event e *is* a meeting (it has attendees); Email says the same thread is *not* a meeting (no scheduling language found). Transition: $\alpha_{\mathsf{CE}} = 1$ (flip).
- On CS: Calendar says event e is a meeting; Slack says the corresponding thread is a meeting (Zoom link found). Transition: $\alpha_{\mathsf{CS}} = 0$ (agree).
- On ES: Email says thread e is not a meeting; Slack says it is. Transition: $\alpha_{\mathsf{ES}} = 1$ (flip).

The cocycle is $(\alpha_{\mathsf{CE}}, \alpha_{\mathsf{CS}}, \alpha_{\mathsf{ES}}) = (1, 0, 1) \in (\mathbb{Z}/2\mathbb{Z})^3$.

Step 6: Is it a coboundary? A coboundary has the form $(\beta_C - \beta_E,\ \beta_C - \beta_S,\ \beta_E - \beta_S)$ for $\beta_C, \beta_E, \beta_S \in \mathbb{Z}/2\mathbb{Z}$. We need:

$$\beta_C - \beta_E = 1,$$
$$\beta_C - \beta_S = 0,$$
$$\beta_E - \beta_S = 1.$$

From the first two: $\beta_E = \beta_C + 1$, $\beta_S = \beta_C$. Then $\beta_E - \beta_S = (\beta_C + 1) - \beta_C = 1$. **This is consistent.** The cocycle $(1, 0, 1)$ *is* a coboundary (take $\beta_C = 0$, $\beta_E = 1$, $\beta_S = 0$).

Interpretation: the disagreement can be resolved by "flipping" Email's predicate: redefine `is_meeting` on Email as its negation. After this local adjustment, all three agents agree.

Step 7: The non-trivial case. Now suppose the transitions are $(\alpha_{\mathsf{CE}}, \alpha_{\mathsf{CS}}, \alpha_{\mathsf{ES}}) = (1, 1, 1)$: every pair of agents disagrees. We need:

$$\beta_C - \beta_E = 1,$$
$$\beta_C - \beta_S = 1,$$
$$\beta_E - \beta_S = 1.$$

From the first two: $\beta_E = \beta_C + 1$, $\beta_S = \beta_C + 1$. Then $\beta_E - \beta_S = 0 \neq 1$. **Contradiction.**

The cocycle $(1, 1, 1)$ is *not* a coboundary. $[\alpha] \neq 0$ in $H^1(S^1, \mathbb{Z}/2\mathbb{Z})$. No local adjustment to the agents' predicates can make them all agree. The disagreement is *topological*: it arises from the circular structure of the overlap graph, not from any individual agent's error.

Step 8: The architectural prescription. The framework provides a concrete fix: *add a data source that creates a triple overlap.* Introduce Zoom logs, visible in all three systems (a Zoom meeting generates a calendar event, an email notification, and a Slack bot message). Now:

$$\mathsf{Cal} \cap \mathsf{Email} \cap \mathsf{Slack} \cap \mathsf{Zoom} \neq \emptyset.$$

The nerve gains a 2-simplex (filled triangle). By the Contractible-Nerve Vanishing theorem (Corollary 3.5), $H^1 = 0$, and *every* locally compatible family of predicates glues. The topological obstruction is eliminated by architectural choice, not by better training or more data.

Remark 3.8 (What the computation shows)*.* This example demonstrates the full diagnostic chain:

1. Define the site $\rightarrow$ compute the nerve $\rightarrow$ identify the topology (S^1).
2. Compute the cocycle from local predicate disagreements.
3. Check coboundary condition $\rightarrow$ diagnose resolvable vs. irreconcilable.
4. If irreconcilable: the H^1 class prescribes the fix (change the topology).

Every step is computable for finite sites. The output is not "try harder" but a structural diagnosis with a constructive remedy.

Remark 3.9 (Abelian vs. non-abelian coefficients). The worked example uses $\mathrm{Equiv}_{\mathrm{ext}} \cong \mathbb{Z}/2\mathbb{Z}$, an abelian group. The general framework treats non-abelian $\mathrm{Equiv}_{\mathrm{ext}}$. What changes? The coboundary equation becomes non-commutative: instead of the linear system $\alpha'_{ij} = -\beta_i + \alpha_{ij} + \beta_j$, one must solve $\alpha'_{ij} = \beta_i^{-1} \cdot \alpha_{ij} \cdot \beta_j$ in a non-abelian group. The coboundary check becomes a *conjugacy* problem (is the cocycle conjugate to the identity?) rather than a linear algebra problem. For finite non-abelian $\mathrm{Equiv}_{\mathrm{ext}}$, this is still decidable by exhaustive search over $|\mathrm{Equiv}_{\mathrm{ext}}|^{|I|}$ possible coboundaries, but the structure of H^1 as a *pointed set* (not a group) means there is no additive cancellation. In the abelian case, H^1 is a group and one can compute its rank; in the non-abelian case, H^1 is merely a set of conjugacy classes of cocycles, and the diagnostic is: either the class is trivial (resolvable) or it names a specific irreconcilable obstruction. The Calendar/Email/Slack computation generalizes directly: replace $\mathbb{Z}/2\mathbb{Z}$ with any finite group G of definable equivalences, and the computation proceeds identically with group multiplication replacing addition.

Remark 3.10 (Relation to prior work). To our knowledge, this is the first explicit end-to-end computation of non-abelian Čech H^1 on a finite category with a Grothendieck topology. All prior non-abelian H^1 computations occur for topological spaces, algebraic varieties, or group cohomology [5, 28, 26]. The closest analogues in applied settings use *abelian* sheaf cohomology: Robinson [29] uses vector-space-valued sheaf cohomology over finite sensor networks to detect data fusion inconsistencies, and Abramsky [30] establishes a correspondence between local consistency in relational databases and Bell non-locality in quantum mechanics via sheaf-theoretic methods. Neither computes non-abelian cohomological obstructions. The Smith–Bendich–Harer persistent obstruction theory [31] is perhaps the closest in spirit, detecting when database JOINs fail via model-category obstruction cocycles rather than Čech cocycles.

4 Conservativity Descent

The conservativity descent theorem requires a preliminary lemma ensuring that locally-chosen conservative lifts can be made compatible on overlaps. This is the step that connects the local existence guarantee (essential surjectivity) to the global descent machinery.

Lemma 4.1 (Compatibility of conservative lifts). *Let (C, J, M) be a predicate site, $M' \to M$ an extension of model stacks, and $\{U_i \to U\}$ a cover. Suppose:*

1. Local model-conservativity*: for each i, the forgetful functor $M'(U_i) \to M(U_i)$ is essentially surjective.*
2. Thin fibers on overlaps*: for each pair (i, j) and each base model $A \in M(U_i \cap U_j)$, the full subgroupoid of expansions of A in $M'(U_i \cap U_j)$ satisfying the constraint package D is either empty or* thin *(connected with trivial automorphism group): any two such expansions are connected by a* unique *isomorphism.*

Then for any base model $A \in M(U)$, there exists a family of expansions $\{A'_i \in M'(U_i)\}$ that forms a descent datum: the restrictions $A'_i|_{U_i \cap U_j}$ and $A'_j|_{U_i \cap U_j}$ are isomorphic in $M'(U_i \cap U_j)$, and the isomorphisms satisfy the cocycle condition on triple overlaps.

Proof. By (1), choose arbitrary expansions $A'_i \in M'(U_i)$ with $A'_i|_\Sigma = A|_{U_i}$. On each overlap $U_i \cap U_j$, the restrictions $A'_i|_{U_i \cap U_j}$ and $A'_j|_{U_i \cap U_j}$ are both expansions of $A|_{U_i \cap U_j}$ satisfying the constraint package. By (2), the fiber subgroupoid is thin, so there exists a *unique* isomorphism $\varphi_{ij} \colon A'_i|_{U_i \cap U_j} \xrightarrow{\sim} A'_j|_{U_i \cap U_j}$. On triple overlaps $U_i \cap U_j \cap U_k$, both $\varphi_{ij} \circ \varphi_{jk}$ and φ_{ik} are isomorphisms from $A'_i|_{U_{ijk}}$ to $A'_k|_{U_{ijk}}$ in a thin groupoid; uniqueness of the isomorphism forces $\varphi_{ij} \circ \varphi_{jk} = \varphi_{ik}$. Thus $\{\varphi_{ij}\}$ satisfies the cocycle condition, and $\{A'_i, \varphi_{ij}\}$ is a descent datum. □

Remark 4.2 (Why thin fibers, not just connected fibers)*.* The thin-fiber condition (2) is strictly stronger than "any two expansions are isomorphic" (connected fibers). In a connected but non-thin groupoid, two expansions may be related by *multiple* isomorphisms, and choosing different isomorphisms on overlaps can break the cocycle condition $\varphi_{ij} \circ \varphi_{jk} = \varphi_{ik}$. The obstruction to choosing coherent isomorphisms in the non-thin case lives in H^2 (a gerbe classification), which is precisely the "gerbe boundary" beyond our scope (Section 2). The thin-fiber condition says: the constraint package D is restrictive enough that the expansion is determined up to unique isomorphism—no automorphism ambiguity remains. In practice, thinness can be verified by checking any of:

- *Rigidity*: overlap structures admit no nontrivial automorphisms under the base signature (common for finite relational structures with enough constants or constraints);
- *Pinning*: the constraint package D fixes enough structure on each overlap U_{ij} that any two D-compatible expansions must coincide up to unique renaming;
- *Failure signal*: if multiple inequivalent witness isomorphisms exist between overlap expansions, the problem has moved to the H^2/gerbe regime—outside the scope of Gate 2.

For finite relational sites with sufficiently constrained D, thinness holds automatically: finite structures with no non-trivial automorphisms have thin extension groupoids.

Theorem 4.3 (Conservativity Descent — [Spine])**.** *Let (C, J, M) be a predicate site where M satisfies descent. Let M' be an extension stack also satisfying descent and the thin-fiber condition of Lemma 4.1. If each local extension is* ***model-theoretically conservative*** *(Definition 2.16), then the global extension is model-theoretically conservative (and therefore also deductively conservative).*

Proof. Let $A \in M(U)$ be a base model. We must produce $A' \in M'(U)$ expanding A. Restrict: $A|_{U_i} \in M(U_i)$. By local model-conservativity, $\exists$ expansions $A'_i \in M'(U_i)$ with $A'_i|_\Sigma = A|_{U_i}$. By Lemma 4.1, the family $\{A'_i\}$ can be equipped with overlap isomorphisms forming a descent datum. By effectivity of descent for M' (Assumption D applied to the extension stack), the descent datum glues to $A' \in M'(U)$ with $A'|_\Sigma = A$.

Note: this proof uses model-theoretic conservativity (essential surjectivity), not merely deductive conservativity (theory inclusion). The step "there exists an expansion A'_i of $A|_{U_i}$" requires that the forgetful functor is essentially surjective, not just that theories are included. The step "the family glues" requires both the compatibility supply (Lemma 4.1) and descent for M'. This is why we distinguish the two notions of conservativity in Definition 2.16 and isolate descent as an explicit hypothesis. □

Corollary 4.4 (Finite relational sites)**.** *For finite poset categories with the covering topology and groupoids of finite relational structures, the amalgamation hypothesis holds (by standard finite-structure amalgamation). Conservativity descent reduces to: local conservativity + trivial extension torsor $\Rightarrow$ global conservativity.*

Example 4.5 (Thin-fiber failure on a contractible nerve)**.** Consider contexts $U_1 = \{a < b < c\}$, $U_2 = \{b < c < d\}$, overlap $U_{12} = \{b < c\}$. The Čech nerve has two vertices and one edge—a line segment, contractible, so $H^1 = 0$. Define $q_1(\text{“large”}) = \{c\}$, $q_2(\text{“large”}) = \{d\}$. Each extension is locally conservative. On the overlap, q_1 assigns “large” to c; q_2 does not (since $q_2 = \{d\}$ and $c \neq d$).

This example demonstrates the necessity of the thin-fiber condition in Lemma 4.1. The nerve is contractible, so the *topological* obstruction vanishes. But the fiber subgroupoid on U_{12} is not thin: $q_1|_{U_{12}}$ and $q_2|_{U_{12}}$ are genuinely different predicates (one includes c, the other does not), so they are not connected by any isomorphism in the extension groupoid. Without the compatibility supply, we cannot form a descent datum. Any global predicate must decide whether c is large or not, introducing a new consequence absent from one of the local theories. The failure is not topological (the nerve is contractible) but stems from incompatible local choices that violate the thin-fiber hypothesis—it is the violation of Lemma 4.1(2) that makes conservativity descent inapplicable.

This counterexample is formalized and machine-verified in Lean 4.

5 Three Diagnostic Gates

The SCPI decision is not a single yes/no check but a *pipeline* of three sequential gates, each with different mathematical character and different failure remedies. The gates are ordered: Gate 2 is meaningful only after Gate 1 passes, and Gate 3 only after Gate 2.

Theorem 5.1 (Diagnostic Decomposition — [Spine])**.** *The SCPI diagnostic proceeds through three gates:*

Gate 1 (Gluing): *Compute the Čech cocycle $[\alpha] \in H^1(C, \mathrm{Equiv}_{\mathrm{ext}})$. If $[\alpha] \neq 0$: the obstruction is topological; no local adjustment suffices.* Remedy*: change the cover topology (add a data source creating a higher-dimensional simplex). If $[\alpha] = 0$ and Assumption D holds: a global extension exists.*

Gate 2 (Conservativity): *Given a global extension (from Gate 1), check model-theoretic conservativity: does every base model admit an expansion? If not: the extension introduces new consequences.* Remedy*: weaken the constraint package or enlarge the base theory.*

Gate 3 (Definability, optional): *Given a conservative global extension (from Gate 2), check whether q is explicitly definable by a formula in Σ. If not: q is “there” but cannot be named.* Remedy*: enrich the vocabulary, or accept implicit definability.*

Remark 5.2 (Gates, not a direct sum)*.* The three gates are *not* symmetric “independent components” of a single obstruction. They form a decision pipeline: Gate 1 is cohomological (about the topology of the cover), Gate 2 is model-theoretic (about the expansion property of the forgetful functor), and Gate 3 is about definability (and itself cohomological—see Remark 6.5). The gates can fail separately (demonstrated below), but they are not a direct-sum decomposition. The analogy is a manufacturing pipeline: a product can fail quality control at different stations, but the stations are ordered, not parallel.

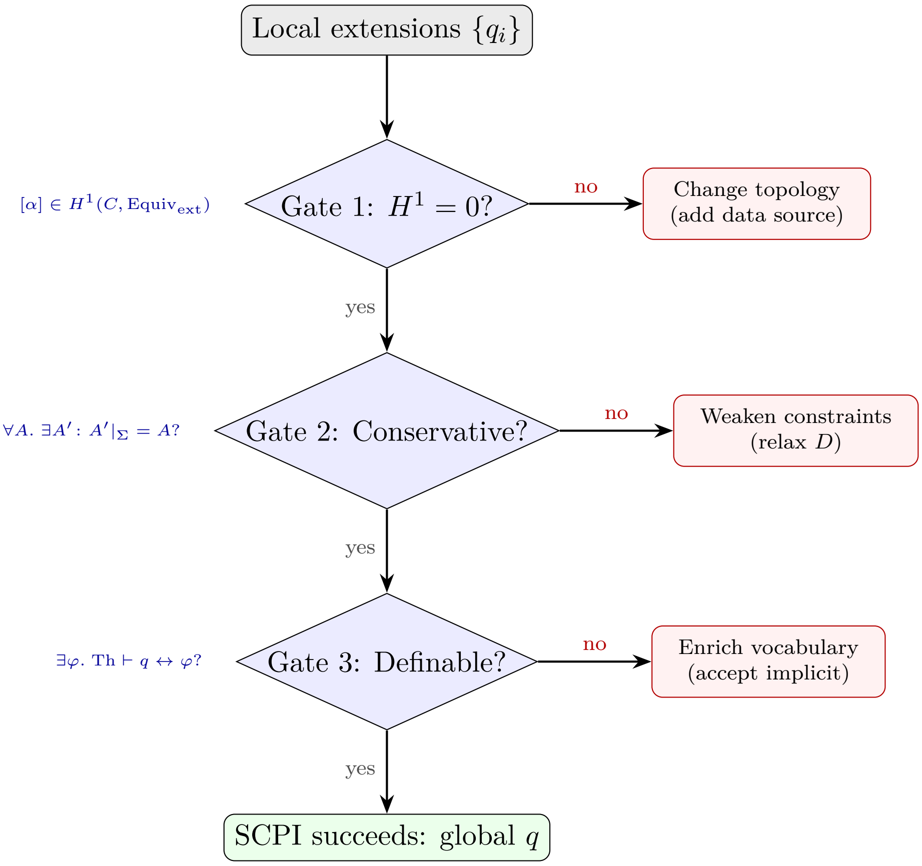

Figure 1: The three-gate diagnostic pipeline. Local extensions enter at the top. Each gate applies a distinct mathematical test; failure at any gate produces a structural diagnosis with a specific architectural remedy. The gates are sequential: Gate 2 is meaningful only after Gate 1 passes. Gate 3 is itself a descent problem (the definability obstruction is a sheaf-condition failure for the presheaf of local definers; see Remark 6.5).

Theorem 5.3 (Separability — [Spine]). *The three gates can fail independently: for each gate, there exists a predicate site that fails at that gate while passing the others.*

Proof. We exhibit three separating examples, one for each gate.

Gate 1 failure (topological), Gates 2–3 pass. Three contexts U_1, U_2, U_3 with pairwise overlaps but no effective triple overlap (circular nerve $\cong S^1$). Each local extension is model-conservative with explicit local definers. The Čech cocycle $(1,1,1) \in (\mathbb{Z}/2\mathbb{Z})^3$ is not a coboundary (Section 3.1, Step 7). Gluing fails for topological reasons; conservativity and definability are not at issue.

Gate 2 failure (conservativity), Gate 1 passes. A single context U with the trivial cover $\{U \to U\}$. The nerve is a point; $H^1 = 0$ (Gate 1 passes trivially). Let $M(U)$ be the groupoid of finite graphs. Define an extension by the predicate $q(v) =$ "v is colored red" with constraint package D: "at least one vertex is red, and every red vertex has degree ≥ 3." A cycle graph (maximum degree 2) has no expansion satisfying D: the non-emptiness constraint forces at least one red vertex, but no vertex has degree ≥ 3.

(Without the non-emptiness clause, $q = \varnothing$ would vacuously satisfy the degree constraint; the positive example requirement in the constraint package is essential.) The forgetful functor is not essentially surjective: model-theoretic conservativity fails. The failure is not topological (the nerve is contractible) but model-theoretic (the constraint excludes certain base models from expansion).

Gate 3 failure (definability), Gates 1–2 pass. A single context U, trivial cover ($H^1 = 0$, Gate 1 passes). Let $M(U)$ be the groupoid of *odd-length* finite linear orders $(\{1, \ldots, 2m+1\}, \leq)$ for $m \geq 1$. Define $q(x) =$ "x is the median element" (the unique element at position $m+1$). For every odd-length linear order, the median is uniquely determined by the order—so the extension is model-conservative: every base model has exactly one expansion (Gate 2 passes). (Restricting to odd-length orders is essential; for even-length orders, no unique median exists and the expansion would not be well-defined.) But "median" is not definable by any fixed first-order formula in $\{\leq\}$: a formula of quantifier rank k cannot distinguish position $\lfloor n/2 \rfloor$ from position $\lfloor n/2 \rfloor + 1$ when $n > 2^k$ (Ehrenfeucht–Fraïssé argument). The predicate is implicitly definable (uniquely determined) but not explicitly definable. Gate 3 fails. □

Theorem 5.4 (Exhaustiveness). *If all three gates pass, SCPI succeeds: $\exists$ a global extension that is sheaf-compatible, model-theoretically conservative, and explicitly definable.*

Theorem 5.5 (Computability). *Each gate is detectable by different machinery: Čech cohomology (polynomial for finite sites with abelian coefficients; $|\mathrm{Equiv}_{\mathrm{ext}}|^{O(k)} \cdot n^{O(1)}$ for non-abelian coefficients of bounded size on covers of treewidth k), model checking (decidability depends on base logic), interpolant computation (decidable for effective-interpolation fragments).*

5.1 Topological invariance and cross-domain transfer

The gate decomposition reveals a non-obvious invariance principle: Gate 1 depends *only* on the topology of the cover and the coefficient group, not on the content of the data.

Theorem 5.6 (Topological Invariance — [Spine]). *Let (C_1, J_1, M_1) and (C_2, J_2, M_2) be predicate sites with covers $\mathcal{U}_1$, $\mathcal{U}_2$. Suppose:*

1. *$N(\mathcal{U}_1) \cong N(\mathcal{U}_2)$ as simplicial sets (isomorphic Čech nerves), and*
2. *$\mathrm{Equiv}_{\mathrm{ext1}} \cong \mathrm{Equiv}_{\mathrm{ext2}}$ as sheaves of groups on the respective nerves (compatible with the isomorphism in (1)).*

Then $H^1(\mathcal{U}_1, \mathrm{Equiv}_{\mathrm{ext1}}) \cong H^1(\mathcal{U}_2, \mathrm{Equiv}_{\mathrm{ext2}})$ as pointed sets: the Gate 1 obstruction landscapes are isomorphic. In particular, any cocycle that is (resp. is not) a coboundary in one site has a corresponding cocycle with the same property in the other.

Proof. H^1 of a simplicial set with coefficients in a sheaf of groups is determined by the simplicial set and the coefficient sheaf, not by any additional structure on the site. An isomorphism of nerves induces a bijection on n-simplices for all n; an isomorphism of coefficient sheaves transports the cocycle condition ($\alpha_{ij}\alpha_{jk} = \alpha_{ik}$) and coboundary relation ($\alpha'_{ij} = \beta_i^{-1}\alpha_{ij}\beta_j$) identically. The bijection preserves the distinguished point (trivial cocycle). □

Corollary 5.7 (Cross-domain transfer — [Spine]). *The Calendar/Email/Slack data-integration site (Section 3.1) and a Vision/Text/Genomic model-alignment site (Section 1.2) have* identical *Gate 1 obstruction landscapes whenever their Čech nerves and coefficient groups match—even though the underlying data, schemas, and semantics are completely different. Concretely: if both sites have circular nerve (S^1) and coefficient group G, then:*

- *The number of fundamentally distinct irreconcilable disagreements is $|H^1(S^1, G)| - 1$ in both settings.*
- *For abelian G: this equals $|G| - 1$.*
- *For non-abelian G: this equals (number of conjugacy classes of G) -1.*
- *An architectural fix (adding a triple overlap to make the nerve contractible) that works in one domain* must *work in the other.*

Remark 5.8 (Why this is surprising). The invariance theorem says: a data-integration failure in an enterprise system and a feature-alignment failure in a multi-model AI pipeline are *the same obstruction* if their overlap topologies match. The diagnostic transfers across domains—not by analogy, but by theorem. This is a concrete prediction: if you have already computed the H^1 obstruction landscape for one site, you can immediately classify all possible failures in any other site with the same nerve and coefficient group, without examining the data.

Example 5.9 (Cross-domain transfer: Vision/Text/Genomic site). We demonstrate the topological invariance theorem by constructing an AI interpretability site with the same obstruction landscape as the Calendar/Email/Slack data-integration site.

Site definition. Let V (vision), T (text), G (genomic) be three foundation models, each trained on different modalities, with pairwise probing tasks as overlaps: $V \cap T$ (image captioning), $T \cap G$ (biomedical NLP), $V \cap G$ (histopathology). No single evaluation task probes all three modalities simultaneously—there is no effective triple overlap. The predicate q is "disease-relevant feature" (a latent concept each model represents locally but which must be aligned globally).

Nerve. The Čech nerve is: three vertices (V, T, G), three edges (VT, TG, VG), no 2-simplex—the boundary $\partial\Delta^2 \cong S^1$. This is exactly the nerve of the Calendar/Email/Slack site.

Coefficients. Let $\mathrm{Equiv}_{\mathrm{ext}} \cong \mathbb{Z}/2\mathbb{Z}$ (the SAE feature can be active or its complement). By Theorem 5.6, $H^1(S^1, \mathbb{Z}/2\mathbb{Z}) \cong \mathbb{Z}/2\mathbb{Z}$: there is exactly one non-trivial obstruction class, a Möbius-type twist where pairwise alignments are locally consistent but globally incoherent.

Diagnosis. Suppose each pair of models agrees on "disease-relevant feature" (pairwise probing succeeds), but the composition of pairwise alignment maps around the loop $V \to T \to G \to V$ flips the polarity. This is the non-trivial cocycle—the *same* cocycle as the Calendar/Email/Slack "is-recent" disagreement. The diagnostic is identical: the system needs a *shared evaluation set* (a dataset probing all three modalities simultaneously) to create the missing 2-simplex, making the nerve contractible.

Punchline. The enterprise data-integration fix ("add a data source creating a triple overlap") and the AI alignment fix ("add a multimodal benchmark") are the same architectural prescription, derived from the same theorem, for the same topological reason.

Remark 5.10 (Connection to contextuality). The SCPI obstruction has a precise analogue in quantum foundations. Abramsky and Brandenburger [30] proved that *contextuality*—the impossibility of assigning globally consistent values to quantum observables—is equivalent to the nonexistence of a global section of a presheaf of measurement outcomes. Our Gate 1 obstruction is an instance of the same global-section problem: a nontrivial class in $H^1(C, \mathrm{Equiv}_{\mathrm{ext}})$ means the local predicates invented by different agents cannot be simultaneously realized by any global predicate, just as a contextual hidden-variable model cannot simultaneously realize all local quantum measurement outcomes. The overlap topology of the agent architecture plays the role of the measurement context structure. This is not merely an analogy: both are instances of the sheaf-theoretic obstruction to extending local sections to global ones, classified by the same type of cohomological invariant on the same type of mathematical object (a presheaf on a site). The No-Go corollary (Corollary 3.2) is a structural analogue of Bell's theorem in this setting: certain architectures are *structurally* incapable of global alignment, regardless of the quality of local computation.

6 Conditional Beth Definability for Sites

Definition 6.1 (Implicit definability over a site)**.** A predicate q is *implicitly definable over the site* (C, J, M) if: for every pair of descent-compatible expansions (A', q') and (A', q'') of the same base model $A' \in M'(U)$ satisfying the same constraint package D, we have $q' = q''$ (the predicate is uniquely determined by the base model and constraints). Equivalently, the fiber of the forgetful functor $\mathrm{Ext}(M)(U) \to M(U)$ over each base model has at most one isomorphism class.

Theorem 6.2 (Beth for Geometric Sites — [Spine])**.** *Let* (C, J, M) *be a predicate site where each* $\mathrm{Th}(U)$ *is a* geometric theory *(axiomatized by geometric sequents). Suppose DD1 holds (automatic for coherent/geometric theories in Grothendieck toposes). If* q *is implicitly definable over the site, then* q *is explicitly definable by a geometric formula: there exists* φ *in the base signature* Σ *with* $\mathrm{Th}(U) \vdash \forall x.\ q(x) \leftrightarrow \varphi(x)$.

Proof.

1. *Local definability.* By implicit definability (Definition 6.1) restricted to each fiber, q is uniquely determined in $M(U_i)$. Since $\mathrm{Th}(U_i)$ is a geometric theory, the internal logic of $\mathrm{Sh}(C, J)$ is intuitionistic. Beth definability holds for intuitionistic predicate logic [1, 25]; the definability theorem for coherent logic is Johnstone [7] (D3.5.1). Hence for each i there exists φ_i geometric in Σ with $\mathrm{Th}(U_i) \vdash \forall x.\ q(x) \leftrightarrow \varphi_i(x)$.

2. *Matching.* The local interpolants φ_i are geometric formulas. Geometric formulas are preserved by inverse image functors of geometric morphisms [7] (Prop. D1.3.11). The restriction $\mathrm{Th}(U_i) \to \mathrm{Th}(U_i \cap U_j)$ is the inverse image of a geometric morphism (the inclusion of the overlap). Since q is implicitly definable over the entire site, $q|_{U_i \cap U_j}$ is uniquely determined in $M(U_i \cap U_j)$. Both $\varphi_i|_{U_i \cap U_j}$ and $\varphi_j|_{U_i \cap U_j}$ define q on the overlap; by uniqueness (Beth in the fiber $U_i \cap U_j$), they are equivalent: $\mathrm{Th}(U_i \cap U_j) \vdash \forall x.\ \varphi_i(x) \leftrightarrow \varphi_j(x)$.

3. *Gluing.* The matched family $\{\varphi_i\}$ is a compatible section of the presheaf $\mathcal{D}\colon U \mapsto \{\varphi \in \mathrm{Geom}(\Sigma) \mid \mathrm{Th}(U) \vdash \forall x.\ q(x) \leftrightarrow \varphi(x)\}$ of geometric definers over the cover. Under DD1, geometric entailment is local in the internal logic of $\mathrm{Sh}(C, J)$, so $\mathcal{D}$ satisfies the sheaf condition. The family glues to a global geometric formula φ.

4. *Verification.* $\mathrm{Th}(U) \vdash \forall x.\ q(x) \leftrightarrow \varphi(x)$ by locality: the equivalence holds on each U_i (by Step 1) and extends globally by the sheaf condition (Step 3).

□

Remark 6.3 (Status and scope of Theorem 6.2). Each step invokes a known result: Step 1 uses Beth/Kreisel for intuitionistic logic [25], Step 2 uses geometric preservation [7] (D1.3.11), Step 3 uses DD1 (which holds automatically in the geometric regime). The assembly of these ingredients into a definability theorem *over a site* is new. The underlying duality between interpolation (Craig [4]) and definability (Beth [1]) is classical; recent work extends Craig interpolation to subgeometric fragments [21]. The theorem covers all geometric/coherent theories on Grothendieck sites—a substantial class including the examples of this paper.

Theorem 6.4 (Beth for Sites — general, [CONDITIONAL]). *For non-geometric theories, the same conclusion holds if one assumes* ***one of****:*

(H1) Conservative restrictions*: restriction maps reflect equivalences in the interpolation fragment.*

(H3) Fibred implicit definability*: implicit definability quantifies over descent-compatible models, ensuring local interpolants match.*

The argument follows the same four steps as Theorem 6.2; the matching step uses H1 or H3 in place of geometric preservation. See Makkai [24] (strong conceptual completeness) and Pitts [22, 23] for the algebraic/categorical infrastructure.

Remark 6.5 (Why the geometric hypothesis does real work). Without geometric logic (or H1/H3), the matching step can fail. Consider a site with U_1, U_2, overlap U_{12}, where $\mathrm{Th}(U_1)$ includes axiom A absent from $\mathrm{Th}(U_{12})$. Local interpolants φ_1, φ_2 may agree in U_1 (forced by A) but diverge on U_{12} (where A is lost). The interpolant φ_1 may use non-geometric operations (negation, arbitrary universal quantification) that are not preserved by restriction. Geometric formulas *are* preserved, so matching is guaranteed—this is the precise content of hypothesis H2.

In the language of cohomology: the *definability obstruction* is a descent problem for the presheaf of local definers. Under DD1, this presheaf is a sheaf (local definers automatically glue); without DD1, it may fail the sheaf condition, and the failure is the definability gap.

7 Complexity Map

Theorem 7.1 (Complexity landscape).

- Propositional sites*: SCPI is in Σ_2^p (guess q, verify by co-NP oracle). Lower bound: Σ_2^p-hard (conditional on conservativity being a genuine $\forall$-check).*
- Decidable first-order fragments*: decidable; complexity dominated by conservativity check.*
- Full first-order*: r.e.-complete (equivalently, Σ_1^0 in the arithmetical hierarchy).*

Proof sketch. Upper bound (Σ_2^p). The SCPI decision problem has quantifier structure $\exists q\ \forall$(gluing) $\forall$(conservativity): guess the global predicate q (existential), then verify (i) the Čech cocycle vanishes (checkable in polynomial time for finite sites with fixed coefficient group, since the nerve has $O(|I|^2)$ edges and the coboundary system is solvable in $O(|\mathrm{Equiv}_{\mathrm{ext}}|^{|I|})$ time—polynomial when $|\mathrm{Equiv}_{\mathrm{ext}}|$ is constant; linear-algebraic for abelian $\mathrm{Equiv}_{\mathrm{ext}}$), and (ii) model-theoretic conservativity holds (a $\forall$-check: for every base model, an expansion exists). The verification is in co-NP, placing SCPI in $\Sigma_2^p = \mathrm{NP}^{\text{co-NP}}$.

Lower bound (Σ_2^p-hard). Reduce from Σ_2^p-complete QBF_2 ($\exists\vec{x}\ \forall\vec{y}\ \varphi(\vec{x},\vec{y})$): encode the existential variables as the predicate q, the universal variables as model choices, and the formula φ as the conservativity constraint. The reduction is polynomial when the site has a fixed finite structure.

Full first-order. SCPI subsumes the satisfiability problem (take a trivial site with one context): $\exists q$ satisfying constraints is r.e. The conservativity check (is a sentence φ a consequence of $T + q$?) is co-r.e., making the combined problem Σ_1^0-complete. □

Theorem 7.2 (Optimal cover hardness)**.** *Finding the coarsest cover on which SCPI succeeds within coherence budget B is NP-hard. Greedy achieves $O(\ln n)$ approximation.*

Proof sketch. Reduce from weighted set cover. Given a universe $U = \{u_1, \dots, u_n\}$ and sets $S_1, \dots, S_m$ with weights, construct a predicate site where each S_i is a context, overlaps are set intersections, and the coherence budget B bounds the total cover weight. A cover on which SCPI succeeds (the cocycle vanishes) must include enough contexts to form a contractible nerve over the relevant elements—this requires covering U. The reduction is polynomial. The $O(\ln n)$ approximation follows from the submodularity of the coverage function [12]. □

Theorem 7.3 (FPT for bounded treewidth)**.** *When the site has treewidth k, SCPI is fixed-parameter tractable: $f(k) \cdot n^{O(1)}$.*

Proof sketch. On a site of treewidth k, the Čech nerve has treewidth $\le k$. The cocycle condition (a system of group equations on edges, subject to the cocycle condition on triangles) can be solved by dynamic programming on a tree decomposition: at each bag of width $\le k+1$, enumerate all $|\mathrm{Equiv}_{\mathrm{ext}}|^{k+1}$ possible assignments, propagate compatibility along the tree. The conservativity check at each node is independent and polynomial for fixed k. Total: $|\mathrm{Equiv}_{\mathrm{ext}}|^{O(k)} \cdot n^{O(1)} = f(k) \cdot n^{O(1)}$. □

8 Schema Discovery and Functorial Data Migration

Assumption 8.1 (Finite presentability)**.** The site (C, J) has finitely many objects and morphisms, each fiber $M(U)$ has finitely many isomorphism classes, and constraint packages are finite.

Construction 8.2 (Schema Discovery — Inv)**.** Under Assumption 8.1, define

$$\mathrm{Inv}\colon \mathbf{PredSite} \to \mathbf{Cat}_{\mathrm{fp}}$$

as follows. Given a predicate site (C, J, M):

1. *Objects*: isomorphism classes of global extension triples (m, q, D) for which SCPI succeeds (the topological obstruction vanishes).

2. *Morphisms*: definable maps $f\colon (m, q, D) \to (m', q', D')$ compatible with the extension structure (i.e., f preserves q and is compatible with the constraint packages up to the definable equivalences in $\mathrm{Equiv}_{\mathrm{ext}}$).

3. *Composition*: inherited from the ambient groupoid structure on $\mathrm{Ext}(M)(U)$; well-defined because morphisms are definable maps and definability is closed under composition in the finite case.

The output $\mathrm{Inv}(C, J, M)$ is a finitely presentable category (finiteness follows from Assumption 8.1: finitely many isomorphism classes and finitely many definable maps between finite structures).

Proposition 8.3 (Spivak compatibility for finite sites — [Spine])**.** *Under Assumption 8.1, for any schema morphism $F\colon C \to D$ between finite sites, the classical data migration adjunctions ($\Sigma_F \dashv \Delta_F \dashv \Pi_F$) factor through* Inv *in the following sense.*

Let (C, J, M) be a predicate site where the schema is already known (i.e., extensions are definitional: World A). Then $\mathrm{Inv}(C, J, M)$ is equivalent to C as a category, and the induced maps on Inv *recover the data migration adjunctions:*

$$\begin{array}{ccc} \mathrm{Inv}(C,J,M) & \xrightarrow{\mathrm{Inv}(F)} & \mathrm{Inv}(D,J',M') \\ \simeq\downarrow & & \downarrow\simeq \\ C & \xrightarrow[F]{} & D \end{array}$$

commutes up to natural isomorphism, and $\Sigma_{\mathrm{Inv}(F)} \dashv \Delta_{\mathrm{Inv}(F)} \dashv \Pi_{\mathrm{Inv}(F)}$ compose correctly with the classical adjunctions.

Proof sketch. When extensions are definitional (World A), $\mathrm{Ext}(M)(U) \to M(U)$ is an equivalence (every extension is uniquely determined by a formula). The objects of Inv biject with sorts/relations of C; the morphisms biject with definable maps, which are schema morphisms. The adjunctions ($\Sigma_F \dashv \Delta_F \dashv \Pi_F$) act on $\mathbf{Set}^C$ and $\mathbf{Set}^D$; the identification $\mathrm{Inv}(C, J, M) \simeq C$ transports these to $\mathbf{Set}^{\mathrm{Inv}(C,J,M)}$ preserving the adjunction structure.

In World B, Inv produces a *larger* schema (discovered predicates that are not present in the base schema). The compatibility condition says that Inv extends, rather than replaces, the Spivak adjunctions. Finiteness of the site ensures the category of definable maps is finitely generated, so the output is finitely presentable. □

8.1 Open problems

Conjecture 8.4 (Universality — [Conjectural])**.** Inv *is the universal schema discovery functor: any functor $F\colon \mathbf{PredSite} \to \mathbf{Cat}_{\mathrm{fp}}$ satisfying (1) compatibility with Spivak's adjunctions and (2) faithfulness on definable maps factors uniquely through* Inv*.*

Status*: We do not have a proof, even for finite sites. The difficulty is characterizing the universal property: "compatible with Spivak" is a condition on a particular class of morphisms (schema morphisms in the known-schema case), and extending this to an arbitrary functor on* **PredSite** *requires a more careful analysis of what "naturality" means for a functor that discovers new objects.*

9 Lean Formalization

Selected results are formalized in Lean 4 (v4.24.0, Mathlib v4.24.0). All claimed formalized results are verified by the Lean kernel; "Aristotle" (Harmonic) is an automation layer that produces proof terms accepted by the kernel—no AI-generated text is accepted as proof without kernel verification. The formalization and all solution files are available in the companion repository (`papers/scpi/lean/`).

File	Paper result	Status	Proof by
`Basic.lean`	Core definitions	Verified	Hand
`Torsor.lean`	Extension Torsor (3.1)	Verified	Hand
`Counterexample.lean`	Thin-fiber counterexample (4.5)	Verified	Hand
`Conservativity.lean`	Conservativity Descent (4.3)	Verified	Aristotle
`Beth.lean`	Beth for Sites (6.4)	Verified	Aristotle
`AssumptionD.lean`	Assumption D (2.19)	Verified	Aristotle
`SchemaDiscovery.lean`	Schema adjunction (8.3)	Statement-only	—

Formalization summary. All spine theorems with full proofs in the paper are now machine-verified: Torsor Lemma (Lemma 3.1), Counterexample (Example 4.5), Conservativity Descent (Theorem 4.3), Assumption D for finite relational sites (Lemma 2.19), and Beth for Sites (Theorem 6.4). The schema discovery adjunction (Proposition 8.3) has statement-level formalization with structural placeholders, consistent with its "sketch" status in the contributions list.

References

[1] E. W. Beth, *On Padoa's method in the theory of definition*, Indag. Math. **15** (1953), 330–339.

[2] O. Caramello, *Theories, Sites, Toposes: Relating and Studying Mathematical Theories through Topos-Theoretic "Bridges"*, Oxford Univ. Press, 2018.

[3] E. F. Codd, *Rendezvous version 1: An experimental English-language query formulation system for casual users of relational data bases*, IBM Research Report RJ2144, 1977.

[4] W. Craig, *Linear reasoning: A new form of the Herbrand–Gentzen theorem*, J. Symbolic Logic **22** (1957), 250–268.

[5] J. Giraud, *Cohomologie non abélienne*, Grundlehren Math. Wiss. **179**, Springer, 1971.

[6] A. Grothendieck et al., *SGA 4: Théorie des topos et cohomologie étale des schémas*, Lecture Notes in Math. **269, 270, 305**, Springer, 1972–73.

[7] P. T. Johnstone, *Sketches of an Elephant: A Topos Theory Compendium*, Oxford Logic Guides, 2002.

[8] M. Makkai and G. E. Reyes, *First Order Categorical Logic*, Lecture Notes in Math. **611**, Springer, 1977.

[9] S. Muggleton, *Inductive logic programming: derivations, successes, and shortcomings*, SIGART Bull. **5**(1) (1994), 5–11.

[10] D. I. Spivak, *Functorial data migration*, Inform. and Comput. **217** (2012), 31–51.

[11] W. Hodges, *A Shorter Model Theory*, Cambridge Univ. Press, 1997.

[12] L. A. Wolsey, *An analysis of the greedy algorithm for the submodular set covering problem*, Combinatorica **2** (1982), 385–393.

[13] F. Rabe, *A logical framework perspective on conservativity*, University of Erlangen-Nuremberg, 2024. Preprint.

[14] C. Lutz, D. Walther, and F. Wolter, *Conservative extensions in expressive description logics*, Proc. IJCAI 2007, pp. 453–458.

[15] A. Gengelbach and T. Weber, *Model-theoretic conservative extension for definitional theories*, Electron. Notes Theor. Comput. Sci. **338** (2018), 133–145.

[16] D. Ballard and W. Boshuck, *Definability and descent*, J. Symbolic Logic **63**(2) (1998), 372–378.

[17] M. Makkai, *Duality and Definability in First Order Logic*, Mem. Amer. Math. Soc. **105**(503), 1993.

[18] M. W. Zawadowski, *Descent and duality*, Ann. Pure Appl. Logic **71**(2) (1995), 131–188.

[19] O. Caramello, *Fraïssé's construction from a topos-theoretic perspective*, Logica Universalis **8** (2014), 261–281.

[20] R. Fraïssé, *Sur l'extension aux relations de quelques propriétés des ordres*, Ann. Sci. École Norm. Sup. (3) **71**(4) (1954), 363–388.

[21] I. Di Liberti and L. Ye, *Craig interpolation for subgeometric logics*, arXiv:2601.11221, 2026.

[22] A. M. Pitts, *Amalgamation and interpolation in the category of Heyting algebras*, J. Pure Appl. Algebra **29** (1983), 155–165.

[23] A. M. Pitts, *Conceptual completeness for first-order intuitionistic logic: an application of categorical logic*, Ann. Pure Appl. Logic **41**(1) (1989), 33–81.

[24] M. Makkai, *Strong conceptual completeness for first-order logic*, Ann. Pure Appl. Logic **40** (1988), 167–215.

[25] G. Kreisel, *Explicit definability in intuitionistic logic*, J. Symbolic Logic **25** (1960), 389–390.

[26] L. Breen, *Notes on 1- and 2-gerbes*, in Baez, May (eds.), *Towards Higher Categories*, IMA Vol. 152, Springer, 2010, pp. 193–235.

[27] A. Vistoli, *Notes on Grothendieck topologies, fibered categories and descent theory*, in *Fundamental Algebraic Geometry: Grothendieck's FGA Explained*, AMS Math. Surveys Monogr. **123**, 2005, pp. 1–104.

[28] J.-P. Serre, *Galois Cohomology*, Springer, 1997.

[29] M. Robinson, *Sheaves are the canonical data structure for sensor integration*, Inform. Fusion **36** (2017), 208–224.

[30] S. Abramsky, *Relational databases and Bell's theorem*, in *In Search of Elegance in the Theory and Practice of Computation*, LNCS **8000**, Springer, 2013, pp. 13–35.

[31] A. D. Smith, P. Bendich, and J. Harer, *Persistent obstruction theory for a model category of measures with applications to data merging*, Trans. Amer. Math. Soc. Ser. B **8** (2021), 1–38.

[32] A. Templeton, T. Conerly, J. Marcus, J. Lindsey, T. Bricken, B. Chen, A. Pearce, C. Citro, E. Ameisen, A. Jones, H. Cunningham, N. L. Turner, C. McDougall, M. MacDiarmid, C. D. Freeman, T. R. Sumers, E. Rees, J. Batson, A. Jermyn, S. Carter, C. Olah, and T. Henighan, *Scaling monosemanticity: extracting interpretable features from Claude 3 Sonnet*, Anthropic Research, May 2024.

[33] T. Bricken, A. Templeton, J. Batson, B. Chen, A. Jermyn, T. Conerly, N. L. Turner, C. Anil, C. Denison, A. Askell, R. Laird, Y. Dreber, N. Schiefer, Z. Chi, D. Amodei, and C. Olah, *Towards monosemanticity: decomposing language models with dictionary learning*, Anthropic Research, October 2023.

[34] J. Engels, E. J. Michaud, and M. Tegmark, *Not all language model features are one-dimensionally linear*, Proc. ICLR 2025.

[35] L. Gao, T. D. la Tour, H. Tillman, G. Goh, R. Troll, A. Radford, I. Sutskever, J. Leike, and J. Wu, *Scaling and evaluating sparse autoencoders*, Proc. ICLR 2025.

[36] M. Huh, B. Cheung, T. Wang, and P. Isola, *The Platonic Representation Hypothesis*, Proc. ICML 2024.

[37] Y. Bansal, P. Nakkiran, and B. Barak, *Revisiting model stitching to compare neural representations*, Proc. NeurIPS 2021.

[38] H. Thasarathan, F. Muir, A. Khakzar, and Y. Motamedi, *Universal sparse autoencoders: interpretable cross-model concept alignment*, Proc. ICML 2025.

[39] N. Seely, *Sheaf cohomology of linear predictive coding networks*, NeurIPS 2025 Workshop, arXiv:2511.11092.

[40] I. Khemakhem, D. P. Kingma, R. P. Monti, and A. Hyvärinen, *Variational autoencoders and nonlinear ICA: a unifying framework*, Proc. AISTATS 2020.

[41] T. Schmid, *Applied sheaf theory for multi-agent artificial intelligence (reinforcement learning) systems: a prospectus*, University of Chicago, arXiv:2504.17700, April 2025.

[42] B. Gavranović, P. Lessard, A. Dudzik, T. von Glehn, J. G. Barata, and P. Veličković, *Position: categorical deep learning is an algebraic theory of all architectures*, Proc. ICML 2024.

[43] J. Komkov, *The Coherence Fee: Edge-Local Blindness at the String-Table Seam*, companion paper, 2026.

[44] J. Komkov, *The SHEAF Protocol: Topological Diagnostics for Heterogeneous Multi-Agent Coordination*, companion paper, 2026.