

Higher Categorical Semantics of an Embedding-Financed Proxy Company

Directed operations, groupoidal equivalence, information retention, and subsidy viability

Anonymous author

25 October 2026

Abstract

We model a proxy company that retains request-content embeddings and sells authorized information services to subsidize request handling. An internal higher category separates execution from equivalence. We prove retention factorization, information-recovery criteria, compositional privacy, audit descent, and cash-flow viability, and construct a finite consistency model.

1. Scope and interpretation

The company forwards requests, returns upstream responses, computes embeddings, and erases transient payloads. Permitted uses of retained representations generate revenue subsidizing the service.

“Only keeps embedding vectors” means that **durable request content is mediated by a declared embedding representation**. Versions, capabilities, expiry, accounting, and deletion handles are separately declared and their leakage counted. Requiring every durable company bit to be a vector would exclude these ordinary operational records.

We assume a bounded horizon, finite encoding bounds, bounded storage, and finite interfaces. Prices, demand, authority, and physical enforcement are external assumptions. No deployment, legal compliance, or market forecast is asserted.

Our contribution is a joint specification and its elementary consequences. Higher categories, descent, stochastic semantics, and differential privacy are established foundations [1-5]; the results below specialize them to this business arrangement.

2. An internal higher category of the company

Definition 2.1 (audit environment). Let $(\mathcal{A}, J)$ be a small site of declared audit contexts: subsystems, regions, and bounded observation intervals, with maps for restriction. Covering families jointly inspect every storage surface relevant to the claim under audit. Set

$$\mathcal{E} = \text{Sh}_\infty(\mathcal{A}, J).$$

This is a chosen semantic environment [1]; an actual deployment may have unlisted storage surfaces.

Definition 2.2 (execution object). An execution semantics is a complete Segal object $\mathcal{C}_\bullet$ in $\mathcal{E}$. Its objects are configurations; its arrows are directed executions; the Segal equivalences

$$\mathcal{C}_k \simeq \mathcal{C}_1 \times_{\mathcal{C}_0} \cdots \times_{\mathcal{C}_0} \mathcal{C}_1$$

encode coherent composition. For configurations s, t , $\text{Map}_{\mathcal{C}}(s, t)$ is the space of executions and certified equivalences between them. Higher paths express coherence between certificates. A concrete starting model is the complete Segal nerve of an ordinary transition category, regarded as constant sheaf data. Nontrivial higher homotopies are optional additional structure.

The maximal subgroupoid $\mathcal{C}^{\simeq}$ retains configuration equivalences; the fundamental groupoid of a mapping space retains executions and paths between them. Storing, deleting, forwarding, and spending remain directed. HoTT identity types describe paths and coherence [2]; they do not make these actions invertible.

Definition 2.3 (company). A company model is a tuple

$$\mathbf{Co} = (\mathcal{C}_{\bullet}, \text{Obs}, \Gamma, T, \text{Allow}, M, g).$$

Here Obs records observations, Γ retention representations, T durable-state channels, Allow authorization, M releases, and g cash/privacy grades. All restrict to contexts with declared descent conditions.

A morphism is a functor with coherent comparison maps preserving these structures; an equivalence has an inverse up to coherence. Identical answers need not imply identical upstream bills: resource preservation is an additional condition.

Organizationally, a functor $\mathcal{O} : \mathcal{B}^{\text{op}} \rightarrow \text{Cat}_{\infty}$ assigns admissible processes to departments or policy contexts. Its Grothendieck construction $\int \mathcal{O} \rightarrow \mathcal{B}$ is a Cartesian fibration recording coherent workflow restrictions.

For an explicit groupoidal symmetry, take two identical replicas and a factor BC_2 , the one-object groupoid with automorphism group C_2 . Its loop exchanges replica labels; all observations and grades ignore this label. The product $\mathcal{C} \times BC_2$ retains directed execution and a nontrivial configuration symmetry.

3. Proxy execution and the retention contract

Let X be encoded requests, Y responses including failures, Θ model versions, and E_{θ} finite quantized vector spaces. An embedding is a declared function

$$e_{\theta} : X \longrightarrow E_{\theta}.$$

The letter e does not assert an injective mathematical embedding. Write $\mathcal{D}(A)$ for probability distributions on a finite set A . An upstream service is a Markov kernel $K : X \rightarrow \mathcal{D}(Y)$. Composition is

$$(L \circ K)(z \mid x) = \sum_y L(z \mid y)K(y \mid x).$$

Finite kernels give symmetric monoidal probabilistic semantics [3]; $\mathcal{C}_{\bullet}$ records execution structure. The kernel category is not asserted to be a topos.

A configuration is $s = (b, V, m, k, a)$: volatile payload buffers b , retained versioned vectors V , declared metadata m , cash k , and privacy expenditure a . A successful request has the protocol

receive ; (forward || encode) ; return ; erase.

Erasure waits for both branches. Errors, cancellations, retries, and crash recovery need explicit erase-or-expire transitions. Responses are transient unless their embeddings are included in Γ .

Let H be terminal histories and $\Gamma : H \rightarrow Z$ retain selected versioned vectors and declared metadata. Ordering, identifiers, lengths, timing, indexes, logs, trained parameters, and backups count as durable observations. The full durable-state channel is $T : H \rightarrow \mathcal{D}(S)$.

Theorem 3.1 (retention factorization). The following are equivalent:

1. $T = \bar{T} \circ \Gamma$ for a channel $\bar{T} : \text{im } \Gamma \rightarrow \mathcal{D}(S)$.
2. For every h, h' with $\Gamma(h) = \Gamma(h')$, $T(h) = T(h')$ as distributions.

Moreover $\bar{T}$ is unique on $\text{im } \Gamma$.

Proof. Factorization gives (2) by substitution. Conversely define $\bar{T}(z) = T(h)$ for any representative with $\Gamma(h) = z$. Condition (2) makes this well-defined. Every z in the image has a representative, which also proves uniqueness. $\square$

This is noninterference relative to Γ , not anonymity. A strict embedding-only content claim also requires metadata to factor through the vectors and public exogenous inputs; otherwise the declared metadata constitute exceptions.

Proposition 3.2 (inductive preservation). Suppose durable state initially depends only on declared public inputs, and transition selection and updates use only the previous durable state, the newly permitted representation, declared metadata, and fresh randomness independent of the payload. Then its law factors through the **ordered** sequence of these permitted inputs. If terminal transitions empty all volatile buffers, the terminal configuration contains no additional volatile payload.

Proof. Induct on the number of transitions, composing their kernels. The inductive state distribution and next permitted input determine the next distribution. Apply Theorem 3.1 to the resulting channel. The buffer assertion is the stated terminal invariant. $\square$

Multiset retention additionally requires permutation invariance; ordered indexes may violate it. Retained fine-tuned parameters must also factor through the representation.

4. What the vectors reveal

Theorem 4.1 (recoverable attributes). For a deterministic representation $e : X \rightarrow E$ and an attribute $q : X \rightarrow A$, a decoder $d : \text{im } e \rightarrow A$ satisfying $q = d \circ e$ exists exactly when q is constant on every fiber of e .

Proof. A factorization is constant on each fiber. Conversely put $d(e(x)) = q(x)$; fiber constancy makes this independent of x . $\square$

Thus an injective representation on the supported request set admits exact reconstruction, irrespective of its numerical appearance. Even a noninjective representation can retain a sensitive attribute exactly. For $X = \{0, 1\}^2$ and $e(p, s) = p$, the first bit is perfectly recoverable and the second is discarded under a uniform product prior. Encoding into fewer coordinates alone proves neither property. Empirical embedding inversion is also documented by Morris et al. [5]; their experimental findings are not universal reconstruction guarantees.

Proposition 4.2 (best attribute prediction). For finite random variables X , $E = e(X)$ and $Q = q(X)$, the optimal probability of correctly predicting Q from E is

$$\text{Succ}^*(Q \mid E) = \sum_v \max_{u \in A} \Pr(Q = u, E = v).$$

Proof. For each v , the decoder maximizes the corresponding joint probability by choosing an attribute of maximal mass. Randomizing is a convex combination of these choices and cannot improve the maximum. Sum over v . $\square$

Utility and leakage depend on the attribute and prior; losing exact text need not eliminate sensitive information. Perfect decoding is characterized by Theorem 4.1.

Homotopical caution. Numerical closeness is not an identity path. Since $\mathbb{R}^d$ is contractible, its ambient homotopy type alone cannot encode text equivalence. Quantized vectors here are discrete. Semantic groupoids require certified invertible transformations preserving observations; higher paths concern their coherence, not cosine similarity.

5. Authorized release and privacy accounting

Revenue may come from aggregate queries, retrieval, or licensed representations. Each release r needs a witness of $\text{Allow}(m, r, t)$ for its purpose, scope, expiry, and downstream permissions. This specifies authority without proving legal compliance. Withdrawal cannot undo information already delivered independently.

For an analytical product let D be a dataset of retained vectors and $D \sim D'$ a declared user-level adjacency relation with contribution caps. A channel M is pure ε -differentially private if

$$\Pr[M(D) \in B] \leq e^\varepsilon \Pr[M(D') \in B] \quad (D \sim D')$$

for every event B , with $\varepsilon \geq 0$ [4]. Include visible refusals and charges in the release transcript. This does not protect unnoised internal databases, raw upstream requests, or undeclared side channels.

Theorem 5.1 (postprocessing and sequential release). If M is ε -private and L is a data-independent channel, then $L \circ M$ is ε -private. If, for every fixed preceding transcript, release i is conditionally ε_i -private, the adaptive transcript is $\sum_i \varepsilon_i$ -private.

Proof. In finite spaces the event inequality implies the singleton inequality. For postprocessing, multiply singleton inequalities by $L(b \mid y)$ and sum over y and the output event. For composition, factor the probability of each transcript into conditional probabilities. Multiply their inequalities, obtaining the factor $\exp(\sum_i \varepsilon_i)$, then sum over transcripts in an event. Zero-probability terms obey the same inequalities. $\square$

Adaptive query choices may depend on previous released outputs; they cannot inspect the private dataset through an unaccounted channel. A conditional budget guard admits a release only when $a + \varepsilon_i \leq \varepsilon_{\max}$. Budget resets require a new privacy analysis. Contribution limits matter: request-level adjacency does not imply user-level privacy for users issuing arbitrarily many requests.

Proposition 5.2 (deterministic release obstruction). On a connected adjacency graph, a deterministic channel with finite pure privacy parameter is constant.

Proof. If adjacent datasets have different outputs, the singleton event containing one output has probabilities 1 and 0, contradicting the privacy inequality. Equality propagates along graph paths. $\square$

Consequently selling exact nonconstant vectors has no finite pure privacy guarantee under such adjacency. An $\varepsilon = 0$ release is dataset-independent on each connected component; a buyer whose only value

is improved prediction of varying dataset attributes receives no such information. Positive revenue can still come from other services or exogenous benefits. This is a conditional information constraint, not a theorem that privacy precludes commerce.

6. Audit descent and deletion

Let $\mathbf{C}fg$ be the sheaf of structured configurations and $P \hookrightarrow \mathbf{C}fg$ the subobject expressing a retention or authorization invariant. A predicate is (-1) -truncated: its witnesses are propositions, rather than potentially many incompatible claims of fact.

Theorem 6.1 (local-to-global audit). For a genuine cover $\{U_i \rightarrow U\}$ and $s \in \mathbf{C}fg(U)$, the predicate $P(s)$ holds iff it holds for every restricted configuration $s|_{U_i}$.

Proof. Restriction proves one direction. For the other, the pullback of $P \rightarrow \mathbf{C}fg$ along s is a subterminal sheaf over U . Its local sections agree on overlaps because their type is propositional. Descent glues them to a global section. $\square$

The sheaf-local hypothesis is essential: regional audits cannot establish omitted cross-region constraints. Sheafification cannot erase hidden backups. Downstream copies need separate contracts and observations.

For implementation equivalence, an identity between configurations transports predicates. More generally, a structured equivalence preserving P induces $P(s) \simeq P(t)$. Univalence can identify equivalence with identity only in an appropriate univalent universe [2]. It does not turn an arbitrary implementation similarity into certified equivalence.

7. Financing and temporal solvency

For period t , let n_t be served requests, p_t the average user payment per request, c_t fully allocated variable cost per request, F_t fixed expenditure, and R_t **received net revenue** from permitted embedding-derived products. Allocated cost must include embedding computation, upstream service, accumulated storage, egress, refunds, and any variable release costs not already netted from R_t . We assume these amounts are well-defined; we do not estimate them from embeddings.

With liquid reserves k_t , cash obeys

$$k_{t+1} = k_t + n_t(p_t - c_t) + R_t - F_t.$$

Thus $p_t < c_t$ defines a variable service subsidy of $n_t(c_t - p_t)$. Revenue is not additive in vector count by default: duplicates, demand, permissions, privacy budgets, and extraction risk can change its value.

Theorem 7.1 (reserve criterion). Put $d_t = n_t(c_t - p_t) + F_t - R_t$. For a horizon of N periods, all period-boundary reserves are nonnegative exactly when

$$k_0 \geq k_{\min} := \max \left(0, \max_{1 \leq j \leq N} \sum_{t=0}^{j-1} d_t \right).$$

Proof. Iterating the recurrence gives $k_j = k_0 - \sum_{t < j} d_t$. Requiring every $k_j \geq 0$ gives exactly the stated inequalities. $\square$

Intra-period solvency needs event-level periods or the same prefix test on individual payments. Positive expected profit implies neither this pathwise condition nor a bound on default probability.

Corollary 7.2 (steady subsidy and bounded revenue). With constant n, p, c, F, R , indefinite operation without reserve depletion requires and suffices for nonnegative period drift that

$$R \geq n(c - p) + F.$$

If $c > p$ and $R \leq R_{\max}$, a necessary bound is $n \leq (R_{\max} - F)/(c - p)$. If $R_{\max} < F$, no positive subsidized volume satisfies the bound.

An execution may be graded by $g(f) = (\Delta k_f, \varepsilon_f)$ with $g(hf) = g(h) + g(f)$ and $g(1) = (0, 0)$. The privacy coordinate is a conservative sequential-composition bound; cash is signed. Cash admissibility is a **prefix** property, not merely a total-grade property. Parallel grades add only under the declared accounting and privacy assumptions. Homotopies admitted as company equivalences preserve these annotations.

8. A finite consistency model and synthetic example

Theorem 8.1 (existence of the combined structure). There is a nontrivial model with a proxy, content retention through noninjective embeddings, informative private releases, and positive subsidy financing under explicitly assumed payments.

Proof. Take requests $x = (p, s) \in \{0, 1\}^2$, a deterministic upstream response $p \text{ xor } s$, and $e(p, s) = p$. Forward using volatile buffers; retain only p and declared policy/accounting data; then erase both request and response. The retention updates satisfy Proposition 3.2.

For one capped user, release randomized response

$$\Pr(M(p) = p) = \frac{e^\varepsilon}{1 + e^\varepsilon}, \quad \Pr(M(p) = 1 - p) = \frac{1}{1 + e^\varepsilon}.$$

Each output likelihood ratio is at most e^ε . Hence the channel is ε -private and informative for $\varepsilon > 0$. With an independent uniform prior and metadata independent of s , the durable analytics transcript reveals nothing about s . This is specific to this example; the service response has a different observation boundary.

Take a finite horizon, rational cash states, explicit authorization, and accounted releases. Assume product payments satisfy Corollary 7.2 and reserves satisfy Theorem 7.1. The transition category, optionally multiplied by BC_2 , has a complete Segal nerve; use the one-context site. This constructs the combined model. Assumed payments establish compatibility, not actual demand. $\square$

For illustration, suppose one month has $n = 1,000,000$, $p = \$0.003$, $c = \$0.012$, $F = \$6,000$, and $R = \$18,000$. The service subsidy is \$9,000 and net cash gain is \$3,000. If fixed costs are prepaid, user fees and variable costs accrue together, and product revenue arrives last, minimum opening liquidity is \$15,000. Every number is synthetic. It is not evidence of willingness to pay, a privacy-parameter recommendation, or a forecast.

9. Evidence, limits, and authorship

The standard-library Python artifact checks finite retention channels, a metadata side channel, attribute decoding, randomized-response privacy, and cash prefixes. These validate examples; the proofs are not proof-assistant certificates. Higher categorical constructions use standard definitions, not a new foundational equivalence theorem.

No production service or revenue experiment was implemented. Deployment requires evidence for crash behavior, retention surfaces, leakage, contribution bounds, and demand. The specification’s consistency supplies neither privacy certification nor commercial validation.

Authorship and assistance. Anonymous human author, responsible for the research direction and submission. OpenAI ChatGPT assisted with formulation, literature retrieval, exposition, and computational checks. No human identity or contact address is included. No prior peer review is claimed.

References

- [1] Jacob Lurie. *Higher Topos Theory*. Annals of Mathematics Studies 170, 2009. Author’s revised text: <https://www.math.ias.edu/~lurie/papers/HTT.pdf>.
- [2] The Univalent Foundations Program. *Homotopy Type Theory: Univalent Foundations of Mathematics*. 2013. <https://homotopytypetheory.org/book/>.
- [3] Tobias Fritz. “A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics.” *Advances in Mathematics* 370, 107239, 2020. <https://arxiv.org/abs/1908.07021>.
- [4] Cynthia Dwork and Aaron Roth. *The Algorithmic Foundations of Differential Privacy*. *Foundations and Trends in Theoretical Computer Science* 9(3-4), 211-407, 2014. <https://www.cis.upenn.edu/~aaronh/privacybook.html>.
- [5] John Morris, Volodymyr Kuleshov, Vitaly Shmatikov, and Alexander Rush. “Text Embeddings Reveal (Almost) As Much As Text.” *EMNLP*, 12448-12460, 2023. <https://aclanthology.org/2023.emnlp-main.765/>.