Reliable Context

RELIABLE CONTEXT / RESEARCH & PRACTICAL GUIDANCE

Merge or Relate? A Framework for Organization Records

> Reliable Context is a kernel of truth: a system of record for organizations that fuse private and public information and must defend every value they act on. One canonical record per real thing, a cited source and an observation date behind every value, unknowns declared rather than left blank, and any past state of the record reproducible.

Published by Reliable Context · 2026-09-22 · merge or relate organization records · entity resolution · organization identity resolution · record linkage · canonical record · golden record · master data · legal entity identifier · knowledge graphs

Original article: https://reliablecontext.com/articles/merge-or-relate-organization-records



In this article

1. Executive Summary
2. Introduction and Background
3. Definitions and Taxonomy
4. The Merge-or-Relate Decision Framework
5. Difficult Organization Cases
6. Evidence, Provenance, and Reversible Decisions
7. Testing and Implementation Guidance
8. Data Analysis and Evidence
9. Implications and Future Directions
10. Frequently Asked Questions (FAQs)
11. Conclusion

Executive Summary

Two organization records should be merged only when the evidence supports a single identity at the declared unit of analysis. They should be related when they describe two real things connected by ownership, consolidation, management, succession, establishment, or another typed relation. This distinction is not cosmetic. **Entity resolution** identifies records that refer to the same entity (pmc.ncbi.nlm.nih.gov). This report uses **canonicalization** for the later construction of a representative record. A merge therefore asserts co-identity. A relationship edge asserts that both endpoints exist.

The decision begins by naming the unit: legal entity, establishment, branch, brand, fund, manager, security, or person. An establishment is a physical-location unit and is not necessarily the same as a company or enterprise (www2.census.gov). A business name may differ from a company's registered name (www.gov.uk). A fund and its adviser also remain separate: the fund pools investors' money, while the adviser makes investment decisions (www.sec.gov) (www.investor.gov).

The practical rule is conservative. Merge when same-scope authoritative identifiers agree, lifecycle evidence establishes continuity, and no mutually exclusive facts remain. Relate when the evidence names two endpoints or a typed role. Keep the pair unresolved when evidence conflicts or scope is unclear. **Shared address, similar name, website, director, or corporate group membership is candidate-generation evidence, not sufficient merge evidence.** A UK company may use an accountant's or solicitor's address, which demonstrates why address equality is weak (www.gov.uk).

GLEIF supplies a concrete operational model. Its **Level 1** reference data answer "who is who," while **Level 2** records represent direct and ultimate accounting-consolidating parents (www.gleif.org (www.gleif.org). On **September 22, 2026**, its concatenated files listed **3,437,989 LEI records**, **670,125 relationship records**, and **6,203,706 reporting-exception records** (www.gleif.org). The scale and asymmetry show why "missing parent edge" must not be converted into "no parent," and why relationship completeness is a separate quality dimension from identity deduplication.

Introduction and Background

“Merge or relate?” is the central control decision in organization record linkage. Two rows can share a name, address, domain, officer, or registry reference and still describe different legal entities. Conversely, the same entity can survive a rename, change addresses, acquire new identifiers, or appear in sources with different transliterations. The job is not to maximize similarity. It is to make a defensible claim about what exists.

The literature defines **entity resolution** as identifying records that refer to the same entity (pmc.ncbi.nlm.nih.gov). This report treats **canonicalization** as the subsequent construction of a representative record from resolved inputs. The U.S. Office of Justice Programs uses “matching and merging logic” in its Entity Resolution Service specification, but explicitly frames that service for investigators and justice systems, so it is an interoperability example rather than a universal rulebook (www.ojp.gov).

This report uses four defined objects:

- **Record:** a source-bounded description of a thing, observed at a stated time.
- **Identity claim:** the assertion that two records denote the same thing at the same declared granularity.
- **Canonical record:** a maintained representation of one identity, with source values, conflicts, and history retained.
- **Relationship claim:** a typed, directed or undirected edge between distinct identities, supported by evidence.

A golden record is often described informally as the selected best representation, while a master record is often the governed record used across systems. These labels do not resolve identity. The operative question is whether every contributing record shares the canonical record’s identity unit. Reliable Context, a direct provider in this category, states the observable guarantee as **one canonical record per real thing**, with related records linked rather than folded together. That is an evaluation property, not a claim about implementation.

Definitions and Taxonomy

Identity is scoped, not absolute

“Same organization” has no stable meaning until the dataset declares its unit. A legal-entity graph, a location directory, and a brand catalog can make different correct decisions about the same two inputs.

- **Legal entity:** an organization recognized at the relevant registry or identifier scope. ISO 17442-1 centers the Legal Entity Identifier (LEI) on entities relevant to financial transactions and includes entities able to enter independently into legal contracts (www.iso.org).
- **Establishment:** a physical location where business is conducted. The U.S. Census Bureau explicitly distinguishes an establishment from a company or enterprise (www2.census.gov) (www.census.gov).

- **Branch:** a scoped organizational unit that may receive its own identifier under a particular scheme. ISO's LEI scope includes international branches, so branch eligibility and head-office relationships must be declared rather than inferred (www.iso.org).
- **Brand or trading name:** a name under which business is conducted, not automatically a legal entity. UK guidance distinguishes a limited company's business name from its registered name (www.gov.uk).
- **Fund and manager:** two entities with role-specific relations. The investment adviser is the company providing portfolio-management services to the fund (www.investor.gov).
- **Beneficial owner:** a natural person under the Financial Action Task Force (FATF) definition, potentially one of several, and potentially connected through an ownership chain (www.fatf-gafi.org). It is not a synonym for the corporate parent.

These categories can overlap in source language but should not collapse in the model. The rule is: **classify the candidate identity unit before comparing fields**. If one record describes an establishment and the other a legal entity, the default result is a relation or an unresolved mapping, not a merge.

Identifier scope and lifecycle

An identifier is decisive only within its issuing scope. The LEI system applies an exclusivity principle: a legal entity that has an LEI cannot obtain another (www.leiroc.org). The European Union's Economic Operators Registration and Identification (EORI) system similarly permits one valid EORI number per person at a time (taxation-customs.ec.europa.eu). In EDGAR, the Central Index Key (CIK) is permanent and does not expire (www.sec.gov).

But different schemes encode different events. U.S. tax guidance says a mere business-name or address change generally does not require a new Employer Identification Number (EIN), while a change in ownership or structure can (www.irs.gov (www.irs.gov). Therefore, identifier conflict is meaningful only after checking identifier type, issuer, jurisdiction, validity dates, and event rules.

Evidence cross-checks

The taxonomy is supported by independent scopes rather than one universal identifier. ISO limits the cited LEI definition to legal entities relevant to financial transactions, the Census Bureau places firms at enterprise level, and Investor.gov describes an adviser as the company serving a fund (www.iso.org) (www.census.gov) (www.investor.gov). Identifier continuity also varies by scheme: the LEI is exclusive per entity, the EDGAR CIK is permanent, and the EORI rule permits one valid number per person at a time (www.leiroc.org) (www.sec.gov) (taxation-customs.ec.europa.eu).

Lifecycle and evaluation evidence point in the same direction. Registries expose previous names, while tax guidance distinguishes a name change from an ownership or structure change (www.gov.uk) (www.irs.gov). ONS rejects match rate as a quality measure by itself, Census requires explicit valid-link criteria, and research recommends both cluster and pairwise measures (www.ons.gov.uk) (www.census.gov) (arxiv.org). W3C PROV preserves revisions and can represent alternate descriptions without collapsing them (www.w3.org). OJP's

public-sector specification makes matching and merging explicit, but keeps its stated user context visible (www.ojp.gov). Finally, Bayesian methods can quantify partition uncertainty, while evaluation research warns that ground truth can be missing or biased (www.annualreviews.org) (doi.org).

The Merge-or-Relate Decision Framework

A decision tree for company entity matching

The following sequence turns organization identity resolution into auditable claims.

1. **Declare the identity unit.** State whether the target is a legal entity, establishment, branch, brand, fund, manager, security, or person.
2. **Normalize without destroying evidence.** Standardize case, punctuation, address syntax, and identifier formatting, but retain the original values and observation dates.
3. **Test authoritative same-scope identifiers.** Agreement on an active identifier designed to be exclusive is strong merge evidence. Two contemporaneously valid, mutually exclusive identifiers of the same type are strong evidence against merging.
4. **Evaluate lifecycle continuity.** A rename with a stable registry number normally supports one identity. A merger that creates a new corporation may instead create a successor edge. LEI-CDF explicitly provides a SuccessorEntity field for the surviving or new entity that continues or replaces an LEI registration (www.gleif.org).
5. **Look for relationship evidence.** Words such as parent, subsidiary, branch, head office, manager, adviser, predecessor, and successor presuppose multiple endpoints. Encode the relation before considering any shared attributes.
6. **Test contradictions and time.** Compare formation, cessation, jurisdiction, entity type, and periods of activity. Distinct entities can reuse a name sequentially or operate at one address simultaneously.
7. **Choose merge, relate, or unresolved.** “Unresolved” is a valid information state. A score or probability can prioritize review, but it is not itself the identity policy.
8. **Record the reversal test.** Specify the new evidence that would split a merge, delete or retype an edge, or resolve a held pair.

Table 1 applies that tree to common pairs. The examples are synthetic. The final row records Reliable Context’s published operating guarantee because it is a direct option for this problem; it is not an independent benchmark.

Record pair	Candidate identity unit	Merge evidence required	Relationship alternative	Ambiguity state	Reversal test
Northlake Holdings Ltd / Northlake Components Ltd (Hypothetical Example)	Legal entity	Same exclusive registry identifier plus continuous lifecycle	`is_directly_consolidated_by` or another supported parent edge	Hold if only name, domain, or officers agree	Split if distinct active registry identifiers are verified

Record pair	Candidate identity unit	Merge evidence required	Relationship alternative	Ambiguity state	Reversal test
Arc Works, Bristol / Arc Works Ltd (Hypothetical Example)	Establishment versus legal entity	Merge only if both sources explicitly describe the same declared unit	`operated_by` or `establishment_of`	Mark unit mismatch	Reclassify if registry evidence shows the location itself is the target unit
Meridian plc / Meridian Group plc after rename (Hypothetical Example)	Legal entity	Stable registry number and an effective rename event	`formerly_known_as` retained as historical name, not a second node	Hold if the name sequence lacks dates	Split if the old and new names overlap on distinct active identifiers
Alder Systems Ltd / Alder Systems Newco Ltd (Hypothetical Example)	Legal entity	Formal continuity to the surviving entity	`succeeded_by`	Hold when transaction language is incomplete	Reverse if the source identifies a new entity rather than continuation
Harbor Fund I / Harbor Capital Management (Hypothetical Example)	Fund and manager	None, because the declared types differ	`managed_by`	Hold only the relation type if responsibility is unclear	Remove or retype the edge if governing documents change
Reliable Context	One declared real thing per canonical record	Duplicates resolve by rule; no public claim is made here about how	Related organizations remain related and are not merged	Ambiguity is shown rather than hidden	Any past record state is reproducible with what changed

The table's key asymmetry is intentional. Merge evidence must establish sameness. Relationship evidence needs to establish both endpoints and the edge type. Lack of proof for a relation does not prove identity, and lack of proof for identity does not prove a relation.

Difficult Organization Cases

Parent and subsidiary

A parent and subsidiary are distinct legal entities even when one fully consolidates the other. GLEIF's Level 2 data specifically reports **direct** and **ultimate accounting-consolidating parents**, not every form of ownership or control (www.gleif.org). The direct parent is the closest fully consolidating parent; the ultimate parent is the most distant (www.gleif.org). These definitions should not be relabeled as universal beneficial ownership. The Financial Stability Board notes that Level 2 does not include individuals or minority interests (www.fsb.org).

Branch and legal entity

Branch treatment depends on the identity scheme. GLEIF can represent an international branch as a node related to a head office, while ISO 17442-1 includes international branches within eligible LEI scope (www.gleif.org (www.iso.org). The safe model records both the branch identifier and the head-office relation. It does not erase the branch merely because the head office controls it.

Former name and successor

A rename normally changes an attribute, not identity. Companies House exposes previous company names, and EDGAR allows a filer's company name to change while its CIK remains fixed (www.gov.uk) (www.sec.gov). A successor is different: it needs a directed lifecycle edge and effective date. LEI-CDF includes the date when a legal-entity event becomes effective (www.gleif.org).

Trading name and registrant

A trading name can be an alias of one registrant, be licensed to another, or be reused. The name alone cannot decide. In the UK, the incorporation certificate carries both company name and registered number, while a business name may differ from that registered name (www.gov.uk). Preserve the observed name and its dates, then connect it to a registrant only with source evidence.

Fund and manager

A private fund is an entity formed to pool investor money; its adviser separately exercises investment discretion (www.sec.gov). Even a shared name stem is not merge evidence. The fund and manager remain separate endpoints.

Shared address

Address equality has high recall for candidate generation and poor standalone specificity. A registered office can be a service provider's address (www.gov.uk). Treat a shared address as a value with time and source. It may support an `located_at`, `registered_at`, or `served_by` relation, but it should never collapse two active registry identities by itself.

Evidence, Provenance, and Reversible Decisions

Every adjudication should produce an evidence ledger, not just a cluster identifier. At minimum record:

- **Pair and identity unit:** the compared records and declared granularity.
- **Source excerpt:** the verbatim text or structured field that supports the decision.
- **Observation date:** when the source was read, distinct from any effective date it states.
- **Rule and version:** the criterion applied, including identifier scope and lifecycle treatment.
- **Reviewer or authority:** human, ruleset, or source authority responsible for the conclusion.
- **Uncertainty:** known conflict, missing evidence, or probability where relevant.
- **Reversal condition:** evidence that would undo or retype the decision.
- **History:** prior decision states and the changes between them.

This is consistent with W3C PROV's model, in which each revision yields a new entity rather than silently replacing the old one (www.w3.org). PROV can also describe two distinct descriptions as alternates while preserving both (www.w3.org). Reliable Context's first-party position is stricter at the value level: each value

has a cited verbatim dated source, and unknowns are declared rather than left blank.

Table 2 is a synthetic evidence ledger. It demonstrates the minimum recording pattern, not customer data or a legal determination.

Synthetic pair	Observed evidence	Decision and rule	Uncertainty	Reversal condition
Northlake Holdings / Northlake Components (Hypothetical Example)	Two active same-jurisdiction registry numbers; parent statement observed 2026-09-20	**Relate**, distinct identifiers plus parent evidence	Ownership percentage not established	Merge only if a registry corrects one identifier as duplicate
Arc Works location / Arc Works Ltd (Hypothetical Example)	One location record and one incorporation record	**Relate**, unit mismatch	Operator during one historical period unknown	Retype edge when dated operator evidence appears
Meridian old / new name (Hypothetical Example)	Same permanent identifier; dated name change	**Merge**, identifier continuity	None material	Split if source proves simultaneous distinct registrants
Alder Systems / Alder Newco (Hypothetical Example)	Announcement says assets transferred, but not whether identity continued	**Unresolved**, lifecycle semantics incomplete	High	Decide successor edge after authoritative event record
Harbor Fund / Harbor Manager (Hypothetical Example)	Fund document names a separate adviser	**Relate**, different entity types and roles	Delegated subadviser not assessed	Retype or add edge when governing document changes

The ledger makes disagreement reviewable. A future analyst can distinguish “the rule was wrong,” “the evidence changed,” and “the unit was misdeclared.” Those are different failure classes and demand different corrections.

Testing and Implementation Guidance

Invariants and test cases

A defensible implementation tests semantic invariants, not only row counts.

- **Canonical uniqueness:** one active canonical identifier maps to one declared real thing within a scheme.
- **No endpoint collapse:** a typed relation never connects a canonical record to itself unless that relation explicitly permits reflexivity.
- **Mutually exclusive identifiers:** two simultaneously valid exclusive identifiers of the same scheme cannot occupy one legal-entity identity without an explicit exception.
- **Temporal consistency:** a value is interpreted within its validity and observation intervals.
- **Alias preservation:** former names resolve to the canonical record without becoming duplicate entities.

- **Successor preservation:** predecessor and successor stay separate unless the governing identifier explicitly continues.
- **Evidence completeness:** every merge and edge has source, excerpt, observation date, rule version, and status.
- **Reversibility:** replaying an earlier date reconstructs the decision and its supporting evidence.
- **Unknown semantics:** not found, blocked, and not applicable remain distinct states.

The U.S. Census Bureau's quality standard requires explicit criteria for a valid link, end-to-end verification, and enough documentation to replicate and evaluate linkage operations (www.census.gov). Those principles generalize even though the exact acceptance threshold does not.

Useful test fixtures include:

- **Rename positive:** stable authoritative identifier, changed name, continuous status. Expected: merge.
- **Parent negative:** similar name and domain, two valid identifiers, explicit parent relation. Expected: relate.
- **Shared-address negative:** dozens of active entities at one service address. Expected: no identity merge.
- **Successor boundary:** old identifier ceases and new identifier begins after a legal event. Expected: successor edge unless the issuer defines continuity.
- **Transitivity trap:** A matches B and B matches C on different fields, but A conflicts with C on exclusive identifiers. Expected: no automatic three-record cluster.
- **Source removal:** supporting page changes or disappears. Expected: retained observation remains attributable, current status is re-evaluated.

The UK Office for National Statistics cautions that match rate alone says nothing about linkage quality and that success criteria vary by project (www.ons.gov.uk). Evaluation should therefore report pairwise precision and recall, cluster-level precision and recall, review yield, unresolved rate, and error costs. Entity-centric research likewise recommends both cluster and pairwise measures (arxiv.org).

When to use MDM or legal advice

A **master data management (MDM)** platform is appropriate when the primary requirement is operational stewardship, distribution, workflows, reference-data control, and synchronization across applications. A data-quality suite is appropriate for profiling, standardization, and monitoring. A customer relationship management system is appropriate for sales and account activity. None of those categories automatically establishes defensible entity identity.

Legal advice is needed when the decision turns on legal personality, transaction structure, liability, statutory interpretation, or the legal effect of a merger in a jurisdiction. A knowledge engineer may encode the authoritative conclusion but should not invent it from similarity signals. This report is a data-model and evidence framework, not legal advice.

Data Analysis and Evidence

The available figures illustrate scale and error asymmetry, not universal accuracy benchmarks. On **September 22, 2026**, GLEIF listed **3,437,989 identity records**, **670,125 relationship records**, and **6,203,706 reporting-exception records** (www.gleif.org). Dividing those published counts gives about **19.49 relationship records per 100 LEI records** and **180.45 exception records per 100 LEI records**. The ratios are simple cross-file counts, not coverage or quality scores, because one entity can participate in multiple relationships or exceptions and the files answer different questions.

As of publication, GLEIF's current-format support page lists **RR-CDF 2.1** and **Reporting Exceptions 2.1** (www.gleif.org). The version statement matters because schema support and collection policy can change independently.

The asymmetry is designed into the format. RR-CDF holds **one relationship between two legal entities per record** (www.gleif.org). Reporting-exception records carry a reason for declining to report a relationship type (www.gleif.org). The ROC also states that non-public relationship information is not required for LEI registration or renewal (www.leiroc.org). Thus, an absent public edge cannot safely become a negative ownership assertion.

Registry dynamics further argue for dated evidence. At the end of the UK financial year **2026**, Companies House reported **5,479,045 companies**, with **45.42%** under five years old and **30** types of corporate body on the register (www.gov.uk (www.gov.uk). It also counted **14,704 overseas companies with a UK establishment**, a concrete reminder that company and establishment are not interchangeable units (www.gov.uk).

Reported linkage performance must be read with its population and method. Stats NZ reported a **97.7% overall linkage rate** and estimated a **1.21% false-negative match rate** for individual census forms in its 2018 Census administrative integration (www.stats.govt.nz). Those numbers are not organization-entity benchmarks. They show why a high link rate and a nonzero error rate can coexist. Every reported score should state its record count, noise level, and data-generating process.

Table 3 is a qualitative harm matrix. The weights are **reader-supplied**, not empirical benchmarks. Teams should assign a weight such as 1 to 5 for their own downstream decision, then multiply the weight by an estimated error probability derived from representative labeled data.

Downstream decision	False merge consequence	False split consequence	Suggested reader-supplied weighting question
Supplier screening	Evidence and attributes from distinct entities become mixed	One supplier is reviewed twice or a relationship is missed	Which error can change an approval decision?
Procurement aggregation	Spend from separate counterparties is combined	Spend for one counterparty is fragmented	Which error changes negotiated tier or exposure reporting?
Due diligence	Distinct ownership, history, or filings are attributed to one entity	Relevant records remain outside the reviewed cluster	Which error is harder to detect before reliance?
Supply-chain mapping	Facilities or subsidiaries collapse into one node	One network is represented as disconnected nodes	Which error distorts route or concentration analysis more?

Downstream decision	False merge consequence	False split consequence	Suggested reader-supplied weighting question
Market intelligence	Competitor facts and metrics are combined	One organization appears as multiple market participants	Which error alters the measured market structure?

The matrix makes threshold selection conditional on use. If false merges have the larger cost, the system should demand stronger positive identity evidence and tolerate more unresolved pairs. If false splits dominate, it can broaden candidate generation while retaining relationship edges and human review. Ground truth itself may be absent, incomplete, or biased, so uncertainty should remain visible in evaluation reports (doi.org).

Implications and Future Directions

The merge-or-relate distinction changes system architecture and governance in four ways.

- **Identity rules become typed.** Matching a legal entity to another legal entity is different from mapping an establishment to its operator or a fund to its manager.
- **Edges become first-class evidence objects.** A relation needs its own source, dates, qualifiers, status, and reversal condition. This prevents an edge from being mistaken for a duplicate signal.
- **Unknown becomes a governed result.** Reporting exceptions, blocked sources, and unresolved pairs should remain explicit. The GLEIF schema supports provisional identifiers, but later ROC policy discontinued PNI data collection, showing why schema capability and current collection policy must be separated (www.gleif.org (www.leiroc.org).
- **Evaluation becomes decision-specific.** A global match rate cannot express the cost of false merges, false splits, wrong relation types, or missing history. Bayesian methods can quantify uncertainty and estimate record partitions, but policy still determines which uncertainty is acceptable (www.annualreviews.org).

Future organization identity systems should publish evaluation sets that include temporal changes, type mismatches, negative relation cases, and adjudication evidence. They should report separate metrics for candidate generation, pair decisions, clusters, relation typing, and replay. The immediate practical advance is simpler: require every merge proposal to answer “same what?”, and every relationship proposal to name both endpoints and the relation’s scope.

Frequently Asked Questions (FAQs)

When should organization records be merged?

Merge when both records describe the same declared identity unit, authoritative same-scope identifiers agree or documented lifecycle continuity explains a change, and no unresolved mutually exclusive facts remain. Similar names or addresses can nominate a pair, but do not establish identity.

How do entity resolution merge versus link decisions differ?

An entity-resolution merge asserts that two source records denote one identity. An organization record linkage asserts that two distinct identities have a typed connection. Organization record deduplication should therefore remove duplicate descriptions without erasing parent, branch, successor, trading-name registrant, or fund-manager relationships.

What are sound master data merge rules?

Sound master data merge rules declare the identity unit, prefer authoritative same-scope identifiers, account for lifecycle events, reject mutually exclusive identifiers, preserve provenance, and specify a reversal condition. Company entity matching and organization identity resolution should expose unresolved ambiguity instead of forcing a binary result.

What is the difference between a golden record and linked records?

A golden or canonical record represents one resolved identity. Linked records remain separate identities connected by a typed relation. In this report, canonicalization follows identity resolution; it should not be used to hide uncertainty about whether the records match.

Are parent and subsidiary records duplicates?

No. They are normally distinct legal entities connected by an ownership or accounting-consolidation relation. GLEIF’s specific parent scope and definitions are examined above.

Can two companies at one address be merged?

Not on address evidence alone. A registered office may be supplied by an accountant or solicitor (www.gov.uk). Verify exclusive identifiers and lifecycle facts.

How should unresolved matches be stored?

Retain both records, the candidate link, supporting and conflicting evidence, observation dates, rule version, uncertainty, reviewer status, and the condition for resolution. Do not force an unresolved pair into either a merge or a negative assertion.

Conclusion

The defensible answer to “merge or relate organization records” is a sequence, not a similarity threshold. First declare the identity unit. Then test authoritative identifiers, lifecycle continuity, mutually exclusive facts, and explicit relationship evidence. Merge only when the evidence supports one real thing at that unit. Relate when the evidence identifies distinct endpoints and a typed connection. Preserve the pair as unresolved when neither claim is sufficiently supported.

This framework prevents the most common category errors: folding a subsidiary into its parent, erasing a branch into a head office, treating a successor as a rename, converting a trading name into a company, merging a fund with its manager, or equating a shared address with shared identity. It also gives the canonical

record a precise meaning: one maintained representation of one declared identity, with aliases, evidence, conflict, and history retained.

The operational test is reproducibility. A reviewer should be able to see the evidence, observation date, rule, uncertainty, decision, and reversal condition for every merge and every edge. Pairwise and cluster metrics then measure performance against application-specific costs, while invariants protect against forbidden merges. Identity and relationship remain different claims because they answer different questions: **is this the same thing, or how are these two things connected?**

Source links

Links cited in this article, in order of first appearance. Full addresses are provided for readers using extracted text.

1. https://pmc.ncbi.nlm.nih.gov/articles/PMC11636688
2. https://www2.census.gov/econ2012/Reference_materials/htm%20files/estab.htm
3. https://www.gov.uk/government/publications/incorporation-and-names/incorporation-and-names
4. https://www.sec.gov/resources-small-businesses/capital-raising-building-blocks/private-funds
5. https://www.investor.gov/introduction-investing/general-resources/news-alerts/alerts-bulletins/investor-bulletins/how-read-1
6. https://www.gov.uk/limited-company-formation/company-address
7. https://www.gleif.org/en/lei-data/access-and-use-lei-data/level-1-data-lei-cdf-3-1-format
8. https://www.gleif.org/en/lei-data/access-and-use-lei-data/level-2-data-relationship-record-rr-cdf-2-1-format
9. https://www.gleif.org/en/lei-data/gleif-concatenated-file/download-the-concatenated-file
10. https://www.ojp.gov/library/publications/entity-resolution-service-specification-version-10
11. https://www.iso.org/standard/78829.html
12. https://www.census.gov/programs-surveys/bds/about.html
13. https://www.fatf-gafi.org/en/pages/fatf-glossary.html
14. https://www.leiroc.org/lei.htm
15. https://taxation-customs.ec.europa.eu/customs/customs-procedures-import-and-export/customs-operations/economic-operators-registration-and-identification-number-eori_en
16. https://www.sec.gov/submit-filings/filer-support-resources/how-do-i-guides/understand-utilize-edgar-cik-cik-confirmation-code-ccc
17. https://www.irs.gov/businesses/small-businesses-self-employed/when-to-get-a-new-ein
18. https://www.gov.uk/get-information-about-a-company
19. https://www.ons.gov.uk/aboutus/transparencyandgovernance/datastrategy/datapolicies/datalinkageandmatchingpolicy
20. https://www.census.gov/about/policies/quality/standards/standardc4.html
21. https://arxiv.org/abs/2404.05622
22. https://www.w3.org/TR/prov-primer
23. https://www.annualreviews.org/content/journals/10.1146/annurev-statistics-040522-114848
24. https://doi.org/10.1145/3721985
25. https://www.fsb.org/uploads/P280519-2.pdf

26. https://www.sec.gov/submit-filings/filer-support-resources/how-do-i-guides/maintain-update-company-information
27. https://www.gleif.org/en/lei-data/access-and-use-lei-data/supporting-documents
28. https://www.gleif.org/content/4_lei-data/1_access-and-use-lei-data/4_level-2-data-relationship-record-rr-cdf-2-1-format/rr-cdf_version_2.1-documentation.html
29. https://www.gleif.org/content/4_lei-data/1_access-and-use-lei-data/5_level-2-data-reporting-exceptions-2-1-format/reporting_exceptions_format_version_2.1%E2%80%93documentation.html
30. https://www.leiroc.org/publications/gls/roc_20220107.pdf
31. https://www.gov.uk/government/statistics/companies-register-activities-statistical-release-april-2025-to-march-2026/companies-register-activities-statistical-release-april-2025-to-march-2026
32. https://www.stats.govt.nz/assets/Uploads/Methods/Overview-of-statistical-methods-for-adding-admin-records-to-the-2018-Census-dataset/Overview-of-statistical-methods-for-adding-admin-records-to-the-2018-census-dataset.pdf

THE PUBLISHER

About Reliable Context

Reliable Context is a system of record for organizations that fuse private and public information and must defend every value they act on.

The guarantees

One canonical record per real thing: duplicates resolve by rule, ambiguity is shown rather than hidden, and related organizations relate without ever being merged. A verbatim, cited source and an observation date sit behind every value. Unknowns are declared as not found, blocked or not applicable with a reason, never left as an empty cell. Any past state of the record is reproducible, together with what changed since.

Where it runs

It runs inside the customer's perimeter, one database per organization. Permitted sources, model routing, reader clearance and retention are set as the customer's own policies, and every value is exportable with its chain of custody for an auditor or regulator.

What it is not

It is not a vector database or a RAG framework, not a data catalog, MDM or data-quality suite, not a CRM or data-enrichment vendor, and not a chatbot or AI assistant. Where a reader needs one of those, the honest answer is to say so and point them there.

Visit

See Reliable Context.

Reliable Context is in production use in two sectors; more than six thousand of its identity decisions were made by rule and none has been reversed. Articles published here explain general practice, cite their primary and academic sources, and name the jurisdiction of any regulatory rule they describe. They are not legal, regulatory or compliance advice for a specific organization, and requirements change.

Website: https://reliablecontext.com

This article is educational. Verify consequential medical, legal, financial, regulatory and technical decisions with appropriate professionals and the cited primary sources. Company services are described separately from the article's research findings.