Reliable Context

RELIABLE CONTEXT / RESEARCH & PRACTICAL GUIDANCE

Privacy-Preserving Truth Discovery: A Utility Audit

> 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 · 2026-10-20 · privacy preserving truth discovery · local differential privacy · truth discovery accuracy · privacy utility tradeoffs · secure aggregation · source reliability · verifiable computation · knowledge fusion

Original article: https://reliablecontext.com/articles/privacy-preserving-truth-discovery-utility-audit



In this article

1. Executive Summary
2. Introduction and Background
3. Key Changes
4. Privacy guarantees and threat models
5. Evidence from truth-discovery methods
6. Implementation Considerations and Process Changes
7. Verifiability and evidence access
8. Data Analysis and Evidence
9. Implications and Future Directions
10. Frequently Asked Questions (FAQs)
11. Conclusion

Executive Summary

Privacy-preserving truth discovery estimates a value from conflicting contributions while limiting specified disclosures about those contributions. The qualification “specified” is decisive. An answer-level local guarantee, protection of a contributor’s complete vector, concealment of source identity, and confidentiality of inferred reliability are different evaluation targets. Differential privacy makes its unit precise through neighboring inputs; the central and local models also place trust in different actors. A meaningful utility comparison must fix both the protected unit and the observer before comparing accuracy. (nvlpubs.nist.gov) (www.nist.gov)

The categorical study by **Li and colleagues (2018)** (www.kdd.org) illustrates why this is an estimation problem as well as a privacy problem. Its one-layer mechanism applies a common perturbation probability; its two-layer design samples private user probabilities. The official abstract reports an advantage for the latter at matched stated guarantees. That finding concerns its mechanism and tasks, rather than arbitrary private knowledge fusion. Its table lists **26 labeled questions** for Peer Grading and **159 labeled questions** for Duchenne Smile, making label coverage an essential qualification. (www.kdd.org) (cse.buffalo.edu)

Secure aggregation addresses a different disclosure boundary: individual inputs can remain concealed while the aggregate becomes available. The aggregate itself remains an information channel. Cryptographic protection can therefore complement differential privacy, but cannot be substituted for a local privacy budget or an output guarantee without a separate argument. Similarly, authenticated provenance establishes a traceable statement, rather than establishing the statement’s truth. (research.google) (www.w3.org)

This report proposes an evaluation card covering input domain, protected unit, adversary, privacy definition, budget, composition, comparator, labels and uncertainty. Its hypothetical calculation shows an answer changing from A to B when perturbation changes reliability weights, even with target answers held fixed. A second calculation shows why **ε = 1 per answer** and **ε = 1 per contributor vector** cannot be treated as equivalent budgets under basic composition. These are illustrative calculations, not published benchmarks.

(differentialprivacy.org) The practical conclusion is to compare methods only within compatible threat models and data types, report answer reversals and abstention alongside error, and treat an empirical privacy audit as a supplement to the formal argument. (papers.neurips.cc) (csrc.nist.gov)

Introduction and Background

Truth discovery separates **source claims** from **source reliability**. A claim states an answer to an object-level question; a reliability estimate influences how much that answer contributes to aggregation. The official description of the categorical method studied here explicitly combines estimated user quality with weighted answers. Agreement can therefore influence both the selected value and the weights used to select it. (www.kdd.org)

Consider independent contributors answering whether a particular identifier refers to the same entity. The object is the identifier pair, the input is a categorical judgment, and the output is an estimated match decision. A numerical question, such as estimating an object's measured magnitude, has a different input domain and loss function. This report uses such questions to define evaluation requirements, without assuming that one estimator works for both.

Two deployment settings must remain distinct. Data already residing inside one organization’s perimeter can be governed through authorized access and a trusted computation boundary. Independent contributors sending perturbed reports to a collector instead require a guarantee against that collector’s observations. NIST’s description of central differential privacy assumes a trusted curator; its local model places perturbation with each contributor before transmission. Neither a perimeter nor access controls alone establish differential privacy. (www.nist.gov)

This distinction matters for organizations combining restricted and public information, including the readership of Reliable Context. The publisher’s position is that of a system-of-record provider. It is not evaluated here as an implementation of a locally private research method. No assertion about its internal mechanism follows from the literature.

The report treats **utility** as preservation of a declared decision task and **privacy** as a declared limitation on an observer. Provenance is a related requirement: W3C describes it through entities, activities and people associated with producing information. A traceable result and a privacy-preserving result consequently need separate evidence. (www.w3.org) This audit is anchored to **October 20, 2026**, its publication date; inaccessible material is identified explicitly, and proposed experiments are distinguished from measured studies.

Key Changes

From a privacy label to a protected unit

The first required change is in reporting practice, rather than a claim that a new standard has replaced an old one. “Private truth discovery” should be expanded into a statement about neighboring inputs. NIST identifies the choice of neighboring datasets as part of the guarantee, while OpenDP requires a dataset metric and

distance. These choices determine what change the observer must have difficulty distinguishing. (nvlpubs.nist.gov) (docs.opendp.org)

For a **local differential privacy (LDP)** evaluation, let x and x′ be permissible alternative inputs for one contributor and S an output event. Pure ε-LDP requires $\Pr[M(x) \in S] \leq \exp(\varepsilon) \Pr[M(x') \in S]$ for the declared alternatives. An approximate version permits an additional δ on the right. Whether x denotes one answer or an entire contribution vector must be stated independently of the formula. An additive δ is a relaxation of the inequality, rather than a universal probability that a particular person's information will be disclosed. (nvlpubs.nist.gov)

A cell-level comparison can vary one answer while fixing the rest. A vector-level comparison can allow the whole contribution to change. These are different adjacency relations. Likewise, concealing values associated with a visible pseudonym does not, by itself, establish anonymity of the source. The proposed evaluation card therefore treats identity and participation as separate observables, consistent with the broader observer model used in protocol privacy analysis. (www.rfc-editor.org)

From a nominal budget to a release history

A displayed ε without a release history is incomplete. OpenDP's basic composition adds privacy expenditure across queries; its adaptive composition also allows later mechanisms to depend on earlier answers. Kairouz, Oh and Viswanath study the degradation associated with interactive querying. A paper should identify the accounting theorem it actually uses, rather than borrowing a more favorable bound informally. (docs.opendp.org) (proceedings.mlr.press)

The distinction between computation and new disclosure is equally important. OpenDP explains that postprocessing an already private result cannot increase its distinguishing divergence. Recomputing a weight from an existing transcript is consequently different from querying the raw contributor again. An audit should identify every fresh raw-data access and every new randomized report, then account for the combined transcript. (docs.opendp.org)

From absence of disclosure to bounded additional inference

Differential privacy limits the additional information attributable to an individual's contribution under its definition. It does not promise that every sensitive attribute remains unpredictable. Harvard's explanation explicitly frames the parameters in terms of additional risk, while Ghosh and Kleinberg distinguish differential privacy from inferential privacy under correlated data. (privacytools.seas.harvard.edu) (arxiv.org)

For truth discovery, the implication is methodological: evaluate attacks on the transcript, source weights, selected values and metadata separately. If contributors share an underlying source, a prediction may exploit that dependence. An unsuccessful attack is useful evidence about that attack's tested conditions, but is insufficient to replace the mathematical guarantee. Statistical auditing research supports checking implementations; it does not turn an unobserved disclosure into a proof of privacy. (papers.neurips.cc)

Privacy guarantees and threat models

Table 1 separates observation boundaries before any utility comparison. The entries describe analytical settings, not recommended architectures.

Setting	Protected unit and observer	Who receives raw inputs?	Permitted output and assumptions to record
One organizational perimeter	Policy-defined records against unauthorized readers; any differential privacy unit must be added explicitly	Authorized systems and personnel under the stated policy	Controlled records or aggregates; access governance and inference protection are separate questions. (www.nist.gov)
Independent local reporters	Declared answer or vector against the collector	Contributor before local randomization	Randomized reports and their postprocessing; enumerate metadata and repeated reports. (www.nist.gov) (docs.opendp.org)
Cryptographic aggregation	Individual inputs against specified protocol participants	Contributors; other access depends on the protocol	Authorized aggregate; information learnable from the aggregate remains in scope. (research.google)
Two-server truth discovery	Sensory values and, where specified, reliability against each server	Depends on the sharing or encryption scheme	Estimated truth under explicit noncollusion assumptions. (researchconnect.suny.edu) (www.cs.newpaltz.edu)

The matrix prevents a common comparison error: identical output error does not establish equivalent confidentiality. A trusted curator, an untrusted collector and noncolluding servers have different knowledge. A study can be internally convincing while being unsuitable for an evaluator’s adversary. The first acceptance criterion should therefore be compatibility of observations and assumptions, followed by utility.

The **L-PPTD** abstract by Miao and colleagues describes additive homomorphic encryption with two noncolluding cloud platforms. It distinguishes protecting sensory data and reliability information from a related setting protecting sensory data alone. These are substantive differences in the protected objects, not interchangeable meanings of the word private. (researchconnect.suny.edu)

The **CEPTD** paper likewise specifies honest-but-curious clouds without collusion, with an inquirer reconstructing estimated truths. That model should be represented as such. It is not evidence that a single collector receiving locally perturbed answers has the same observation boundary. (www.cs.newpaltz.edu)

Bonawitz and colleagues’ secure aggregation work provides a useful complementary distinction: private inputs are protected except for information learnable from the aggregate. Their later protocol abstract reports proofs for honest-but-curious and malicious server settings. This audit did not retrieve the full cryptographic theorem, so it does not extend those abstracts into specific collusion or dropout thresholds. (research.google)

The evaluator should ask which party sees contribution counts, object selection, timing and participation history. RFC 6973 observes that the existence of information can remain detectable even when contents are unreadable. A value-hiding channel and a participation-hiding channel consequently require separate tests. (www.rfc-editor.org)

Evidence from truth-discovery methods

Table 2 records what the retrieved evidence supports. It deliberately omits a shared utility score where the input domains or guarantees are not matched.

Method or evidence boundary	Input and privacy statement	Comparator, labels and measured utility	Audit disposition
Li and colleagues' categorical mechanisms	Answer-domain theorem; a private user probability is sampled from a public hyperdistribution. (cse.buffalo.edu)	One-layer versus two-layer perturbation; utility evaluated as error-rate change. (cse.buffalo.edu)	Inspect joint reporting separately; do not promote the answer-level statement into a vector guarantee.
L-PPTD	Encryption with two noncolluding platforms; sensory data and reliability are specified protected objects. (researchconnect.suny.edu)	Retrieved institutional abstract does not supply a matched local-budget experiment	Different threat model; not ranked against LDP.
BLIND	Homomorphic computation of clustering on encrypted user data. (scholarsmine.mst.edu)	Abstract describes real and synthetic data evaluations, without a verified matched LDP budget	Relevant cryptographic approach; no cross-model utility winner inferred.

The evidence supports comparison of evaluation obligations, rather than an overall league table. In particular, an unavailable theorem is an unknown. It is neither a negative result nor evidence that a method has a weaker guarantee.

The categorical result and its interpretation

Li and colleagues sample a user-specific perturbation probability. Their theorem assumes that this probability is private while the hyperdistribution is known to the server. The official abstract attributes improved utility to using estimated quality to absorb perturbation loss. That is the paper's reported finding; this report makes no quantitative transfer claim to another task. (cse.buffalo.edu) (www.kdd.org)

The audit question is what the observer learns jointly from the reports available under that sampling procedure. The retrieved proof concerns answer values; a whole-vector or repeated-release account was not verified. This is a boundary on the conclusion drawn here, rather than an assertion that the paper's stated theorem is false.

Continuous answers require a separate specification

A numerical-input experiment must declare bounds, measurement units, clipping and treatment of missing values. OpenDP notes that unbounded numerical inputs can produce unbounded sensitivity and describes public clamping bounds as a way to make sensitivity finite. A private estimate on a clipped interval and a non-private estimate on an unrestricted interval are not yet a matched utility comparison. (docs.opendp.org)

For reproducibility, **Shi's separate sensing artifact** illustrates what can be inspected: its repository supplies code and data and describes measurements from **10 mobile devices**. Those measurements are a distinct benchmark, not a substitute for the categorical study's datasets. Reusing them would require a newly specified task and truth-label procedure. (data.mendeley.com)

Implementation Considerations and Process Changes

The proposed **research evaluation card** is the deliverable of this audit. It should accompany every reported privacy-versus-utility point. NIST's guidelines explicitly address pitfalls between mathematical privacy and implementation; Census guidance separately distinguishes implementation choices from a disclosure method's inherent properties. The card makes those distinctions inspectable. (csrc.nist.gov) (www.census.gov)

Table 3 defines mandatory reporting fields. "Not established" should remain an admissible finding.

Card field	Required content	Comparison consequence
Protected unit	Single answer, full vector, identity, participation or reliability; exact neighboring alternatives	Equal ε is comparable only after adjacency is aligned. (nvlpubs.nist.gov)
Observer and transcript	Collector, servers, output recipient; metadata, auxiliary knowledge and collusion	A mismatch creates separate experimental strata. (www.rfc-editor.org)
Definition and accounting	Pure ε or approximate ε, δ; release schedule and accounting theorem	Report cumulative exposure, including adaptive reports. (docs.opendp.org)
Input domain	Categorical alphabet or numerical bounds, units, missingness and clipping	Re-evaluate on the same admissible domain. (docs.opendp.org)
Comparator and labels	Same fusion task without privacy; origin, coverage and separation of labels	Report uncertainty about the reference truth itself.
Utility and uncertainty	Task error, abstention, weight stability, reversals and paired intervals	Pair runs on the same test objects; state resampling assumptions. (docs.scipy.org)

The card is a reporting proposal, not a certification. A complete card can reveal that two methods should not be compared. Conversely, a missing field should not be silently filled by assuming the most favorable interpretation.

Before running an experiment, investigators should register the following design decisions:

- **Objects:** Fix the question set and eligible candidate values.
- **Sources:** State the sampling frame and independent sampling unit.
- **Labels:** Identify who established the reference values and when.
- **Coverage:** Distinguish labeled objects from all collected objects.
- **Overlap:** Measure which sources answer the same questions.
- **Dependence:** Identify shared upstream material or source clusters.
- **Missingness:** Separate no contribution from a perturbed contribution.
- **Comparator:** Freeze a non-private method using the same inputs.

These are proposed safeguards for interpreting results. Correlation is specifically relevant because inferential privacy can differ from the differential privacy statement when data are dependent. It also motivates reporting source-density strata rather than assuming all observations provide independent evidence. (arxiv.org)

The execution record should then contain:

- **Adjacency:** Write the alternatives the guarantee compares.
- **Observer:** Enumerate the information the attacker receives.
- **Randomization:** Record mechanism versions and reproducible seeds.
- **Bounds:** Freeze numerical clipping before test evaluation.
- **Budget:** Declare ε and δ for the protected unit.
- **Composition:** Include every fresh report and release.
- **Stopping:** Record adaptive stopping and budget exhaustion.
- **Access:** State who can inspect raw contributions.

The composition and input-bound requirements follow the relevant formal accounting and sensitivity concerns; the remaining items make the proposed experiment reconstructible. (docs.opendp.org)

A proposed study could test vector budgets **0.5, 1, 2 and 4**, using **50 perturbation repetitions** per fixed source-object matrix. These are design choices, not literature benchmarks or universally acceptable privacy parameters. Source-density strata should be defined from the task's observed overlap structure.

Abstention should be evaluated as a policy fixed before test scoring. An unanswered difficult object cannot be credited as a correct answer. Presenting coverage together with error prevents an apparent utility gain from being explained entirely by withholding more outputs.

Verifiability and evidence access

Verifiable privacy-preserving truth discovery needs a named verification proposition. Was a report submitted by an authorized contributor? Was a transformation executed correctly? Was the privacy accountant applied to every release? Was an estimated answer factually correct? These questions are distinct.

W3C's Verifiable Credentials model explicitly separates a credential's verifiability from the truth of its claims. Likewise, the PROV overview distinguishes provenance representation and validity constraints from the underlying factual question. A researcher should avoid treating either a valid provenance graph or an authenticated statement as an independent truth label. (www.w3.org)

A cryptographic example makes the distinction concrete. RFC 9497 describes a verifiable oblivious pseudorandom function in which a client checks that a server used a particular private key. That is a precise proposition about execution. It does not certify an external-world assertion supplied as input. This report uses the document as an example of scoped verifiability, not as a proposed truth-discovery construction. (datatracker.ietf.org)

Evidence access should therefore be described in explicit categories:

- **Public evidence:** Material whose wording and location can be independently inspected.
- **Restricted evidence:** Material available only to authorized evaluators.
- **Private reports:** Perturbed or encrypted contributions with stated guarantees.
- **Execution evidence:** Records or proofs of a specified computation.
- **Reference labels:** Adjudicated answers with a documented origin.
- **Unresolved evidence:** Material inaccessible or insufficient to support a claim.

These categories are a proposed reporting vocabulary. They separate availability from truth and make it possible to describe what an external reviewer can actually reproduce. RFC 6973's data-minimization principle supplies a reason to avoid disclosing additional material merely to make a report convenient to inspect. (www.rfc-editor.org)

A practical tension remains: verbatim source access can assist scrutiny while exposing information a local mechanism was intended to conceal. The evaluation should state which proposition can be checked without raw access and which requires a privileged reviewer. Selective disclosure and unlinkable disclosure are also separate capabilities in W3C's model, so a partially disclosed credential should not automatically be described as unlinkable. (www.w3.org)

Finally, secure aggregation should be evaluated at its output boundary. Kariyappa and colleagues explicitly discuss why aggregation alone is insufficient for provable privacy in their federated-learning setting. That result is not a truth-discovery benchmark, but it reinforces the need to specify output disclosure instead of equating encrypted computation with a complete privacy guarantee. (proceedings.mlr.press)

Data Analysis and Evidence

Label coverage is part of the denominator

Li and colleagues' dataset table reports **72 users, 360 answers and 26 questions** for Peer Grading, with **26 groundtruth labels**. (cse.buffalo.edu) For Duchenne Smile it reports **64 users, 17,729 answers and 2,134 questions**, with **159 groundtruth labels**. (cse.buffalo.edu) Its reported perturbation results are means over **100 independent trials**. (cse.buffalo.edu) Confidence intervals were not verified in this audit.

Calculated from those counts, labeled-question coverage is **100%** and approximately **7.45%**, respectively. These are coverage calculations, not accuracy estimates. They show why the total number of collected questions cannot automatically become the denominator of a labeled evaluation. No new error-rate effect size is inferred here. (cse.buffalo.edu)

Weight changes can reverse the answer (Hypothetical Example)

Consider three sources, four calibration objects with known labels, and one target object. Assign each source a weight equal to its fraction of correct calibration answers. This intentionally simple supervised calibration rule isolates the mechanism; it is not a reproduction of a published estimator.

Initially, the sources have calibration correctness vectors (1,1,1,1), (1,1,1,0), and (0,0,0,0). Their weights are 1, 0.75 and 0. The first two sources answer A on the target; the third answers B. Weighted support is therefore 1.75 for A and 0 for B.

Now consider one possible perturbation realization in which the correctness vectors become (0,0,1,0), (0,1,0,0), and (1,1,1,0). Keep the target answers unchanged. Re-estimated weights are 0.25, 0.25 and 0.75. Support becomes 0.5 for A and 0.75 for B: the selected answer reverses.

All numbers in this example are constructed arithmetic. The example establishes possibility, not frequency, expected loss or an attack. It motivates measuring **weight-rank stability** and **answer reversals** alongside task error. A method could have a similar average error while moving different objects across the decision

boundary.

Equal ε can describe unequal protection (Hypothetical Example)

Take a binary randomized response that reports the true bit with probability exp(ε)/(1 + exp(ε)). The ratio of output probabilities for alternative bits is exp(ε), establishing the single-bit pure local bound directly. At ε = 1, the flip probability is approximately 0.269.

If a contributor independently reports five answers at ε = 1 each, basic composition supplies a vector bound of ε = 5. To obtain a vector bound of ε = 1 through equal allocation under that same accounting, each answer receives ε = 0.2; the flip probability becomes approximately 0.450. This is a derived example using basic composition, not a recommended budget. (differentialprivacy.org)

The calculation explains why a privacy axis must state its unit. It also does not license multiplying marginal bounds when a mechanism shares private random state across reports. Such a design needs an argument for the joint transcript.

A matched experimental protocol

Report the following outputs for every stratum:

- **Categorical error:** Include the exact class weighting.
- **Numerical loss:** Report absolute and squared error in declared units.
- **Coverage:** Count answers issued and answers withheld.
- **Selective risk:** Measure error among issued answers.
- **Weight stability:** Record rank correlation and large rank changes.
- **Reversals:** Count decisions changed relative to the frozen comparator.
- **Uncertainty:** Give paired differences and interval assumptions.
- **Cost:** Report messages, bytes and runtime separately from accuracy.

For class imbalance, scikit-learn defines balanced accuracy as average recall across classes; use an explicit definition rather than an unexplained label. For uncertainty, SciPy documents paired bootstrap resampling through common indices. Object-level resampling is appropriate only under the declared independence assumptions; dependent sources or object blocks require a corresponding design. (scikit-learn.org) (docs.scipy.org)

Implications and Future Directions

The most useful future comparison would hold the information problem fixed while varying the privacy construction. It would use identical objects, candidate domains, contributor availability and reference labels, then separate experimental strata by observer. Combining incompatible threat models into one score would hide the central decision: which disclosure must be prevented.

Source reliability needs its own evaluation target. A system may conceal raw answers yet expose a ranking that a contributor regards as sensitive. L-PPTD’s explicit distinction between protecting sensory data and protecting reliability illustrates why the two should be named separately. If reliability is postprocessed

entirely from an LDP transcript, the transcript's guarantee is relevant; if raw information is accessed again, the accounting must reflect that access. (researchconnect.suny.edu) (docs.opendp.org)

Dependence requires both statistical and privacy analysis. Shared source material changes the evidential interpretation of agreement, while correlated data complicate the relationship between differential and inferential privacy. A defensible report should disclose the dependence assumptions rather than infer independence from separate contributor identifiers. (arxiv.org)

Reproducibility should be scoped to clearance. A public reader may be able to verify the protocol specification and aggregate outputs while a privileged evaluator checks reference labels. That division can be legitimate, but it must remain visible. Provenance representation and credential verification provide useful vocabulary for recording activities and authenticating statements; neither supplies the external truth judgment by itself. (www.w3.org)

The immediate research agenda is therefore narrower than choosing an overall winning method: populate evaluation cards from exact definitions, and run matched experiments only after the privacy unit and output loss have been aligned. Empirical auditing can test implementations, while a formal proof supplies the quantified guarantee under its assumptions. NIST's implementation-focused guidance makes this separation particularly relevant. (papers.neurips.cc) (csrc.nist.gov)

Frequently Asked Questions (FAQs)

What is privacy-preserving truth discovery?

It is estimation from conflicting contributions with a specified protection for inputs, participation, reliability or another declared object. The categorical formulation examined here uses inferred quality to weight answers. Different mechanisms protect different observations; the phrase alone does not establish an LDP guarantee. (www.kdd.org) (researchconnect.suny.edu)

Does ε identify the amount of information protected?

It quantifies a bound relative to declared alternatives. The input unit and composition determine the interpretation. One answer, a contributor vector and participation are not interchangeable neighboring relations. (nvlpubs.nist.gov) (docs.opendp.org)

Does secure aggregation provide differential privacy?

A secure aggregation protocol can conceal individual inputs while releasing an aggregate. Its confidentiality statement permits information learnable from that aggregate, so a differential privacy claim needs its own definition and analysis. (research.google)

Can a private result also be verifiable?

Yes, if the verification proposition is precisely stated. Authenticating a statement, verifying a particular computation and establishing factual truth remain separate tasks. W3C's credential model explicitly cautions against equating verifiability with truth. (www.w3.org)

Conclusion

Privacy-preserving truth discovery should be evaluated through a complete statement of the information boundary. The protected unit, observer, input domain, privacy definition and release history determine what a budget means. The comparator, truth labels, loss and uncertainty determine what a utility claim means. Without both sets of fields, an accuracy number cannot answer whether a method preserves the right information while concealing the right contribution.

The accessible evidence supports scoped findings rather than a universal ranking. The categorical study provides a concrete perturbation-based example. Cryptographic truth-discovery and secure aggregation work provide different confidentiality boundaries.

The strongest evaluation will also preserve unknowns: an unread theorem, unavailable reference label or undeclared composition rule should remain visible until evidence resolves it. Completing those fields is part of the research result.

The proposed card makes those limitations actionable. It requires matched experiments, separate source-weight and output tests, a visible abstention policy, and an account of what reviewers can inspect. The hypothetical calculations explain why both source weighting and budget units can change the practical conclusion. A defensible result is therefore more than a private-looking answer: it is an answer whose utility and disclosure claims can each be checked against a declared model.

Source links

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

1. https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-226.pdf
2. https://www.nist.gov/blogs/cybersecurity-insights/threat-models-differential-privacy
3. https://www.kdd.org/kdd2018/accepted-papers/view/an-efficient-two-layer-mechanism-for-privacy-preserving-truth-discovery
4. https://cse.buffalo.edu/~lusu/papers/KDD2018Yaliang.pdf
5. https://research.google/pubs/practical-secure-aggregation-for-federated-learning-on-user-held-data
6. https://www.w3.org/TR/vc-data-model-2.0
7. https://differentialprivacy.org/composition-basics
8. https://papers.neurips.cc/paper_files/paper/2022/hash/1add3bbdbc20c403a383482a665eb5a4-Abstract-Conference.html
9. https://csrc.nist.gov/pubs/sp/800/226/final
10. https://www.w3.org/TR/prov-overview
11. https://docs.opendp.org/en/stable/api/user-guide/framework-api/a-framework-to-understand-dp.html
12. https://www.rfc-editor.org/rfc/rfc6973
13. https://docs.opendp.org/en/stable/api/python/opendp.combinators.html
14. https://proceedings.mlr.press/v37/kairouz15.html
15. https://privacytools.seas.harvard.edu/differential-privacy?page=2
16. https://arxiv.org/abs/1603.01508
17. https://researchconnect.suny.edu/en/publications/a-lightweight-privacy-preserving-truth-discovery-framework-for-mo

18. https://www.cs.newpaltz.edu/~lik/publications/Xingting-Liu-IEEE-SJ-2023
19. https://scholarsmine.mst.edu/comsci_facwork/1393
20. https://data.mendeley.com/datasets/93kr7pb7dc/1
21. https://www.census.gov/about/policies/privacy/statistical_safeguards.html
22. https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.bootstrap.html
23. https://datatracker.ietf.org/doc/html/rfc9497
24. https://proceedings.mlr.press/v202/kariyappa23a.html
25. https://scikit-learn.org/stable/modules/generated/sklearn.metrics.balanced_accuracy_score.html

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.