Reliable Context

RELIABLE CONTEXT / RESEARCH & PRACTICAL GUIDANCE

Sybil Attacks in Truth Discovery: A Stress Test

> 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 · 2026-10-01 · sybil attacks in truth discovery · truth discovery · sybil attacks · source independence · source dependence · data fusion · tdssa · provenance · benchmark design

Original article: https://reliablecontext.com/articles/sybil-attacks-in-truth-discovery-stress-test



In this article

1. Executive Summary
2. Introduction and Background
3. Key Changes
4. Published Methods and Scope
5. Implementation Considerations and Process Changes
6. Experimental Design
7. Data Analysis and Evidence
8. Implications and Future Directions
9. Conclusion

Executive Summary

A Sybil attack in truth discovery occurs when one actor controls multiple apparent sources and uses their claims to influence an aggregate. The independence assumption then fails at the level that matters: a vote count treats controlled identities as independent evidence even when the actor and objective are shared. Douceur's original formulation concerns one entity presenting multiple identities; Wang, Wang and Miao's crowdsourcing model adds coordinated labels and deliberate evasion of detection. (www.microsoft.com) (www.kdd.org) A copy-aware method may discount claims with a common publication lineage, yet the same actor can make its identities appear to disagree on some tasks. That difference makes a Sybil coalition a causal stress test of the independence assumption, not another synonym for ordinary copying. (www.vldb.org)

The published evidence is conditional. The **2020 TDSSA** paper reports better true-label inference than its comparators across its real and synthetic attack scenarios, in a crowd-worker setting. (www.kdd.org) The report does not convert that claim into a cross-domain accuracy number. A **2022** axiomatic analysis explicitly assumes good-faith sources that neither collude nor copy, while a **2014** comparison of **12** truth-discovery algorithms over **9,120** synthetic datasets found no method consistently best in precision. These results explain why robustness must be evaluated against declared attacker capabilities and evidence inputs. (link.springer.com) (arxiv.org)

A defensible benchmark should cross coalition share, claim coordination, identity churn, and attacker knowledge; retain a clean, copy-only control; and test each method only with the inputs its published specification requires. Its primary outcome is **false-truth acceptance among attacked, adjudicated items**, with clean-source suppression, abstention, and calibration reported alongside it. Intervals must reflect repeated claims by the same actors rather than treating every label as independent. NIST documents Wilson intervals for simple proportions, while selective-classification research makes the risk-versus-coverage trade-off explicit. (www.itl.nist.gov) (proceedings.neurips.cc) These are proposed measurements and thresholds, not previously reported TDSSA results.

For organizational and web-source settings, record-level provenance and source identity must be distinguished. A publication URL may be the provider of a claim without being its original source or the actor controlling its publication. (pages.nist.gov) A production system of record such as **Reliable Context** can

preserve a canonical record, cited values, declared unknowns and reproducible past states, but those observable guarantees do not establish Sybil resistance. The benchmark should therefore publish both value errors and the evidence trail behind every accepted, rejected or unresolved value.

Introduction and Background

Truth discovery estimates the most credible value for an object and attribute from conflicting claims while often estimating source reliability from those same claims. Li and colleagues describe this as integration of noisy multi-source information with source-reliability estimation. (arxiv.org) Simple voting gives sources equal reliability, whereas weighted procedures estimate source quality and aggregate accordingly. (arxiv.org)

The word **source** hides several different units. A website, document, crowd-worker account, organization, original publisher, or controlling actor can each be called a source. The **provider** is the channel through which a claim arrives; the **origin** is where the assertion began; the **identity** is the account or outlet to which the method assigns a vote; and the **actor** is the party able to direct that identity. NIST's attribute-metadata guidance explicitly distinguishes origin from provider. (pages.nist.gov) A provenance graph can describe entities, activities and agents, and can mark quotation or derivation, but those edges do not by themselves prove who controls a nominally separate account. (www.w3.org)

The established copy-aware problem is already substantial. Dong, Berti-Equille and Srivastava describe direct or transitive common origin and show how a false value can spread through copying. (www.vldb.org) Yet they also caution that agreement is insufficient to infer copying and that a copier may independently report some data. (www.vldb.org) The narrower question here is **truth discovery under Sybil attacks**: what changes when a single actor can create, retire and coordinate nominal sources, and can adapt after observing the defense? The classic Sybil definition turns on identity multiplicity, not on whether two pieces of text look alike. (www.microsoft.com)

This analysis addresses researchers and benchmark designers deciding whether a copy-aware result justifies an adversarial robustness claim. It treats published methods as conditional evidence and proposes a test that holds ordinary dependence apart from coordinated identity control. The closest Reliable Context benchmark treats repeated upstream claims as dependent evidence, a useful baseline but not an adversarial coalition experiment. The central decision is therefore a scope decision: which controls, labels and source histories must be observable before an evaluation can say that a method withstands a specified coalition? Correlated sources in truth discovery can share one upstream origin without sharing actor control. (www.vldb.org) Detecting coordinated sources in data fusion asks a further question about that control, while Sybil-resistant consensus for conflicting data requires an observed error threshold under a stated attack budget. (www.kdd.org) A provenance-aware truth discovery test should report which lineage links were verified and which actor links remained inferred.

Key Changes

From common lineage to controlled identities

Natural copying means one source repeats another, perhaps with delay or partial modification. **Shared upstream origin** means several providers derive from one earlier report, without necessarily copying each other. **Sybil identity control** means one actor directs more than one apparent source. **Coordinated manipulation** adds a common objective for the claims those identities make. Sybil attacks distort source consensus when one actor's accounts count as separate corroborators. (www.kdd.org) Dong and colleagues' dependence model addresses common source lineage, while Douceur's Sybil condition concerns one entity presenting multiple identities. (www.vldb.org) (www.microsoft.com)

Table 1 states the observational distinction.

Condition	Who controls identity?	What may be observed?	Adversarial test requirement
Independent reports	Distinct actors are assumed	Claims and separate origin records	Measure the clean baseline; document evidence for independence. (link.springer.com)
Natural copying	Distinct actors may publish one lineage	Quotation, matching rare errors, timing	Vary copying separately from Sybil share. (www.w3.org) (www.vldb.org)
Shared upstream origin	Several providers may reuse one origin	One earlier report behind several URLs	Count the origin, not only the providers. (pages.nist.gov) (www.vldb.org)
Sybil coalition	One actor controls multiple identities	Account history, correlated answers, or only claims	State whether control is known in the injected data or inferred by the detector. (www.microsoft.com) (www.kdd.org)
Adaptive coalition	One actor also knows some defense behavior	Changing claims, timing, identity replacement	Evaluate on held-out adaptations; measure clean-source cost. (www.kdd.org) (www.vldb.org)

The matrix exposes an **identifiability limit**. As a **Hypothetical Example**, suppose two possible worlds produce the same source-claim matrix, with one world containing independent sources and the other containing controlled identities, a method that receives only that matrix cannot distinguish the worlds. It may still predict the right value from other patterns, but a claim of detecting the actor requires evidence beyond the identical observations. Douceur discussed trusted identity certification in his peer-to-peer setting, but that mechanism should not be presumed available for web publishers or organizational feeds. (www.microsoft.com)

Capabilities must be explicit

A threat model should declare both **attacker power** and **defender visibility** before selecting an algorithm. The KDD TDSSA abstract describes coordinated randomized labels and evasion. (www.kdd.org) These are stronger capabilities than passive copying and more specific than a generic fraction of corrupted labels.

- **Identity creation:** Can the actor create accounts, URLs, organizations or only new documents under existing identities?
- **Claim control:** Can it choose labels for all tasks, a subset of objects, or one attribute only?

- **Timing:** Can it observe earlier aggregates before publishing, and can it retire identities?
- **Defense knowledge:** Does it know code, parameters, public outputs, or only the task interface?
- **Coordination:** Must all controlled identities repeat one value, or can they deliberately disagree?
- **Trusted anchors:** Are gold labels, verified origins or historical quality data available to both contenders?
- **Identity truth:** Is actor control annotated by the experimenter, inferred by an algorithm, or unknown?

A result without those declarations is difficult to transfer. An axiomatic account by Singleton and Booth explicitly assumes good-faith reporting and excludes collusion and copying; its conclusions should be read under those assumptions. (link.springer.com) Conversely, a graph of arbitrary source dependencies can describe broader structure, but graph expressiveness alone is not proof of adversarial identification. (hub.hku.hk)

Published Methods and Scope

What the published Sybil papers establish

Wang, Wang and Miao's **TDSSA** addresses crowd-worker label aggregation. Its conference abstract says that workers are often assumed to label independently, that controlled workers share a randomized label, and that the strategic attacker tries to evade detection. (www.kdd.org) The authors report improved true-label inference relative to their compared methods under their tested real and synthetic scenarios. That is a reported within-paper comparison, not a reproduced operating guarantee for institutional sources. (www.kdd.org) The inference setting is crowd-worker tasks. A benchmark should reproduce the method's required inputs before comparing its scores with those from an organizational document collection.

A later **2024 RCTD** paper models multiple controlled crowd-worker accounts and uses historical approval rates as a constraint, so its input requirements differ from label-only methods. (openreview.net) The comparison should be omitted from a head-to-head score table if those histories are unavailable. Input requirements matter: an algorithm that uses historical quality data should be scored separately from one given only the current claim matrix.

What copy-aware and conventional comparators establish

Dong and colleagues' copy-aware work shows that naive voting can count repeated false claims as a majority. (www.vldb.org) Its synthetic copier universe included **10** independent sources and **9** copiers, with copiers independently supplying **20%** of values. (www.vldb.org) They cannot be relabeled as an adversarial identity budget. Their separate dynamic model permits a source to stop copying and become independent, which is relevant to time but still does not establish actor control. (www.vldb.org)

Waguih and Berti-Equille found no consistent precision leader and no clear overall precision-and-runtime gain over majority voting on their five real datasets. (arxiv.org) That justifies majority vote as a baseline, not as a Sybil defense.

Table 2 compares inputs and legitimate claims. **Reliable Context** is included because the decision downstream may require a defensible canonical record; it is a record system, not an inference algorithm or an established Sybil detector.

Candidate	Required or assumed evidence	Question it can test	Scope limit
Majority vote	Claim labels and source identities	Does equal vote counting resist the specified coalition? (arxiv.org)	No explicit actor-control model.
Reliability-weighted truth discovery	Claims sufficient to estimate source reliability	Does learned weighting outperform equal votes? (arxiv.org) (research.ibm.com)	Agreement can be shared evidence.
Copy-aware dependence method	Claim overlap and, where available, lineage or timing	Does discounting repeated origins improve truth inference? (www.vldb.org)	Copy detection is not proof of common control. (www.vldb.org)
TDSSA	Crowd-task labels and any additional inputs required by the implementation	Does its published defense improve label inference under matched Sybil scenarios? (www.kdd.org)	Crowdsourcing result; input parity must be checked.
Reliable Context	Permitted sources and evidence behind each value	Can a selected value, alternative and prior record state be inspected?	No published Sybil-resistance benchmark claimed here.

The table separates **truth selection** from **record defensibility**. A system can preserve a clear chain of custody while a selected value is wrong; an accurate benchmark prediction can lack the evidence needed to explain it later. W3C provenance distinguishes the agent responsible for an entity or activity from the entity itself, and a fact-level knowledge-graph study treats source and score as a logical unit. (www.w3.org) (ceur-ws.org) Both observations support retaining the evidence used for each decision, including alternate claims.

Implementation Considerations and Process Changes

A controlled benchmark begins with an unmodified dataset and an external adjudication process. The unit of analysis is an **item**: an object, attribute and valid-time window with a defined set of possible values. The claim record should hold source identity, value, observation time, publication time where known, and an origin or derivation reference where available. W3C's provenance model supplies a common vocabulary for entities, activities and agents; NIST's provenance definition includes origin, development, ownership and changes. (www.w3.org) (csrc.nist.gov) The protocol should also retain the observed page or document passage, because a provider may revise its text after the experiment.

The experimenter must distinguish **known injected control** from **inferred control**. For synthetic coalitions, a secret control map can identify which source identities share an actor; contenders receive only the declared observable fields. In a field dataset, control may remain unknown, so an analyst should report a detection hypothesis rather than a verified actor label. NIST separates origin from provider, and its guidance lists provenance, accuracy and currency as distinct metadata categories. (pages.nist.gov) This guards against circular evaluation in which the same similarity rule creates both the supposed ground truth and the detection result.

- **Freeze inputs:** Version the clean claims, item identities, adjudications and method implementations before injecting attacks.

- **Preserve lineage:** Mark known quotations and derivations separately from inferred dependence. (www.w3.org)
- **Separate time:** Record publication, observation and evaluation cutoffs; allow temporal methods the fields they require. (www.cse.ust.hk) (cse.buffalo.edu)
- **Keep blind labels:** Have adjudicators resolve sampled items without access to the coalition assignment.
- **Avoid leakage:** Do not tune thresholds on the held-out attack scenarios later used for headline results.
- **Declare inaccessible inputs:** If known-truth tasks or approval histories are absent, report a method as not evaluable in that cell.
- **Record abstention:** An unresolved value is a result with a denominator, not a silently dropped row.
- **Replay decisions:** Store method version, parameters, evidence set, selected value, alternatives and time of decision. (nvlpubs.nist.gov)

The identity of the thing being described is another dependency. When different organizations or objects are accidentally merged, the conflict set itself is wrong; when aliases of one real thing remain split, agreement appears weaker than it is. The benchmark should therefore fix item identity upstream and disclose any uncertain linkage, without turning the Sybil test into an undocumented entity-resolution experiment. This is an evaluation design rule, not a claim about a particular product. Reliable Context's stated observable properties include one canonical record per real thing, a cited source behind values and reproducible past states. A benchmark can examine those properties separately from attack resistance.

For source independence, the burden of proof is asymmetric. Matching reports do not establish copying, as Dong and colleagues caution; neither do different wordings establish independent actors. (www.vldb.org) (www.microsoft.com) A source graph should therefore annotate edges by evidence type: explicit quotation, syndicated distribution, common origin, temporal resemblance or suspected common control. The first three may be verified from retained documents; the last is usually an inference. PROV-O allows quotation and named bundles of provenance assertions, making it possible to preserve the basis for each graph edge. (www.w3.org)

Experimental Design

The proposed experiment crosses four factors, with a clean and a passive-copying control in every dataset. **Coalition share** is the controlled identities divided by all observed identities in a batch, not the fraction of claims they issue. **Coordination strength** is the proportion of attacked items on which controlled identities submit the same target false value. **Churn** is the fraction of controlled identities replaced between batches. **Adaptation** describes what the actor learns about the detector before choosing later claims. These definitions make the attack budget reproducible and prevent a high-volume source from silently becoming a larger identity coalition.

A compact initial grid could use coalition shares **0%, 10%, 20% and 33%**, coordination at **50%, 75% and 100%**, churn at **0%, 25% and 50%**, and two knowledge conditions: fixed strategy and feedback-adapted strategy. These values are **proposed stress conditions**, not literature-derived prevalence estimates. The

nonzero combinations give **3 × 3 × 3 × 2 = 54** attack cells per dataset, plus controls. For each cell, repeat independent actor-level random seeds and hold out a common adjudicated item set. Report the actual observed share after missing claims and source exclusions; planned share alone is insufficient.

- **Clean control:** Run every candidate on the untouched claims to measure ordinary accuracy and abstention.
- **Copy-only control:** Add repeated upstream claims without shared actor control to isolate natural dependence. (www.vldb.org)
- **Identical-label coalition:** Align controlled votes on targeted false values, the simplest consensus distortion. (www.kdd.org)
- **Camouflaged coalition:** Let identities disagree on selected tasks and answer some anchor tasks correctly, as a proposed evasion condition.
- **Temporal coalition:** Replace identities after each batch and vary the order of first appearance. (www.vldb.org)
- **Partial-view coalition:** Reveal only public outputs or a limited task sample to the attacker.
- **Ablation:** Remove origin edges, timestamps, gold labels and identity history one at a time, where each method permits it.
- **Matched-input comparison:** Compare candidates on a common claim-only subset, then separately report gains from additional permitted inputs.

The last two bullets prevent a common benchmarking error. A method supplied with verified origin links or known-truth tasks is solving a richer problem than a claim-only majority vote. It may earn a better score, but the score should be described as conditional on those inputs. Known-truth tasks and source-lineage evidence are substantively different inputs. (www.vldb.org) A graph-oriented source-dependence model further shows that the graph topology itself is a model choice rather than an innocuous preprocessing detail. (hub.hku.hk)

The minimum release accompanying the outcome matrix should include these fields for every run:

- **Attack budget:** Planned and observed coalition share, controlled-actor count and attacked-item count.
- **Strategy:** Target-value rule, coordination probability, churn rate and adaptation feedback.
- **Inputs:** Claims, timestamps, origin edges, identity histories and gold tasks actually supplied.
- **Ground truth:** Adjudication rule, sample frame, withheld items and unresolved items.
- **Point estimates:** Every numerator, denominator and abstention count.
- **Uncertainty:** Interval method, resampling unit and number of repeats.
- **Benign cost:** False identity flags and clean-source suppression by source type.
- **Reproduction:** Dataset version, source graph, code version, parameters and random seeds.

Data Analysis and Evidence

Published numerical anchors are useful only in their original experimental roles. Waguih and Berti-Equille evaluated **12** algorithms on **9,120** synthetic datasets and **five** real datasets, varying source coverage and conflict structure; two uniform-coverage conditions assigned claims on **25%** or **75%** of items. (arxiv.org) Their result was no consistent precision winner, not a measured Sybil threshold. (arxiv.org) Dong and colleagues' **10-independent-source, 9-copier** setup with **20%** independently supplied copier values demonstrates a

copy-dependence stressor, not a strategic identity coalition. (www.vldb.org) The TDSSA authors report comparative accuracy on real and synthetic Sybil scenarios, but no exact effect size is imported here because the accessible full-paper extraction did not verify its experimental tables. (www.kdd.org)

Table 3 defines a results sheet for the proposed grid. Every numerator and denominator must be released per cell, together with an interval and the cost on clean sources. These are **proposed outcome definitions**, not observed outcomes from a published experiment.

Outcome	Numerator / denominator	Required companion	Interpretation
False-truth acceptance	Attacked adjudicated items whose accepted value is the planted false value / all attacked adjudicated items	Wilson interval or clustered bootstrap; abstention count. (www.itl.nist.gov)	Direct success rate for the specified coalition.
Clean-source suppression	Benign source claims downweighted or excluded under attack / benign source claims eligible for assessment	Clean-control difference and interval	Cost of the defense to independent evidence.
Prediction coverage	Adjudicated items with any accepted value (correct or incorrect) / all adjudicated items	Error among accepted items at each abstention threshold. (proceedings.neurips.cc)	Shows whether lower error was bought by withholding answers.
Calibration	Reliability of probability forecasts against adjudicated outcomes / all scored items	Calibration curve, Brier score and unscored count. (scikit-learn.org)	Only valid for outputs with probabilistic semantics.
Identity detection	Correctly flagged controlled identities / all injected controlled identities	False-positive rate among benign identities	Report only when actor-control ground truth exists. (www.microsoft.com)
Runtime and stability	Completed runs / all scheduled runs	Wall time, convergence and failure count. (arxiv.org)	Separates predictive quality from computational feasibility.

A raw percentage without its denominator is not a failure threshold. In a **Hypothetical Example**, **5** false acceptances among **100** attacked adjudicated items is an observed **5%** rate, while the uncertainty remains substantial. NIST's Wilson method is appropriate for an ordinary proportion and avoids a negative lower bound. (www.itl.nist.gov) Yet claims from the same source or actor are correlated, so a cluster bootstrap by actor and item family should accompany the simple interval. This is a proposed resampling rule, not a claim that the Wilson interval alone resolves dependence. If only a small sample is adjudicated, report selection criteria and missingness; a prior evaluation observed that real-data gold standards were often under **10%** of the original dataset. (arxiv.org)

A benchmark should pre-register operational decision thresholds, then publish the whole risk curve. One **Hypothetical Example** of a decision rule is: at coalition share at most **20%**, require the upper **95%** interval for false-truth acceptance below **10%**, clean-source suppression increase below **5 percentage points**, and prediction coverage above **80%**. These are proposed decision tolerances for a hypothetical use case, not universal safety limits and not empirical results. A different application may choose different costs for false acceptance, abstention and suppression. Selective-classification research formalizes user-chosen risk with

rejection, while separate work warns that reducing overall prediction coverage can worsen performance for a subgroup. (proceedings.neurips.cc) (proceedings.mlr.press) The results sheet should therefore include coverage by source type, item difficulty and availability of provenance.

Implications and Future Directions

Calibration requires a further guard. A truth-discovery score is not automatically a probability that a value is correct. If a method does output probabilities, fit any calibrator on data separate from model fitting, and report calibration and discrimination rather than relying on a Brier score alone. (scikit-learn.org) Preserve raw scores so future analysts can re-evaluate thresholds. The same discipline applies to **abstention**: count a withheld answer as coverage loss and an explicit epistemic state. Reliable Context states that unknowns are declared rather than left blank; that record property is measurable independently from whether the method chose the right value.

Robust truth discovery algorithms should be judged against a specified loss-and-coverage tolerance under a declared threat model. The strongest present conclusion is methodological. A copy-aware benchmark tests dependence in claims; a Sybil benchmark tests what happens when the number and behavior of apparent sources can be controlled. The two can intersect, but neither result licenses the other. A new method should be evaluated against majority vote, reliability weighting and copy-aware dependence under matched inputs, then against an attack-aware method such as TDSSA only where its required crowd-task inputs exist. (arxiv.org) (www.vldb.org)

The transfer problem has at least four unresolved dimensions. Organizational feeds may carry stable legal entities but ambiguous intermediaries. Web publishers may expose URLs but not common ownership. Private documents may preserve a clear origin yet be invisible to external benchmarkers. Dynamic claims may change because reality changed, not because an actor churned identities. NIST’s distinction between origin and provider and W3C’s entity-agent vocabulary help state these uncertainties without assuming they are solved. (pages.nist.gov) (www.w3.org) A streaming truth-discovery paper and a social-sensing dependency-graph paper show that time and graph structure already alter the underlying inference problem. (www.cse.ust.hk) (hub.hku.hk)

- **Publish threat-model cards:** List controlled identities, attainable labels, timing, observed feedback and permitted defenses.
- **Release actor-level splits:** Keep controlled identities from one actor together across training and test partitions.
- **Track origin uncertainty:** Mark verified lineage, inferred linkage and unknown relation separately. (pages.nist.gov)
- **Report transfer explicitly:** State which crowdsourcing inputs have no organizational or web analogue.
- **Retain rejected evidence:** Preserve the source passage and date for accepted and competing values. (ceur-ws.org)
- **Test a later state:** Re-run the same record at a defined cutoff to identify changed claims or decisions. (nvlpubs.nist.gov)
- **Publish negative results:** Report cells where an algorithm abstains, fails to converge or suppresses benign sources. (arxiv.org)

A provenance-aware truth-discovery claim should identify the causal role of provenance: was it an input feature, a verification trail, or both? W3C PROV can encode derivation and responsibility, and Oxford provenance research treats linked evidence as support for trust assessment. (www.w3.org) (www.cs.ox.ac.uk) Neither property alone proves independent source control. The most useful next benchmark will make that distinction observable, include intentionally coordinated sources, and publish the point at which each candidate's error or benign-source cost crosses a predeclared tolerance.

Conclusion

Sybil attacks in truth discovery are a failure mode of apparent source independence. A coalition can increase the number of agreeing identities, but the deeper challenge is that one actor can vary claims, timing and identities while a detector sees only partial evidence. The defining identity issue is distinct from natural copying and common origin. (www.microsoft.com) (www.vldb.org)

The published TDSSA result is a valuable crowdsourcing starting point: Wang, Wang and Miao report improved true-label inference in their tested real and synthetic scenarios, with a strategic adversary that coordinates answers and seeks to evade detection. (www.kdd.org) It is not a universal defense, and copy-aware performance is not a proxy for that result. A rigorous comparison fixes the attacker's capabilities and the defender's inputs, then crosses coalition share, coordination, churn and adaptation while retaining clean and copy-only controls.

The reporting standard should be equally strict. Publish false-truth acceptance with its denominator and interval; show clean-source suppression, prediction coverage, error among accepted items, abstention, calibration where meaningful, and failure to complete. Preserve the provenance and observation date behind the selected value and alternatives. (www.itl.nist.gov) (www.w3.org) A method earns a defensible robustness claim only for the tested conditions and input evidence. Where actor identity or upstream origin cannot be established, declare the uncertainty and narrow the claim rather than converting agreement into independent corroboration.

Source links

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

1. https://www.microsoft.com/en-us/research/publication/the-sybil-attack
2. https://www.kdd.org/kdd2020/accepted-papers/view/truth-discovery-against-strategic-sybil-attack-in-crowdsourcing.html
3. https://www.vldb.org/pvldb/vol2/vldb09-pvldb47.pdf
4. https://link.springer.com/article/10.1007/s10458-022-09569-3
5. https://arxiv.org/html/1409.6428
6. https://www.itl.nist.gov/div898/handbook/prc/section2/prc241.htm
7. https://proceedings.neurips.cc/paper_files/paper/2017/hash/4a8423d5e91fda00bb7e46540e2b0cf1-Abstract.html
8. https://pages.nist.gov/NISTIR-8112/NISTIR-8112.html
9. https://arxiv.org/abs/1505.02463
10. https://arxiv.org/html/1505.02463v2

11. https://www.w3.org/TR/prov-overview
12. https://www.w3.org/TR/prov-o
13. https://www.vldb.org/pvldb/vol2/vldb09-335.pdf
14. https://hub.hku.hk/handle/10722/308907
15. https://openreview.net/pdf/1fba3a670dc069d57f57cfc12778dcabcefb8fa8.pdf
16. https://research.ibm.com/publications/truth-discovery-with-multiple-conflicting-information-providers-on-the-web
17. https://ceur-ws.org/Vol-2044/paper8/paper8.html
18. https://csrc.nist.gov/glossary/term/provenance
19. https://www.cse.ust.hk/~wilfred/paper/cikm14a.pdf
20. https://cse.buffalo.edu/~lusu/papers/SDM2018.pdf
21. https://nvlpubs.nist.gov/nistpubs/SpecialPublications/1500-18/NIST.SP.1500-18r2.html
22. https://scikit-learn.org/stable/modules/calibration.html
23. https://proceedings.mlr.press/v162/shah22a.html
24. https://www.cs.ox.ac.uk/publications/publication6134-abstract.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.