Reliable Context

RELIABLE CONTEXT / RESEARCH & PRACTICAL GUIDANCE

Federated Query Completeness in Knowledge Graphs: A Certificate

> 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 · 2026-11-02 · federated query completeness · knowledge graph completeness · sparql federated query · completeness statements · open world assumption · query completeness certificate · truth discovery · rdf provenance

Original article: https://reliablecontext.com/articles/federated-query-completeness-knowledge-graphs



In this article

1. Executive Summary
2. Introduction and Background
3. Open-World Answers, Certain Answers and Local Completeness
4. Key Changes: From Graph-Wide Scores to Query-Level Proof Obligations
5. Worked Examples: Single-Source, Federated and ASK
6. A Query-Level Certificate
7. Counterexamples and an Adversarial Fixture
8. Formal Completeness versus Estimated Recall
9. Data Analysis and Evidence
10. Implications and Future Directions
11. Frequently Asked Questions (FAQs)
12. Conclusion

Executive Summary

A federated SELECT or ASK result can be called **complete** only when declarations about the participating sources entail completeness for the exact query, the exact population and the exact time. Without such an entailment the honest verdict is **incomplete** or **unknown**. The most developed theory for this decision is the completeness-statement framework of Darari, Nutt, Pirrò and Razniewski, published at the International Semantic Web Conference in 2013 and extended in ACM Transactions on the Web in 2018 (Darari et al. 2018). A completeness statement says that a source holds all triples matching a pattern under a condition. A query is complete when it returns the same answers over the available graph as over a hypothetical ideal graph (www.inf.unibz.it). Entailment is decided by evaluating the statements' CONSTRUCT queries over a frozen copy of the query pattern, a check the authors show is NP-complete in combined complexity (as this report reads it, the check runs over the statements and the frozen query pattern, not over the data). For federations, a complete query can be rewritten so that each triple pattern is answered by a single source.

The guarantee has a hard boundary. The 2013 formalism considers only sources that may miss information but contain no wrong information (www.inf.unibz.it). Once two sources disagree, or one carries an erroneous assertion, the premise fails and a completeness certificate says nothing about soundness. The truth discovery literature starts from exactly that situation, where collected data conflict (arxiv.org), and large-scale fusion work reports that source errors exist alongside extraction errors (lunadong.com). Tanon et al. (2018) address a different problem: they use cardinality information to score mined rules, not to certify a query (www.ijcai.org), and their completeness metadata was derived (10 million statements for WikidataPeople) rather than independently curated (www.ijcai.org).

Federation adds operational failure modes that declarations cannot see. Under SERVICE SILENT, a failed call becomes a single solution with no bindings (SPARQL 1.1 Federated Query); endpoints often cap result sizes; and in one monitoring study 233 of 427 public endpoints returned no results to a bulk test query (link.springer.com). Recall is a ratio and completeness is binary, and most statistical estimators are described

as proofs of concept. For an explicitly bounded reference set, 34 confirmed returned answers out of 40 confirmed eligible ones give an answer recall of 85% (95% Wilson interval 71% to 93%, calculated here), which is never a completeness claim.

This report therefore proposes a falsifiable, query-level certificate: source identity and version, declaration scope, query and mapping version, entailment result, evidence for each declaration, expiry and reviewer. Standards status as of November 2026: SPARQL 1.1 Federated Query is a W3C Recommendation dated 21 March 2013 (www.w3.org), and the SPARQL 1.2 Federated Query draft is a Working Draft (www.w3.org).

Introduction and Background

This report examines one decision: whether a particular SELECT or ASK result computed over heterogeneous sources may be represented as complete, incomplete or unknown for a named population and time. It is about **query-answer completeness**, not about a product architecture and not about a claim that any graph is globally complete.

Seven terms are fixed before anything else:

- **Query:** the exact text, including its mapping to each source's vocabulary.
- **Answer unit:** the row (for SELECT) or the Boolean (for ASK).
- **Eligible population:** the set of real-world entities the question is about, for example the datasets published by one named laboratory.
- **Source snapshot:** the identified version of each source at a retrieval time.
- **Time boundary:** the date up to which the answer is claimed to hold.
- **Complete:** no eligible answer is omitted.
- **Sound:** every returned value is true.

Motro separated these two properties for databases in 1989: validity excludes all false information, and completeness includes all true information (www.scilit.net). Levy's 1996 work framed the query-level question directly, asking whether the answer to a specific query is complete even when the database is incomplete (www.vldb.org). Razniewski and Nutt (2011) carried the question into practice with a decentralised school database that is notoriously incomplete yet queried for official statistics (Razniewski and Nutt 2011), arguing that it is crucial to guarantee completeness of query answers (Razniewski and Nutt 2011).

The question is not new, but it is under-served. Issa et al.'s systematic literature review unifies terminology across 56 articles and identifies seven types of completeness (cris.maastrichtuniversity.nl), which is taxonomy rather than a tested certificate. Razniewski, Arnaout, Ghosh and Suchanek's survey (arXiv 2305.05403) covers logical foundations, statistical estimation and negation (arxiv.org). Note that this arXiv survey is a separate work from the Issa et al. review. The survey's authors argue that the first problem is to know whether and where a knowledge base is incomplete (arxiv.org). Query-level entailment is the formal answer to that problem; estimation is the pragmatic one.

A note on perspective. Reliable Context is a system of record built around properties that a certificate needs to be able to point to: a verbatim, cited source and an observation date behind every value, and unknowns declared as not found, blocked or not applicable with a reason rather than left as an empty cell. This report

makes no claim that any product computes completeness entailment, and its evidence comes only from third-party papers and standards. Readers who need a vector database, a data catalog or a master data management suite for other purposes should use one; none of those is a substitute for a declaration about a source. Further information is at reliablecontext.com.

Open-World Answers, Certain Answers and Local Completeness

Why an empty result proves nothing by default

RDF graphs are sets of triples, and the RDF semantics offers no way to assert that a container has only a fixed number of members (RDF 1.1 Semantics). The OWL 2 Primer states the consequence plainly: a statement that is missing may simply be missing but possibly true, following the open-world assumption (www.w3.org). SPARQL, by contrast, evaluates against a dataset (SPARQL 1.1 Query Language) by matching the active graph. An ASK query tests whether a pattern has a solution (SPARQL 1.1 Query Language), and FILTER NOT EXISTS tests whether a pattern does not match the dataset (SPARQL 1.1 Query Language). A `false` from ASK is therefore a statement about the queried dataset, and a negative conclusion about the world requires something more. The Wikidata completeness tool paper makes the same point: SPARQL computes answers under the closed-world assumption even though those answers may be incorrect under the open-world assumption (CEUR-WS Vol-1666 paper).

Two extreme stances exist. Under the closed-world assumption, absence means falsity; under the open-world assumption, absence means unknown. Partial-closed-world assumptions sit between them. The partial completeness assumption used for rule mining (if a source knows some value of a property for a subject, it knows all) (Galárraga et al. 2013) is one instance of these intermediate stances. The AMIE authors call it still reasonable for knowledge bases drawn from a single source (Galárraga et al. 2013), and note that in a knowledge base an absent statement is not necessarily false but just unknown (Galárraga et al. 2013).

Completeness statements and the ideal graph

The Darari et al. framework models an **incomplete data source** as a pair of an available graph and an ideal graph, with the available graph a subset of the ideal one. The ideal graph is a possible conceptualization of the world, and a completeness statement `Compl(P1 | P2)` says that the source contains all triples of the pattern P1 that satisfy the condition P2 (Darari et al. 2018). Because entailment quantifies over all ideal graphs consistent with the statements, reasoning about completeness does not depend on any specific ideal graph (Darari et al. 2018). The authors also describe their position as the middle ground between the classic open and closed world views (Darari et al. 2018).

A small hypothetical graph

(Hypothetical Example) Laboratory L publishes four datasets, d1 to d4. Source A, an institutional repository, lists `d1..d4 publishedBy L`. Source B, a license registry, lists a license for each: CC-BY for d1, d2 and d4, CC0 for d3. Neither source is known to be complete until it says so. Suppose A declares `Compl({?d publishedBy L} | {})` and B declares `Compl({?d license ?l} | {?d publishedBy L})`, both dated

2026-09-30. If a fifth dataset d5 existed in the world but in neither source, a query that returned four rows would be correct but incomplete, unless the declarations had ruled d5 out. This graph is used throughout the report; it is invented for illustration and corresponds to no real repository.

Key Changes: From Graph-Wide Scores to Query-Level Proof Obligations

Change 1: the unit of assessment becomes the query

Graph-wide quality scores answer a question nobody asks of a single answer. The query-level view is that a query is complete whenever all answers we would obtain over the ideal graph are also obtained over the available graph (www.inf.unibz.it). The same distinction between a Boolean completeness and a quantitative recall is drawn in the recall section below.

Change 2: closed regions replace a global closed-world switch

Statements close only the portions of the data they cover, and queries about uncovered data remain open. This is the property that makes a certificate possible: the proof obligation is local to the triple patterns of the query.

Change 3: federation is handled by flattening and single-source rewriting

For non-federated queries over a federation, completeness is decided on the union of all sources' statements (Darari et al. 2018). For the conjunctive federated fragment (Darari et al. 2018), a complete query can be rewritten so that each triple pattern is answered by a single source whose join equals the original result (Darari et al. 2018).

Change 4: time becomes part of the statement

The 2018 extension adds timestamped statements and a guaranteed completeness date, and deciding whether that date is at least a given value stays in NP (www.inf.unibz.it). The survey makes the reason explicit: areas formerly complete might become incomplete later on (arxiv.org).

Federated query versus data integration

A federated query evaluates SERVICE clauses at query time, extending SPARQL to queries that merge data distributed across the Web (www.w3.org). A materialized integration copies data into one graph, so a declaration must then be about the integrated snapshot rather than about the origin source. In both cases the certificate attaches to a source identity and version; the difference is which identity is named. Levy's motivation, that systems over multiple heterogeneous sources often meet incomplete sources (www.vldb.org), applies to both.

What the single-source theory covers

The 2018 paper covers basic queries, DISTINCT queries, well-designed OPTIONAL queries and RDFS (Darari et al. 2018). FILTER is only partly handled and its full treatment is left for future work, negation is future work, and the authors warn that aggregates over incomplete data may produce incorrect results (Darari et al. 2018). This report does not generalize beyond those constructs.

Worked Examples: Single-Source, Federated and ASK

Single source

(Hypothetical Example) Query Q0: `SELECT ?d WHERE { ?d publishedBy L }` against A. The frozen pattern contains one triple, `d0 publishedBy L`. A's statement yields exactly that triple, so the pattern is returned unchanged and completeness is entailed. The authors' own running example has the same shape: LinkedMDB can guarantee the completeness of the answer because it holds the relevant statements, whereas DBpedia is complete for a director's movies but nothing can be said about the same person's participation as an actor (www.inf.unibz.it).

Federated join

(Hypothetical Example) Query Q1: `SELECT ?d ?l WHERE { SERVICE <A> { ?d publishedBy L } SERVICE <B> { ?d license ?l } }`. A alone does not cover the license pattern, so Q1 is not entailed at A. With B's statement added, the license triple's condition (`?d publishedBy L`) is satisfied by the first frozen triple, and the transfer operator, which moves information from the ideal graph to the available graph, returns both frozen triples, which is the entailment test of evaluating the statements' CONSTRUCT queries over the frozen pattern (www.inf.unibz.it). Completeness is entailed under the declarations as written, and each triple pattern has a single covering source, so the four returned rows could be certified for the snapshot date once the evidence for each declaration has been reviewed and identifier agreement between A and B has been verified; until then the certificate verdict is conditional, and unknown if either check fails. The entailment holds only if both sources use compatible identifiers for the dataset entities, which is the global-schema assumption and is outside the theory when sources have different structures (www.inf.unibz.it). *Table 1* below lays out the obligations.

Triple pattern	Eligible entities	Source	Completeness statement	Version / date	Mapping	Entailment	Unresolved gap
`?d publishedBy L`	Datasets published by L	A (repository)	`Compl({?d publishedBy L} | {})`	snapshot A-2026-09-30	identity mapping, v1	Covered	Evidence for the statement not yet reviewed
`?d license ?l`	Datasets in the first row	B (registry)	`Compl({?d license ?l} | {?d publishedBy L})`	snapshot B-2026-09-30	`license` v1	Covered, given the first row	Identifier agreement between A and B unverified

Triple pattern	Eligible entities	Source	Completeness statement	Version / date	Mapping	Entailment	Unresolved gap
Join on `?d`	All of the above	A and B	none (a join is not a declaration)	n/a	IRI equality	Entailed only under the shared-schema assumption	Sameness of entities is assumed, not proved
`?d reviewedBy ?r` (added for the ASK below)	Datasets in the first row	none	none	n/a	none	Not covered	No source declares reviewers

Table 1 shows that completeness is a property of the whole obligation set, not of any source. The last row is the useful one: a pattern that no declaration covers makes the entire query unprovable, and the certificate must say which pattern failed.

ASK and the unsupported negative

(Hypothetical Example) Query Q2: `ASK { ?d publishedBy L . ?d license GPL }` returns `false`. Because the pattern is covered in the same way as Q1, the available and ideal answer sets coincide, so the empty result extends to the ideal graph: no dataset published by L carries a GPL license, for that snapshot. This step is derived here from the definition of query completeness, not stated as a separate theorem in the papers. Query Q3: `ASK { ?d publishedBy L . ?d reviewedBy ?r }` also returns `false`, but the reviewer pattern is uncovered, so the result must be labeled unknown. The two Booleans are identical on the wire and opposite in what they license. Arnaout et al. make the same distinction at the level of statements: a Wikidata SPARQL query for "actors without Oscars" returns plausible results that have no formal foundation under the open-world assumption (arxiv.org).

A Query-Level Certificate

The certificate is the artifact that a reviewer can accept or refute. Each field below exists because a failure case in the next section needs it.

- **Query and mapping version:** the exact query text, a hash of it, and the version of every predicate mapping applied to each source.
- **Answer unit and population:** SELECT row or ASK Boolean, plus the written definition of the eligible population.
- **Source identity and version:** endpoint or dataset identifier, snapshot identifier, retrieval time, and response status (answered, truncated, blocked, timed out).
- **Declaration scope:** pattern, condition, source index, issue date and the time boundary the declaration claims.
- **Evidence for the declaration:** who issued it and how. The literature distinguishes domain experts, crowd workers, text extraction and bots (ceur-ws.org).
- **Entailment result:** entailed or not entailed, with the first uncovered frozen triple when not entailed.
- **Soundness stance:** assumed (the 2013 premise) or checked, with the check named.

- **Expiry and reviewer:** under the policy proposed in this report (not part of the cited theory), the date after which the verdict reverts to unknown, and the named reviewer who accepted the evidence.
- **Verdict:** complete, incomplete or unknown, stated for the named population and time.

Provenance vocabularies supply part of this: PROV-O defines attribution as the ascribing of an entity to an agent (www.w3.org), generation time as when an entity was completely created and available for use (www.w3.org), and derivation as a transformation of one entity into another. VoID can carry dataset metadata such as the total number of triples. These terms record where a declaration came from; they do not make it true.

Source withdrawal as a test

Re-run the entailment with each source's statements removed. If withdrawing source A leaves the first pattern uncovered, then A is a necessary witness and its version and expiry are critical fields. A source whose removal changes nothing is redundant for this query, which is useful when choosing which declaration to audit first.

The verdict table

Table 2 below contrasts four answer states, plus a fifth that the 2013 premise excludes, using the vocabulary of the certificate.

State	Returned vs eligible set	Sound?	Complete?	Certificate verdict	What may be said
Complete answer	Equal	Yes	Yes (entailed)	Complete	"These are all eligible datasets as of the snapshot date."
Incomplete answer	Returned set is missing known eligible items	Yes	No (known omission)	Incomplete	"At least the omitted item exists; the list is partial."
Correct but incomplete	Every returned item is true; omissions possible but not shown	Yes	Not entailed	Unknown	"These are true; others may exist."
Unsupported empty answer	Empty, but pattern uncovered	Vacuously yes	Not entailed	Unknown	"Nothing found in the queried sources."
Complete by declaration, unsound	Includes a wrong item	No	Fails set equality	Not certifiable	"Source assertions conflict; resolve before certifying."

Table 2 makes the central point visible: "correct" and "empty" are not synonyms for "complete". Only the first row supports a negative conclusion, and the fifth row is where the framework's own assumption gives out.

Counterexamples and an Adversarial Fixture

Where the guarantee stops: erroneous and conflicting assertions

The 2013 paper states that it considers only sources that may miss information but do not contain wrong information (www.inf.unibz.it). The 2018 paper formalizes this as the available graph being a subset of the ideal graph, so the information in it is assumed correct (www.inf.unibz.it), and it notes that the problem of query soundness might also arise. The survey makes the same simplifying assumption, that the knowledge base is correct (arxiv.org).

Consider a derivation. If source B asserts a license for d3 that is not in the ideal graph, then B is not a subset of the ideal graph, the premise fails, and query answers over B can contain a tuple absent from the ideal answer. Equality of answer sets then fails even if every true tuple is present. A statement that B is complete for licenses does not repair this, because completeness bounds what is missing, not what is wrong.

The truth discovery literature begins at this point. Li et al.'s survey states that collected data may conflict due to errors, missing records and other causes (arxiv.org), that truth discovery integrates multi-source noisy information (arxiv.org), and that source reliability is usually unknown a priori (arxiv.org). Copying can spread false values and defeat majority voting (lunadong.com). In Knowledge Vault, an error mode is an erroneous statement on a spammy web site (lunadong.com). In the Dong et al. fusion study, a sample of 25 false triples found that only 4% were actually provided by the web sources, with extraction errors accounting for the rest (arxiv.org); the sample is small, and the lesson is that error attribution is itself uncertain. Local closed-world labeling is imperfect in the same data: half of 20 sampled false positives were not errors but artifacts of the assumption (arxiv.org). Knowledge Vault avoids the pure closed-world assumption because it would be rather dangerous when the reference base is very incomplete (lunadong.com), and its local closed-world labels can be wrong in the other direction because Freebase can contain errors (lunadong.com).

Tanon et al. (2018) sit on the other side of this boundary. Their contribution is a set of completeness-aware scoring functions for relational association rules (www.ijcai.org), motivated by the observation that no systematic way of using completeness or cardinality metadata in rule learning existed (www.ijcai.org). Their baseline, PCA confidence, guesses negative facts by assuming that data is usually added in batches (www.ijcai.org) and may give high scores to rules about very incomplete relations (www.ijcai.org); their completeness confidence generalizes standard and PCA confidence (www.ijcai.org). It is evaluated by the Pearson correlation between each ranking measure and rule quality (www.ijcai.org), on datasets such as WikidataPeople with 2.4 million facts over 9 predicates (www.ijcai.org). Their cardinality statements were derived for WikidataPeople, 10 million of them (www.ijcai.org), and for the synthetic LUBM data by assuming the PCA on the whole dataset (www.ijcai.org). The ideal graph is approximated using solid rules from the family domain (www.ijcai.org). The abstract claims remarkably higher accuracy than the state of the art in uncovering missing facts (www.ijcai.org). This is a ranking result on a rule-learning task. It is not a guarantee about any query result, and its assumptions about metadata differ from the entailment setting.

The adversarial fixture

Table 3 below (Hypothetical Example) fixes six cases on the hypothetical graph, with the expected certificate outcome. A certificate procedure that cannot reproduce these outcomes should not be trusted.

Case	Setup	Observed output	Expected outcome	Reason
F1 Silent omission	B drops d3's license; its declaration stays	3 rows	Certificate issued, then contradicted by audit; verdict revoked	The theory trusts declarations; only a cross-check (count, sample) can catch a false one
F2 Disagreement	A second registry C says d2 is CC0, B says CC-BY	5 rows, d2 twice	Not certifiable; soundness unresolved	Premise "no wrong facts" fails; truth discovery needed
F3 Mapping change	B renames `license` into two predicates	0 rows from B	Unknown until entailment is re-run on mapping v2	Declaration was about the old predicate
F4 Blocked endpoint	B returns an error; Q1 run with and without SILENT	Without SILENT the query fails (www.w3.org); with it, 4 rows, `?l` unbound	Unknown; response status recorded	A failed call counts as one solution with no bindings
F5 Expired declaration	B's statement valid to 2026-09-30; query run 2026-11-02	4 rows	Unknown after expiry	Expiry is this report's proposed policy, layered on the timestamped statements of the 2018 extension
F6 Truncated result	Endpoint caps rows silently	Partial rows	Unknown unless count cross-check passes	Caps are documented behavior (www.w3.org)

Table 3 shows that only F3 to F6 are detectable from metadata the certificate already holds. F1 and F2 need external evidence, which is why the "evidence for the declaration" and "soundness stance" fields exist. The failure in F4 deserves emphasis: the SERVICE SILENT form ignores errors encountered while accessing a remote endpoint (www.w3.org), so a naive consumer sees four plausible rows with a blank license, which reads as "no license" rather than "endpoint unavailable".

Formal Completeness versus Estimated Recall

Recall in information retrieval is the size of the relevant set that was returned divided by the size of the relevant set (www.cl.cam.ac.uk). For knowledge bases the survey defines recall as a real-valued ratio and completeness as the binary case in which recall equals 1 (arxiv.org). Cardinality assertions can be used to infer recall for a given selection condition (arxiv.org). Estimators exist for the case where no assertion is available, and their limits are documented.

Bounded answer recall

(Hypothetical Example) The calculation has three steps.

- **Fix the reference set:** a bounded list that reviewers can enumerate.
- **Confirm membership:** each reference item and each returned item is checked as eligible.
- **Divide:** answer recall = confirmed eligible returned answers / confirmed eligible reference answers.

Suppose reviewers confirm 40 eligible datasets in an explicitly bounded reference set (all datasets listed on L's own publication page for 2025) and the federated query returns 34 of them. Answer recall is confirmed eligible returned answers divided by confirmed eligible reference answers, 34 / 40 = 85%. If the 40 were a random sample of a larger population, the 95% Wilson score interval is about 71% to 93% (calculated here), and with 20 reference items and 17 hits the same 85% would span roughly 64% to 95%. In neither case is this a completeness statement: it does not license a negative, and a reference set that is itself incomplete biases it upward.

What estimators can and cannot say

Galarraga et al. report that their learned oracle achieves an F1 above 90% for more than half of the relations tested (arxiv.org), but for hasChild and marriedTo, predicting completeness is close to impossible even for a human (arxiv.org). Luggen et al. warn that a high number of singletons might result in significant overestimation by the Chao92 estimator (arxiv.org). The survey cautions that most estimators are proofs of concept (arxiv.org) and that statistical estimators should be used last because they are least reliable (arxiv.org). A high estimated recall is therefore evidence worth recording in the certificate, never a substitute for entailment.

Negatives

Absence is not a negative statement. All popular knowledge bases capture virtually only positive statements (arxiv.org). In a crowd assessment of inferred negations, 63% were correct, 31% incorrect and 6% ambiguous (arxiv.org). Only a covering completeness declaration turns an empty answer into an evidenced negative.

Data Analysis and Evidence

The quantitative record is mixed in vintage and method, and the sources are cited with their dates.

- **Cost and speed of entailment (2018):** the authors generated about 485,000 unique completeness statements; because they were generated from the queries, all queries are guaranteed to be complete (www.inf.unibz.it), so the evaluation measures reasoning speed, not real-world coverage.
- **Incompleteness magnitudes (2013 to 2018):** in Freebase, 71% of people had no known place of birth and 75% no known nationality (data of 2013) (lunadong.com); only 0.5% of football players in Wikidata had a highly complete profile (www.inf.unibz.it); DBpedia held 160 of 199 physics Nobel laureates in one example (aclanthology.org); in Wikidata, 2.7% of humans had an assertion about their father (www.ijcai.org).
- **Endpoint behavior (2013 to 2016):** 233 of 427 endpoints (54.6%) returned no results to a bulk test (link.springer.com), 24.3% were always down (link.springer.com), and 32.2% had very high reliability (link.springer.com). At least 68 endpoints implemented fixed result-size thresholds (link.springer.com), and clients may receive silently thresholded results when more than 10,000 are requested (link.springer.com). SPARQLES has monitored public endpoints since October 2013 (www.semantic-web-journal.net) with hourly availability tests (www.semantic-web-journal.net). No current availability study was retrieved for this report.

- **Federation engines (published online 2015, journal issue 2016):** in a controlled six-engine comparison, only FedX always retrieved complete results (www.semantic-web-journal.net); because other causes were controlled, the sole possible reason for missing results was the join implementation (www.semantic-web-journal.net); on a sliced setup one engine selected 30 sources instead of 5 (www.semantic-web-journal.net); and FedX's completeness claim holds only while its cache is up to date (www.semantic-web-journal.net). The result-size limit was set to 100,000 so that it could not cause truncation (www.semantic-web-journal.net), meaning live caps were not tested.
- **Public endpoint limits (as of the retrieved manual):** the Wikidata Query Service limits query runtime to 60 seconds (www.mediawiki.org) and allows 5 parallel queries per IP (www.mediawiki.org).
- **Declarations in the wild:** COOL-WD was prepopulated with over 10,000 statements (ceur-ws.org), about 7,600 of them Wikidata no-value statements (ceur-ws.org); crowdsourced statements cost about 10 cents each (ceur-ws.org), and about 84,000 extracted phrases had a precision of 94% (ceur-ws.org). Wikidata property constraints are hints, not firm restrictions (www.wikidata.org), and the no-value mechanism covers only the absence of a property for one item (www.wikidata.org).
- **Fusion error rates:** in a study of 1.6 billion extracted triples, 44% of predicates had accuracy below 0.3 while 13% were above 0.7 (arxiv.org).

Read together, the declarations that exist are small in number relative to the sources they would need to cover, and the evaluation that shows entailment is cheap used statements synthesized from the queries. Both facts bound what can be concluded: the reasoning is tractable, the supply of trustworthy declarations is the constraint.

Implications and Future Directions

Obtaining truthful declarations. The framework rests on the assumption that a domain expert has the necessary background knowledge, and the authors concede that incorrect completeness annotations may occur (www.inf.unibz.it). The survey adds that completeness metadata produced by language models itself comes with huge credibility issues (arxiv.org). A practical consequence is that each declaration needs an issuer, an evidence type and an expiry, which is what the certificate records.

Non-RDF and federated semantics. As noted in the federated join example, the federated results are stated for a conjunctive fragment, assume a global schema and leave rule-based mappings beyond the scope of the work. Real federations also meet endpoints whose ontologies differ from the home endpoint's (www.mediawiki.org). SPARQL 1.1 remains the Recommendation, dated 21 March 2013 (www.w3.org), and the 1.2 federated query document is a Working Draft (www.w3.org). The 1.2 query language page states it may be updated, replaced or obsoleted at any time (www.w3.org). Any certificate must therefore name the specification version it assumes. Relational and document sources need their own mapping and a statement of what the mapping preserves; no tested method for this was found in the sources reviewed.

Evolving schemas. A mapping change invalidates a declaration made against the old predicate (case F3). The survey's expectation that complete areas may become incomplete (arxiv.org) and the 2018 time extension (www.inf.unibz.it) suggest keeping declarations version-scoped, but no source reviewed here evaluates declaration lifecycle at scale.

Source errors. Combining entailment with truth discovery is open. A certificate can honestly record soundness as assumed, as the 2013 premise does, or checked by a named method, but there is no theorem that composes the two for the SPARQL constructs discussed here. Unsupported constructs (negation, aggregation, FILTER) remain future work.

Property exposure. A system of record is useful in this program when it exposes, per value, the source, the observation date and a reason for every unknown. Those are inputs to the certificate; they are not the entailment itself.

Frequently Asked Questions (FAQs)

Can a federated query be complete? Yes, under stated conditions. Completeness is entailed when the sources' declarations cover every triple pattern of the query; the check is NP-complete in combined complexity (www.inf.unibz.it), and a complete query admits a rewriting that sends each pattern to one source. Without covering declarations, the answer is incomplete or unknown.

What does completeness mean for a SPARQL federated query specifically? SERVICE sends part of a query to a remote endpoint, and without SILENT a failing call makes the query fail as a whole (www.w3.org). VALUES may be used to constrain what a remote endpoint returns (www.w3.org). The specification does not define completeness; it defines evaluation, so completeness must come from declarations plus a record of what each call returned.

Which completeness guarantees can a federated knowledge graph give? Only conditional ones: for named patterns, populations, sources, mapping versions and dates, assuming no wrong facts. Duplicates are handled by the evaluation semantics: DISTINCT eliminates duplicate solutions (www.w3.org), while REDUCED merely permits their elimination, so a row count under REDUCED is not a stable measure.

How do provenance and completeness relate? Provenance records where a value came from and when. A completeness declaration is a separate claim about what the source holds; a certificate cites provenance as evidence for it.

How is a federated query different from data integration? A federated query merges data at query time (www.w3.org), so each certificate names the live sources and their snapshots. A materialized integration freezes one graph, so the declaration must be about that integrated snapshot and its refresh date. The entailment test is the same; what changes is which identity and version the certificate names.

How can result completeness be verified in practice? Follow these steps:

- **Run the entailment** for the exact query and mapping versions, and keep the first uncovered frozen triple if it fails.
- **Withdraw each source** in turn and record which patterns lose coverage.
- **Execute the six fixture cases** against the pipeline and compare outcomes with Table 3.
- **Record endpoint status** for every call: answered, truncated, blocked or timed out.
- **Check declared cardinalities** against returned counts where any exist; cardinality assertions can be used to infer recall for a given selection condition (arxiv.org).
- **Check expiry** of every declaration against the query date.

- **Report answer recall** for a bounded reference set with an interval, labeled as an estimate and never as completeness.

What does an open-world knowledge graph say about a missing fact? That it may simply be missing but possibly true (www.w3.org). An ASK that returns `false` over such a graph therefore means no solution in the queried data, and nothing more unless a covering declaration exists.

Conclusion

A federated SELECT or ASK result is complete for a named population and time when, and only when, declarations about its sources entail completeness for the exact query at the stated mapping versions. The Darari et al. theory supplies the entailment test, the federated rewriting and the complexity bounds; it supplies no way to know that a declaration is true, and it does not cover erroneous or conflicting assertions. Tanon et al. show that cardinality information improves rule scoring, which is a different use of completeness metadata. Estimated recall can accompany a certificate and never replace it.

The recommended practice is simple to state. Name the query, the population and the date. Enumerate the declarations and their evidence. Compute the entailment and keep the failing pattern when it fails. Test source withdrawal, silent omission, disagreement, mapping change, a blocked endpoint and expiry. State the verdict as complete, incomplete or unknown, and let the unknown stand where the evidence ends. The main uncertainty is not computational. It is whether reliable, time-scoped declarations can be obtained for the sources that matter, and no source reviewed here shows that at scale.

Source links

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

1. https://www.inf.unibz.it/%7Enutt/Publications/18-compl-tweb.pdf
2. https://www.inf.unibz.it/%7Enutt/Publications/13-iswc.pdf
3. https://arxiv.org/pdf/1505.02463
4. https://lunadong.com/publication/kv_kdd.pdf
5. https://www.ijcai.org/proceedings/2018/0749.pdf
6. https://www.w3.org/TR/sparql11-federated-query
7. https://link.springer.com/content/pdf/10.1007/978-3-642-41338-4_18.pdf
8. https://www.w3.org/TR/sparql12-federated-query
9. https://www.scilit.net/publications/e4ad9e46ec2fb56e9f07be62796086eb
10. http://www.vldb.org/conf/1996/P402.PDF
11. https://www.vldb.org/pvldb/vol4/p749-razniewski.pdf
12. https://cris.maastrichtuniversity.nl/en/publications/knowledge-graph-completeness-a-systematic-literature-review
13. https://arxiv.org/abs/2305.05403
14. https://www.w3.org/TR/rdf11-semantics
15. https://www.w3.org/TR/owl2-primer

16. https://www.w3.org/TR/sparql11-query
17. https://ceur-ws.org/Vol-1666/paper-02.pdf
18. https://suchanek.name/work/publications/www2013.pdf
19. https://arxiv.org/pdf/2305.05403
20. https://arxiv.org/pdf/2001.04425
21. https://www.w3.org/TR/prov-o
22. https://arxiv.org/abs/1505.02463
23. https://lunadong.com/publication/dependence_vldb.pdf
24. https://arxiv.org/pdf/1503.00302
25. https://www.cl.cam.ac.uk/teaching/0910/InfoRtrv/lecture3.pdf
26. https://arxiv.org/pdf/1612.05786
27. https://arxiv.org/pdf/1909.01109
28. https://www.inf.unibz.it/%7Enutt/Publications/18-recoin-wikiWS.pdf
29. https://aclanthology.org/W16-1308.pdf
30. https://www.semantic-web-journal.net/system/files/swj1381.pdf
31. https://www.semantic-web-journal.net/system/files/swj954.pdf
32. https://www.mediawiki.org/wiki/Wikidata_Query_Service/User_Manual
33. https://www.wikidata.org/wiki/Help:Property_constraints_portal
34. https://www.wikidata.org/wiki/Help:Statements
35. https://www.w3.org/TR/sparql12-query

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.