Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

From corpus to claims: The TRACE framework for auditing inferential dependencies in bibliometric research

Greta Baldanzi

grebaldan@yahoo.it

Italian Ministry of Education and Merit

Research Article

Keywords: Bibliometrics, scientometrics, research methodology, inferential validity, research evaluation, reproducibility

Posted Date: September 29th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11175948/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Bibliometric research has developed increasingly explicit standards for study design, retrieval, reporting, reproducibility, indicator use, and interpretation. Yet procedural transparency does not by itself establish whether a bibliometric output warrants the substantive claim attached to it. This conceptual article introduces TRACE, a claim-centred architecture linking Theoretical Framing, Retrieval Sensitivity, Analytical Output, and Claim Calibration, with Epistemic Audit operating across the resulting stages and dependencies. ECART audits the same architecture in reverse. It first separates auditability from epistemic support, then completes the reverse traversal before identifying consequential problems and selecting a proportionate remedy. TRACE was developed through conceptual synthesis of adjacent methodological guidance rather than systematic review or validation. Four published cases illustrate different diagnoses at the output-to-claim and retrieval levels, while a reflexive audit constrains claims made for the framework itself. TRACE's contribution is to connect claim-level diagnosis with proportionate revision decisions through an explicit audit procedure that complements existing methodological guidance. Its reliability, usability, diagnostic performance, and effects on authoring and review decisions remain to be empirically evaluated.

1. Introduction

Bibliometric methods are now used across a wide range of disciplines to describe publication growth, map intellectual and social structures, examine collaboration, and assess patterns of scholarly visibility and influence. Their expansion has been accompanied by increasingly sophisticated software, accessible bibliographic databases, and methodological guidance intended to improve study design and reproducibility (Donthu et al., 2021; Öztürk et al., 2024). At the same time, the spread of bibliometric analysis beyond specialist scientometric communities has intensified concern about studies in which technically correct outputs are assigned meanings that the underlying measures, corpora, or research designs do not fully support (Ellegaard, 2026; Watson & Hayter, 2025).

A recent exchange in Scientometrics illustrates the issue without reducing it to a judgement about bibliometrics as a field. Ellegaard (2026) argues for stronger standards in the preparation and interpretation of bibliometric analyses, including clearer search-set construction, attention to database dependence, transparent parameters, and conclusions that remain justified by the results. Nazarovets (2026), while accepting the prevalence of low-quality and purely descriptive studies, cautions against treating such work as evidence that bibliometrics itself is a weak scientific tradition. The disagreement is therefore less about whether bibliometric methods can produce valid knowledge than about how methodological quality should be secured and how unevenly existing standards are applied across disciplines and journals.

This article addresses a narrower problem within that debate: the inferential dependency between a bibliometric output and the substantive claim formulated from it. A co-occurrence cluster is an analytical relation among terms; describing it as a theoretical tradition adds a substantive property. A publication

trend records indexed activity; describing the field as mature adds a developmental construct. A country ranking by publication count identifies relative output in a defined corpus; describing the country as scientifically dominant or as a research leader introduces a stronger evaluative proposition. These interpretations may sometimes be defensible, but they are not contained automatically in the outputs from which they are derived.

Existing guidance already addresses many conditions that bear on these transitions. Design frameworks connect research questions to data collection and analytical choices (Donthu et al., 2021; Öztürk et al., 2024). Work on search quality and field delineation addresses the validity of the literature set on which an analysis depends (Farooq et al., 2023; Zitt & Bassecoulard, 2006). GLOBAL strengthens reporting transparency (Ng et al., 2025, 2026). Hoang (2025) offers broader guidance for evaluating bibliometric reviews across data selection, cleaning, analysis, interpretation of conceptual and intellectual structures, and peer review, with VALOR providing an additional quality-control structure for multi-source studies. Lim and Kumar (2024) explicitly address the move from description to interpretation through scanning, sensing, and substantiating, including triangulation and return to the source literature represented in bibliometric structures. Responsible-metrics guidance, meanwhile, has long emphasised contextual interpretation and restraint in the use of quantitative indicators (Hicks et al., 2015; Waltman, 2016).

Across the methodological guidance examined in this conceptual synthesis, the remaining gap is not an absence of good practice. Rather, the relevant conditions are not explicitly organised as a claim-centred architecture for asking how epistemic stages and the dependencies between them jointly support, constrain, or fail to support a specified substantive claim. Transparency can reveal the choices made without establishing that the resulting claim is warranted. Reproducibility can show that the same procedure generates the same output without establishing that the output measures the property named in the conclusion. Sensemaking can enrich and substantiate interpretation without necessarily organising the entire path from framing and retrieval to output and claim as a claim-specific audit structure.

TRACE addresses this gap by treating bibliometric research as a claim-centred architecture composed of epistemic stages, inferential dependencies, and a control layer. The framework does not score whole studies or classify them as valid or invalid. Its unit of scrutiny is the substantive claim. TRACE asks whether a claim can be reconstructed through the methodological path that generated it, while ECART audits that path in reverse. ECART does not stop at the first problem encountered: it completes the reverse traversal, distinguishes non-assessability from insufficient support, and then identifies where intervention should begin. This claim-specific structure is intended to preserve findings that remain warranted while identifying only the analyses, reporting, or conclusions that require additional evidence, clarification, reframing, or narrower language.

The article first positions TRACE in relation to adjacent methodological approaches and explains the conceptual derivation of its architecture. It then formalises T, R, A, and C as epistemic stages, the connections between them as inferential dependencies, and E as the control layer that evaluates both

auditability and support. ECART is specified as a complete reverse audit with a controlled re-entry protocol. An analytic demonstration applies the framework to four published cases, including one retrieval-level case designed to distinguish a reporting limitation from an epistemic failure. The final sections consider implications for bibliometric research, peer review, and research evaluation, followed by limitations and a validation agenda.

2. Conceptual position and derivation

2.1 From methodological quality to inferential dependency

Methodological quality in bibliometric research is multidimensional. A transparent search can still retrieve a poorly delineated field; a well-delineated corpus can be analysed with an indicator that does not address the research question; a technically appropriate output can be interpreted beyond what it measures. These are not interchangeable weaknesses. Their importance depends on the claim that is ultimately made.

The distinction becomes clearer when reproducibility and inferential warrant are separated. Reproducibility asks whether a documented procedure can be repeated and whether it produces the same or a sufficiently comparable output. Inferential warrant asks whether that output, under the evidential conditions that produced it, supports the substantive proposition advanced by the researcher. A co-word analysis can therefore be reproducible while a claim that its clusters constitute theoretical schools remains under-justified. Likewise, a transparent Scopus search can support conclusions about the retrieved Scopus corpus while remaining insufficient for an unrestricted claim about an entire field.

This distinction also changes the role of limitations. Limitations are often placed at the end of an article after substantive conclusions have already been stated. From a claim-centred perspective, a limitation should alter the wording or scope of a claim whenever it constrains the evidence available for that claim. A known database coverage issue, unstable terminology, or indirect indicator is therefore not merely contextual information. It is part of the evidential conditions under which a conclusion becomes defensible.

2.2 Conceptual derivation of TRACE

TRACE was developed as a conceptual synthesis rather than through a systematic review, Delphi process, or empirical validation study. The aim was not to produce an exhaustive taxonomy of bibliometric quality criteria. Instead, the framework was derived functionally by asking which epistemic stages and inferential dependencies must remain inspectable and justifiable for a substantive bibliometric claim to be reconstructed from its evidential basis.

Four epistemic stages precede the audit itself and are represented as nodes in the TRACE architecture. Theoretical Framing (T) specifies and bounds the phenomenon. Retrieval Sensitivity (R) concerns the construction and stability of the claim-relevant corpus under plausible database, terminology, field,

language, time, and eligibility choices. Analytical Output (A) identifies the relation, measure, or structure directly produced by the selected procedure. Claim Calibration (C) formulates the substantive proposition at an explicit object, scope, and level of certainty. The dependencies are the links between these stages: T→R asks whether retrieval adequately represents the framed phenomenon; R→A asks whether the analytical procedure and resulting output are adequate to the corpus and question; A→C asks whether the output warrants the property, scope, and certainty asserted in the claim. Epistemic Audit (E) is structurally different. It is a control layer that evaluates both the internal adequacy of the nodes and the sufficiency of the dependencies connecting them.

The order is therefore functional rather than mnemonic. A claim cannot be audited without an output to which it can be traced; the output cannot be interpreted independently of the corpus on which it was calculated; and the corpus cannot be evaluated for representational adequacy without knowing what phenomenon it was intended to represent. TRACE names this forward architecture. ECART names its reverse traversal from an already formulated claim. The distinction also fixes the language of diagnosis: a problem 'at A' concerns the Analytical Output node itself, whereas a problem 'at A→C' concerns the inferential dependency from an otherwise interpretable output to the substantive claim.

2.3 Relation to adjacent methodological guidance

Table 1 is therefore presented as a derivation matrix rather than as a gap table. The functions on which TRACE depends are not claimed as novel in isolation: existing guidance already addresses study design, retrieval quality, interpretation, triangulation, reporting, reproducibility, and peer-review evaluation. The matrix makes explicit which functions TRACE draws from these contributions and where it reorganises them around a different unit of analysis: the individual substantive claim, evaluated through explicit nodes and dependencies rather than through a study-level score.

This distinction is particularly important in relation to Lim and Kumar (2024). Their sensemaking approach does more than recommend cautious interpretation: sensing and substantiating involve returning to the literature represented in bibliometric structures, triangulating interpretations, documenting supporting evidence, and checking whether an interpretation can be sustained. TRACE treats these functions as directly relevant to A→C and to Epistemic Audit rather than as absent from prior guidance. Its proposed additional organisation lies in making the individual claim the audit unit, completing the reverse traversal across upstream stages and dependencies, distinguishing auditability from support, and linking any consequential problem to a proportionate remedy and controlled re-entry.

The same caution applies to GLOBAL and Hoang (2025). GLOBAL makes methodological decisions visible through reporting guidance, while Hoang provides a broad framework for critical reading and peer review across data selection, cleaning, performance analysis, conceptual, intellectual and social structures, interpretation, and multi-source quality control through VALOR. TRACE relies on many of these existing functions. Its proposed contribution is their claim-centred procedural organisation: explicit localisation of a non-assessable or insufficiently supported node or dependency, separation of affected

from preservable downstream claims, selection of a proportionate remedy, and controlled re-entry into the forward chain before the claim is audited again.

Table 1

Derivation matrix: functions drawn from adjacent bibliometric guidance and their organisation in TRACE.

Source / guidance	Functions already addressed	Function taken up in TRACE	Additional organisation in TRACE
Donthu et al. (2021); Öztürk et al. (2024)	Research design, method selection, analytical workflow, and alignment between research questions and bibliometric techniques.	Provides design decisions that populate T, R, and A and make later inferential dependencies inspectable.	Reorganises those decisions around whether they support a specified substantive claim rather than evaluating the study only at workflow level.
Farooq et al. (2023); Zitt & Bassecoulard (2006)	Search quality, field delineation, and validity of the retrieved literature set.	Directly informs Retrieval Sensitivity and the relation between conceptual boundaries and corpus construction.	Makes retrieval adequacy claim-dependent: a retrieval choice matters insofar as it changes the evidential basis of the claim under audit.
GLOBAL (Ng et al., 2025, 2026)	Transparent and complete reporting of bibliometric methods and decisions.	Provides the visibility required for dependencies to be reconstructed and audited.	Treats transparency as an enabling condition for auditability rather than as inferential warrant: visible decisions can be assessed, while missing reporting is not equated with lack of epistemic support.
Hoang (2025), including VALOR	Critical evaluation of data selection, cleaning, performance analysis, conceptual/intellectual/social structures, interpretation, peer review, and multi-source quality control.	Contributes evaluative functions across R, A, C, and E, including verification, interpretation, and reproducibility concerns.	Uses the individual claim as the audit unit and adds explicit node/dependency localisation, downstream propagation, proportionate remedy, and controlled re-entry.
Lim & Kumar (2024)	Scanning, sensing, and substantiating; return to source literature; triangulation, documentation, and cross-checking of interpretations.	Directly informs A→C and the evidential work required to substantiate interpretation.	Makes the individual claim the audit unit; extends the audit upstream through R and T; distinguishes auditability from support; and localises consequential node/dependency problems only after complete reverse traversal.
Hicks et al. (2015); Waltman	Responsible use of indicators, contextual interpretation, methodological quality, and	Provides normative and methodological	Converts these constraints into a claim-specific audit

Source / guidance	Functions already addressed	Function taken up in TRACE	Additional organisation in TRACE
(2016); Ellegaard (2026)	restraint in substantive conclusions.	constraints on indicator use and claim strength.	procedure that distinguishes preservable findings from claims requiring qualification, reformulation, or withdrawal.
TRACE / ECART	Synthesises rather than replaces the above functions.	T-R-A-C are epistemic stages represented as nodes; T→R, R→A, and A→C are dependencies; E is the control layer; ECART is the complete reverse audit.	Combines claim-level audit, explicit node/dependency localisation, preservation of unaffected findings, proportionate remedy, and controlled re-entry after the full reverse traversal.

3. The TRACE framework

TRACE formalises four epistemic stages, three inferential dependencies, and one control layer. T, R, A, and C are represented as nodes; T→R, R→A, and A→C are the links whose sufficiency must be justified; E evaluates both nodes and links. Table 2 fixes this ontology and the corresponding audit questions. The elements are not independent quality criteria and are not intended to produce a score.

Table 2
TRACE architecture: epistemic stages, inferential dependencies, and control layer.

Element	Architectural role	Primary audit question	Characteristic risk
T - Theoretical Framing	Epistemic stage / node	Is the phenomenon explicitly defined and bounded?	Conceptual ambiguity or construct substitution before retrieval.
T→R	Inferential dependency	Does the retrieval strategy adequately represent the framed phenomenon?	Searchable representation under- or over-includes the conceptual object.
R - Retrieval Sensitivity	Epistemic stage / node	Is corpus construction coherent, documented, and robust to consequential plausible alternatives?	Database, terminology, field, language, time, or eligibility decisions materially alter the claim-relevant corpus.
R→A	Inferential dependency	Are the selected procedure and analytical conditions adequate to the corpus and research question?	The method or parameters construct an output that does not answer the intended analytical question.
A - Analytical Output	Epistemic stage / node	What relation, measure, or structure did the method directly produce, and is it correctly reported?	The operational output is misdescribed, unstable, or conflated with interpretation.
A→C	Inferential dependency	Does the output warrant the property, scope, and certainty asserted in the claim?	Interpretive or generalising claim exceeds the evidential range of output and corpus.
C - Claim Calibration	Epistemic stage / node	Is the substantive claim explicit in object, scope, and certainty?	Compound or ambiguous wording conceals a stronger proposition than the evidence supports.
E - Epistemic Audit	Control layer	Is each relevant node and dependency assessable, and where assessable, sufficiently supported?	Reporting opacity is mistaken for lack of support, or an unsupported link is left unpropagated.

3.1 T - Theoretical Framing

Theoretical Framing specifies the phenomenon before it becomes a search string. In bibliometric research this requires more than naming a topic. The researcher must state what counts as an instance of the phenomenon, what neighbouring constructs are excluded, and how conceptual boundaries relate to the vocabulary through which records can be retrieved. Field delineation research has long shown that bibliometric indicators depend critically on how the relevant publication set is defined (Zitt & Bassecoulard, 2006).

T becomes inferentially consequential when a claim is broader than the concept operationalised in retrieval. If a study defines a field through one contemporary label, a field-wide conclusion may silently

exclude earlier or discipline-specific terminology. Conversely, an intentionally narrow framing can be entirely appropriate when later claims remain bounded to that operational definition. The diagnostic question at T is therefore not whether the framing is maximally inclusive, but whether the conceptual object named in the claim corresponds to the object the study actually set out to represent.

3.2 R - Retrieval Sensitivity

Retrieval Sensitivity examines how dependent the evidential corpus is on plausible alternatives in database, terminology, search fields, time window, language, document type, and eligibility decisions. It goes beyond reporting the final search strategy. Farooq et al. (2023) show that weaknesses in synonym selection, Boolean logic, and search space can materially affect the quality of bibliometric search sets, while database-comparison studies demonstrate that coverage can alter the literature that becomes visible (Martín-Martín et al., 2018; Mongeon & Paul-Hus, 2016).

TRACE does not require exhaustive testing of every conceivable search configuration. The relevant alternatives are those that are plausible in relation to the theoretical framing and potentially consequential for the intended claim. Sensitivity should also be considered in terms of composition rather than record count alone: two retrieval strategies may return similar numbers of documents while changing disciplinary, linguistic, geographic, or thematic representation. Retrieval Sensitivity does not denote formal statistical sensitivity analysis; it asks whether plausible retrieval alternatives materially affect the corpus or the claims that depend on it. R is therefore a claim-dependent robustness question, not a requirement for maximal retrieval or mandatory use of multiple databases.

3.3 A - Analytical Output

Analytical Output identifies the relation, measure, or structure directly generated by the selected procedure under the reported analytical conditions. Different bibliometric techniques construct different representations of the same corpus. Co-citation represents patterns of joint citation; bibliographic coupling represents shared references; keyword co-occurrence represents joint term appearance; publication counts represent indexed output; citation indicators represent recorded citation relations under defined database and temporal conditions (Aria & Cuccurullo, 2017; Szomszor et al., 2021; Zupic & Čater, 2015).

A deliberately retains the operational description before substantive meaning is assigned. A community detected in a co-citation network is, at this stage, a structural community under a particular clustering procedure. A country with the largest document count is the largest contributor by that measure in the retrieved corpus. A rise in annual output is an increase in indexed publication activity. This restraint does not prevent interpretation. It creates a stable reference point from which the additional inferential work performed at C can be evaluated.

3.4 C - Claim Calibration

Claim Calibration concerns the transition from analytical output to substantive assertion. It asks what property is named in the claim, how far the claim extends beyond the analysed corpus, and what level of

certainty is expressed. These dimensions are diagnostic prompts rather than a new taxonomy. Their purpose is to make inferential distance visible.

The distinction between A and C is central to TRACE. Publication growth may contribute to an argument about the development of a field, but growth alone does not define maturity. A cluster may be consistent with an intellectual tradition, but the label 'tradition' ordinarily requires substantive knowledge about shared theories, concepts, or citation practices. High publication output may indicate productivity, while scientific leadership can imply influence, agenda setting, collaboration position, quality, or intellectual authority. Citation counts likewise reflect multiple citing and database-related factors and do not directly measure research quality (Bornmann & Daniel, 2008; Waltman, 2016). Calibration therefore concerns correspondence between the property measured and the property asserted.

Calibration does not require weak claims. Strong claims can be justified when the chain contains sufficient evidence, including qualitative validation, triangulation, theory, or sensitivity checks where appropriate. The requirement is proportionality. When evidence supports a narrower proposition than the one originally stated, the appropriate response is to narrow the claim rather than to dismiss the underlying analysis.

3.5 E - Epistemic Audit

Epistemic Audit is logically prior to an epistemic verdict because a dependency that is not visible cannot be evaluated. E therefore begins by assigning an auditability status to each relevant node and dependency: assessable when the published documentation permits a substantive judgement, partially assessable when only part of the relevant path can be reconstructed, and not assessable when reporting is insufficient for a judgement. Non-assessability is not evidence of non-support. It is a reporting or reconstructability limitation and must be recorded as such.

Where auditability is sufficient, E then evaluates the internal adequacy of each node and the inferential sufficiency of each dependency. A node can therefore be problematic in its own right, or the problem can lie in the link between otherwise acceptable nodes. E can be applied prospectively while a study is being designed and interpreted or retrospectively by authors, reviewers, editors, and readers. This two-step logic prevents a reporting deficit from being converted automatically into an epistemic failure.

3.6 TRACE as a dependency structure

Figure 1 represents the architecture. T, R, A, and C are epistemic stages represented as nodes; T→R, R→A, and A→C are inferential dependencies; E operates as a control layer across both. ECART begins from the target claim and traverses the complete sequence C, A→C, A, R→A, R, T→R, and T. Failure localisation and controlled re-entry are assigned only after this full reverse traversal.

Because nodes and dependencies are claim-specific, the same methodological condition can have different consequences within one study. A database limitation may be immaterial to a statement explicitly restricted to the analysed corpus but consequential to a claim about the global field. A

clustering choice may affect conclusions about intellectual structure while leaving publication-trend findings intact. The distinction between node and dependency matters diagnostically: a failure at A means that the analytical output itself is inadequate or insufficiently documented, whereas a failure at A→C means that the output can be sound while the substantive claim exceeds what it warrants. TRACE therefore evaluates propagation rather than assigning a single quality score to the paper as a whole.

4. ECART and controlled re-entry

4.1 Reverse traversal and failure localisation

Where TRACE describes forward construction, ECART formalises the reverse audit. The target is first fixed as a specific substantive claim at C. ECART then traverses the architecture in a fixed order: C, A→C, A, R→A, R, T→R, and T. At each node and dependency, the auditor first asks whether the element is assessable from the published documentation and only then, where assessment is possible, whether it is adequately supported. The sequence is fixed so that downstream symptoms are not mistaken for upstream causes.

ECART therefore does not stop at the first problem encountered during reverse reading. A problem at A→C may coexist with an earlier problem at R or T→R, and stopping at the proximal problem would conceal a more upstream condition that may also matter to the target claim. After the full traversal, ECART distinguishes the proximal consequential problem, meaning the issue closest to the claim, from the upstream-most consequential intervention point, meaning the furthest upstream non-assessable or insufficiently supported element whose clarification or correction could alter the evidential basis, scope, or certainty of that claim. An issue is consequential only when resolving it could change what the target claim is warranted to say; disclosed limitations that do not bear on that claim are recorded but do not determine re-entry.

Localisation is recorded at the level at which the diagnosis is made. 'At R' denotes a problem with corpus construction or its documentation; 'at T→R' denotes a mismatch between the framed phenomenon and its retrievable representation; 'at A' denotes a problem with the analytical output itself; and 'at A→C' denotes an inferential overreach from output to claim. Where reporting is insufficient, ECART records the element as not assessable rather than unsupported.

Localisation also distinguishes a node problem from a dependency problem. A node problem concerns the specification, construction, or reporting of T, R, A, or C itself; a dependency problem occurs when adjacent nodes may each be intelligible but the transition between them is insufficiently justified. The diagnostic traversal is complete when every claim-relevant node and dependency has an auditability status, every sufficiently assessable element has an epistemic status, and all consequential problems have been localised. Diagnostic completion does not imply that a remedy has already been applied.

4.2 Auditability, epistemic status, and remedial disposition

ECART separates three questions that should not be collapsed. Auditability concerns whether a node or dependency can be examined from the available documentation. Epistemic status concerns whether an assessable element provides sufficient support for the target claim. Remedial disposition concerns what should be changed once a consequential problem has been identified. This separation prevents absence of reporting from being treated as evidence that support is absent and prevents a diagnostic label from being confused with the revision it triggers.

For epistemic status, supported means that the available assessable evidence provides affirmative and sufficient warrant for the target claim at its stated scope and certainty; it is not assigned merely because no problem was detected. Qualified means that support holds only under explicit conditions or a narrower scope; unsupported means that the available assessable evidence is insufficient. Reformulation is not itself an epistemic status: it is one possible remedy, alongside preservation, additional reporting or clarification, reconstruction of the corpus, reanalysis, or additional substantiation. Table 3 makes these dimensions explicit.

Table 3
ECART auditability, epistemic status, and remedial dispositions.

Dimension	Status / disposition	Meaning and proportionate response
Auditability	Assessable	Documentation is sufficient to evaluate the relevant node or dependency. Proceed to epistemic judgement.
Auditability	Partially assessable	Only part of the relevant path can be reconstructed. State the boundary of the audit and avoid conclusions beyond the visible evidence.
Auditability	Not assessable	Reporting is insufficient for an epistemic verdict. Seek clarification or additional documentation; do not relabel the element as unsupported.
Epistemic status	Supported	Available assessable evidence provides affirmative and sufficient warrant for the target claim at its stated scope and certainty; absence of an identified problem alone is not enough.
Epistemic status	Qualified	Support holds under explicit conditions, a narrower scope, or stated uncertainty. Calibrate the claim accordingly.
Epistemic status	Unsupported	The available assessable evidence is insufficient for the target claim. Identify the upstream intervention point and revise.
Remedial disposition	Preserve / qualify / reformulate	Retain unaffected material; narrow scope or certainty; or replace the claimed property with one closer to the evidence.
Remedial disposition	Clarify / reconstruct / reanalyse / substantiate	Restore auditability or support at the relevant node/dependency, then propagate the change forward and re-audit.

4.3 Controlled re-entry protocol

TRACE also specifies how revision should proceed once the diagnostic traversal is complete. The controlled re-entry protocol is rule-governed rather than mechanical: (1) fix the target claim; (2) complete the reverse traversal and record auditability before epistemic status at every relevant node and dependency; (3) identify the proximal consequential problem and the upstream-most consequential intervention point; (4) preserve findings and claims that do not depend on that point; (5) choose the minimum proportionate remedy in light of the claim the author intends to retain; (6) propagate any resulting change forward through dependent stages; and (7) re-run ECART on the revised claim. Two routes are therefore possible. If the original claim is to be preserved, re-entry begins at the upstream-most point whose repair is required to support it. If the claim is narrowed or replaced so that it no longer depends on that upstream condition, the proportionate remedy may instead occur at C, while the unresolved upstream limitation remains explicitly bounded outside the revised claim. Reanalysis is required only when clarification or recalibration cannot restore an adequate claim-evidence correspondence. The audit cycle is closed only after the selected remedy has been applied, any resulting change has been propagated through dependent stages, and the revised claim has been subjected to a new ECART traversal. Closure does not require every methodological limitation to disappear: an unresolved condition may remain if the revised claim is explicitly bounded so that it no longer depends on that condition.

This distinction makes proportionality claim-dependent: the same methodological limitation may require reanalysis for a broad field-level claim but only explicit scope restriction for a corpus-bounded descriptive claim.

5. Analytic demonstration on published bibliometric claims

To examine whether TRACE can do more than restate familiar methodological cautions, this section applies ECART to four published bibliometric cases. The purpose is analytic demonstration, not empirical validation of the framework and not a ranking of the studies concerned. Three cases examine output-to-claim transitions involving field maturity, theoretical traditions, and research leadership. A fourth follows the audit further upstream to a documented cross-database retrieval asymmetry, while distinguishing the existence of that asymmetry from any untested effect on the corpus or conclusions.

The demonstration asks a deliberately limited operational question: can the same audit architecture distinguish a claim that requires qualification, a stronger interpretation that is substantively supported, a claim that should be reformulated, and a retrieval-level problem whose epistemic status depends on what is actually assessable from the published record? It does not estimate prevalence, establish inter-rater agreement, demonstrate usability, or validate diagnostic performance across the full TRACE architecture.

5.1 Case selection and audit procedure

Cases were selected purposively to provide diagnostic coverage rather than a representative sample of bibliometric studies. Each case had to be recent and publicly accessible, contain an identifiable substantive proposition or field-level inference, report the evidence invoked in support of it, and provide enough methodological detail for at least part of the relevant TRACE path to be reconstructed. The first three cases were selected to contrast different A→C transitions; the fourth was selected specifically because the published retrieval description makes an R-level dependency visible. Selection was based on diagnostic contrast, not journal prestige or a presumption that the published study was defective.

The audit procedure was held constant across cases. The target proposition was fixed at C, after which ECART traversed C, A→C, A, R→A, R, T→R, and T. At each element, auditability was considered before epistemic status. The main text reports the diagnostically material elements, while Online Resource 1 provides the complete audit record for every case, including evidence location, auditability, epistemic status, localisation, and remedial disposition. After the complete traversal, consequential issues were distinguished from non-consequential limitations and the minimum proportionate response was specified.

The supplementary audit records are included so that readers can verify not only the local diagnosis reported in the narrative but also why no more upstream element was treated as consequential for the target proposition.

The analysis is deliberately internal to each published article: it uses the authors' reported search, methods, outputs, and substantive statements and does not reconstruct underlying datasets, infer author intentions, or judge the overall quality of the paper. Where the published report does not make an element inspectable, TRACE records that limit as partial or non-auditability rather than inferring that the study lacks support. This restriction is important because the demonstration illustrates claim-specific diagnosis, not post-publication peer review of the selected studies.

5.2 Case 1: thematic evolution and field maturity

Chen et al. (2025) analyse research on gut microbiota during pregnancy using Web of Science Core Collection and Scopus. Their analyses include publication trends, collaboration, co-citation, keyword co-occurrence, thematic evolution, and citation bursts. The conclusion states that the field has matured from descriptive ecology toward mechanistic and translational research. ECART separates two propositions embedded in that sentence: that the thematic focus changed in the stated direction, and that this change establishes field maturity.

At A, the reported keyword and timeline analyses provide direct evidence that research topics shifted from microbial composition toward mechanisms linking the microbiota to pregnancy complications, neurodevelopmental outcomes, and translational applications. That evidence supports the directional transition. At A→C, however, 'maturity' introduces a broader construct that is not operationally defined by the reported bibliometric measures. The audit therefore assigns different epistemic statuses to the two

propositions: the thematic transition is supported, whereas maturity is not established by the evidence considered.

The proportionate remedy is not to weaken the supported transition but to separate it from the maturity claim. A calibrated formulation can state that the thematic focus of the retrieved literature shifted from descriptive microbial composition toward increasingly mechanistic and translational topics. Retaining 'maturity' would require an explicit definition and evidence aligned with that construct. This is reformulation as a remedy following an unsupported A→C dependency for the maturity proposition, not an epistemic status in itself.

5.3 Case 2: co-citation communities and theoretical traditions

Kurmasheva et al. (2026) map digital adaptive-learning research. The target claim is directly anchored in Section 3.5, 'Intellectual foundations and knowledge flows', immediately after Fig. 5 ('Co-citation analysis of authors'), where the authors state that 'the co-citation network reveals three additional theoretical traditions'. They then identify those traditions through the scholars and literatures represented in the clusters, including cognitive science and instructional design, self-regulated learning and metacognition, and context-aware or ubiquitous learning. The same section reports a modularity of 0.395 and substantially stronger within-cluster than between-cluster link strength. TRACE therefore audits the authors' own interpretive label rather than introducing a stronger one.

At A, the analysis establishes structurally differentiated co-citation communities. At A→C, 'theoretical tradition' adds a substantive interpretation concerning intellectual coherence. In this case, however, the authors do not rely on cluster geometry alone: they return to the scholars and conceptual orientations represented within the clusters and use those bodies of work to substantiate the interpretation. Under the evidence reported in Section 3.5, the A→C dependency is therefore supportable rather than categorically failing.

The claim can therefore be retained, provided its substantive basis remains explicit. A formulation such as 'the co-citation structure is consistent with several partially distinct theoretical traditions, as substantiated by the literatures represented within the clusters' makes the dependency visible without weakening the authors' underlying interpretation. If a study supplied only cluster membership and modularity, the same label would require additional content-level substantiation. This case shows that TRACE is not designed to collapse interpretation back into operational description; it distinguishes an interpretive claim that is supported by additional evidence from one inferred from clustering alone.

5.4 Case 3: publication output, collaboration centrality, and research leadership

You et al. (2024) examine higher-education leadership-development research in Scopus. Section 3.3 reports that the United States is connected with 36 countries and is in a leadership position for global

research collaboration. The same discussion then extends the network result to broader claims about research prowess and leadership in the field. TRACE therefore audits that broader extension rather than treating the narrower collaboration-position statement as a claim of general scientific superiority.

At A, publication counts support statements about indexed output and the co-authorship network supports statements about collaboration connectivity. The narrower collaboration-position claim can therefore be interpreted in relation to the reported network. At A→C, however, extending those outputs to broader 'research prowess' or field-wide research leadership adds properties that the reported indicators do not independently define. ECART also identifies an upstream scope condition: the study uses Scopus, English-language journal articles, and a 2013–2023 window, and the authors themselves note that these choices may omit relevant literature and bias coverage toward English-speaking countries. That limitation is potentially consequential to a genuinely global leadership claim, but not necessarily to a corpus-bounded statement about the observed network. Bibliometric leadership indicators additionally require explicit conceptualisation and are not interchangeable with productivity or connectivity measures (Moya-Anegón et al., 2018).

The epistemic diagnosis and the remedy must remain separate. Collaboration prominence within the retrieved corpus is supported by the reported network; the broader research-prowess/field-leadership extension is unsupported by those indicators alone, and its global scope is additionally qualified by retrieval coverage. This case therefore illustrates the two remedial routes specified in Section 4.3. Preserving a broad global leadership claim would require both a defined leadership construct and evidence that retrieval choices do not materially change the country-level result. Narrowing the claim to publication output and collaboration prominence within the analysed Scopus corpus avoids that stronger dependency and does not require unnecessary reanalysis. The paper also reports different corpus totals in its abstract and Methods; Online Resource 1 records this as a reporting inconsistency whose effect on the specific network result is not assessable from the published record.

5.5 Case 4: retrieval asymmetry and field-level topic evolution

Zhang et al. (2025) examine the evolution of AI-literacy research from 2015 to 2024 using Web of Science and Scopus. Section 3.2, 'Data Sources and Preprocessing', reports a Web of Science query joining 'Artificial Intelligence literacy' and 'AI literacy' with OR, whereas the Scopus search joins the same expressions with AND in Title, Abstract, and Keywords. After database-field standardisation and merging, the authors report an assembled dataset of 3,801 publications; subsequent cleaning and integrity checks retain 3,361 validated publications for topic analysis. The target proposition for TRACE is the field-level trajectory inferred from that analysed set.

At A, topic modelling and semantic analyses can describe patterns within the 3,361-publication analysis set. ECART nevertheless continues upstream. At R, the documented search logics are not equivalent: a record can enter the Web of Science set by containing either expression, while the reported Scopus logic requires both. This establishes a documented retrieval asymmetry. It does not establish that the

asymmetry changed the merged corpus materially, because records omitted by Scopus may already have been recovered by Web of Science, nor does it establish that any corpus difference would alter the topic trajectory. The published report also does not show whether the AND expression reflects the query actually executed or a reporting error.

The diagnosis is therefore deliberately staged. The documented asymmetry is assessable; its effect on the merged corpus and on the substantive conclusions is not established from the published record. The first remedy is clarification of the executed query. If the reported logic is accurate, the next step is a harmonised retrieval check comparing the incremental record set and its composition. Reanalysis is warranted only if that check shows a consequential difference for the target claim. If the harmonised retrieval produces no material change, the original analysis may be retained with the robustness result reported. If the difference is consequential, the affected analyses should be rerun and ECART repeated. This graduated response distinguishes documented asymmetry, corpus effect, and conclusion effect rather than treating non-equivalent syntax as proof of distorted findings.

Case 4

also makes the decision sequence visible. C is assessable as a field-level trajectory claim; A and A→C are assessable for the analysed set; the audit then reaches R, where the published search description reveals the OR/AND asymmetry. Because the effect of that asymmetry on the merged and cleaned corpus is not reported, the audit records partial assessability rather than an unsupported field-level conclusion. The remedial branch is then determined by the result of clarification and harmonised retrieval testing: either preserve the original claim with documented robustness, recalibrate its scope, or rerun the affected analysis if the corpus difference proves consequential.

5.6 What the demonstration adds

The four cases produce different diagnoses without using remedy labels as epistemic verdicts. Case 1 separates a supported thematic transition from a maturity proposition not established by the reported outputs. Case 2 retains the authors' theoretical-traditions interpretation because co-citation structure is accompanied by substantive cluster-level corroboration. Case 3 preserves collaboration prominence while distinguishing it from a broader research-prowess/field-leadership extension. Case 4 reaches R but stops short of inferring an unobserved corpus or conclusion effect from the documented query asymmetry. Table 4 summarises these distinctions; Online Resource 1 makes the complete reverse audits inspectable.

The procedural contribution illustrated here is the linkage between evidence, diagnosis, and intervention. Existing guidance already supports careful interpretation, triangulation, transparent reporting, and critical evaluation. TRACE adds a claim-specific record of where a consequential problem is localised, whether the problem is epistemic or merely non-assessable, which conclusions remain preservable, and whether proportional correction requires clarification, claim recalibration, substantiation, retrieval testing, or

reanalysis. The demonstration illustrates this organisation; it does not establish usability, reliability, validity, or superiority over adjacent guidance.

Table 4
Summary of the analytic demonstration.

Study / audited proposition	Evidence / location	Auditability	Epistemic status / localisation	Remedial disposition
Chen et al. (2025): thematic shift	Keyword/timeline analyses; Results/Conclusion	Assessable	Supported	Preserve the directional transition.
Chen et al. (2025): field maturity	Same outputs; Conclusion	Assessable	Unsupported at A→C for the maturity construct	Reformulate to thematic transition or define and substantiate maturity separately.
Kurmasheva et al. (2026): theoretical traditions	Section 3.5, Fig. 5; cluster interpretation and structural metrics	Assessable	Supported at A→C	Preserve, with the substantive cluster-level basis explicit.
You et al. (2024): broader research prowess / field leadership	Section 3.3; collaboration network; Scopus-only English-language corpus	Assessable at A→C; retrieval scope only partly resolves the global claim	Unsupported for the broader A→C extension; global scope additionally qualified at R/T→R	Either substantiate leadership and retrieval robustness, or narrow to output/network prominence within the analysed corpus.
Zhang et al. (2025): field-level topic trajectory	Section 3.2; OR/AND retrieval asymmetry; 3,801 assembled records, 3,361 analysed	Partially assessable at R	No claim-level verdict on retrieval effect: asymmetry is documented, corpus/conclusion effect is not established	Clarify executed query; test harmonised retrieval; preserve, recalibrate, or reanalyse only according to observed consequentiality.

5.7 Reflexive audit: applying TRACE to the present manuscript

As a reflexive check, TRACE was applied to three claims made in the present manuscript. This is not validation of the framework. It is a consistency test asking whether the paper applies the same distinction between auditability, epistemic status, and remedy to its own claims that it requires of the external cases. Table 5 reports the result in the same three-part form.

The audit led to three constraints. First, the literature-gap claim is restricted to the guidance examined because the conceptual synthesis is non-systematic. Second, the four cases illustrate application at A→C and an R-level auditability problem, but they do not demonstrate validated localisation across the full architecture or at T. Third, author-applied cases illustrate procedural applicability but do not establish usability, reliability, or inter-rater agreement.

These constraints were propagated into the manuscript through the derivation matrix, the formal separation of nodes and dependencies, complete reverse traversal, the auditability distinction, and the more limited description of what the analytic demonstration establishes. The reflexive audit is therefore evidence of self-constraint in claim formulation, not evidence that TRACE has validated itself.

Table 5
Reflexive TRACE audit of claims made in the present manuscript.

Claim audited	Evidence basis	Auditability	Epistemic status	Remedial disposition
Existing guidance lacks a claim-centred structure.	Non-systematic conceptual synthesis of adjacent guidance.	Not assessable at field-wide scope; assessable for the guidance examined.	No field-wide verdict; the organisational distinction is supported only within the guidance examined.	Restrict the claim to the examined guidance and specify TRACE's organisational difference.
ECART can localise consequential problems across the full architecture.	Four author-applied cases: A→C coverage and one R-level auditability case; no T-level case or independent raters.	Partially assessable.	No full-architecture performance verdict; application is illustrated at selected A→C and R points.	State illustrated coverage only; reserve full-chain performance and agreement for validation.
The analytic demonstration provides evidence of usability.	Author application to purposively selected cases; no independent users or usability measures.	Assessable.	Unsupported as a usability claim.	Reformulate as procedural illustration; retain usability as a future validation question.

6. Implications for bibliometric research, peer review, and research evaluation

TRACE has implications for bibliometric research at three levels. First, during study design it makes the intended claim part of methodological planning. The question is not only which analysis can be performed on an available dataset, but what evidential chain would be required to justify the proposition the study ultimately seeks to make. This can alter retrieval testing, indicator choice, triangulation, and the level at which conclusions are formulated.

Second, TRACE provides a vocabulary for peer review that is more specific than labelling a conclusion 'too strong'. A reviewer can distinguish a node problem from a dependency problem and, separately, a non-assessable element from an unsupported one. This matters when reporting is incomplete: a reviewer can request the information needed to make the audit possible without claiming that the hidden methodological choice was necessarily wrong. ECART converts these concerns into a reconstructable sequence and a proportionate request.

Third, the framework connects bibliometric field mapping with principles already familiar in responsible research evaluation. The Leiden Manifesto and subsequent guidance emphasise that indicators require contextual interpretation and should not substitute automatically for expert judgement (Hicks et al., 2015). TRACE operates at a different level but shares the same resistance to metric reductionism: it asks whether the property asserted in a claim is actually supported by the indicator and evidential context used to produce it.

The framework may also help distinguish requests for additional evidence from routine methodological maximalism. Multiple databases, qualitative reading, robustness checks, and triangulation can all strengthen bibliometric research, but TRACE does not prescribe them universally. They become necessary when a consequential node or dependency remains unsupported for the claim being advanced. Additional reporting becomes necessary when that node or dependency is not assessable. A corpus-bounded descriptive claim may require no cross-database triangulation; a field-wide generalisation based on unstable or non-equivalent retrieval may. This claim-dependent logic is intended to make both authoring and reviewing more proportionate.

For editors, the same structure offers a way to separate deficiencies that affect the scientific contribution from those that can be corrected through calibration. A paper need not be treated as globally invalid because one interpretive claim overreaches. Conversely, a limitations paragraph should not be treated as sufficient when a known methodological constraint continues to underpin an unrestricted conclusion.

7. Limitations and validation agenda

TRACE remains a conceptual and methodological framework. The present article does not establish inter-rater reliability, construct validity, predictive validity, usability, or an effect on the quality of bibliometric review. The four purposively selected cases illustrate differentiated claim-level use at A→C and show how an R-level reporting asymmetry can be audited without assuming a material corpus or conclusion effect. They do not estimate the prevalence of inferential overreach, demonstrate a T-level case, or provide evidence that independent users will localise the same consequential problems.

The conceptual derivation also has limitations. The architecture was developed through functional synthesis of methodological problems and adjacent guidance rather than through a systematic evidence review or formal consensus process. Its defensibility therefore rests at this stage on the coherence and completeness of the dependency argument. Future work should test whether additional nodes or

dependencies are needed, whether different users agree on auditability status and node/dependency localisation after a complete reverse traversal, and whether remedial dispositions can be applied consistently.

Validation should proceed in stages. Retrospective studies could ask independent reviewers to audit a common set of bibliometric claims and measure agreement on auditability, node versus dependency localisation, consequentiality, and remedial choice. Prospective studies could compare projects developed with and without TRACE to determine whether the framework changes retrieval checks, indicator selection, substantiation, reporting, or conclusion wording. Comparative studies could also test whether TRACE yields decisions that are more transparent or reproducible than those reached through existing guidance alone. Such studies would evaluate the procedural advantage proposed here rather than assuming it from the conceptual architecture.

A further limitation concerns expertise and scale. T and C often require substantive knowledge of both the mapped field and bibliometric methods; they cannot be reduced safely to mechanical scoring. The auditability judgement can likewise require contextual expertise when reporting is partial rather than absent. Claim-specific auditing across large manuscripts may also be resource-intensive. Digital or AI-supported tools could eventually assist with traceability, citation-to-claim mapping, and documentation of nodes and dependencies, but conceptual framing and inferential judgement should remain human responsibilities.

Finally, TRACE has been developed for bibliometric research. The broader logic of reconstructing claims through framing, evidence construction, analysis, and calibration may be relevant to other forms of evidence synthesis, but transfer beyond bibliometrics requires separate conceptual and empirical examination.

8. Conclusion

The current methodological debate in bibliometrics is not adequately described as a choice between confidence in bibliometric methods and scepticism about their proliferation. A more tractable question is whether particular substantive claims remain justified by the methodological conditions that produced them. Transparency, reproducibility, responsible indicator use, and careful interpretation all contribute to that judgement, but none alone provides a claim-centred audit.

TRACE formalises bibliometric inference as an architecture of four epistemic stages, three inferential dependencies, and an audit control layer. ECART traverses the complete structure in reverse, separates auditability from epistemic support, and localises consequential problems before a remedy is chosen. The analytic demonstration shows how this procedure can preserve supported propositions, isolate unsupported output-to-claim extensions, and handle a retrieval asymmetry without treating it as evidence of an unobserved distortion. The reflexive audit applies the same constraints to claims made for TRACE itself. Demonstrating T-level application, inter-rater agreement, usability, and reliable full-chain performance remains part of the validation agenda.

Rather than replacing existing design, reporting, evaluation, or sensemaking guidance, TRACE reorganises their inferential relevance around the reconstructability of individual substantive claims. Its empirical test is whether independent researchers and reviewers can apply the audit categories consistently, make consequentiality judgements transparently, and use claim-preserving repair or claim-restricting recalibration proportionately. Until those tests are performed, TRACE should be read as a specified audit method and validation agenda rather than as a demonstrated improvement in review performance.

Declarations

Supplementary Information

Online Resource 1: Complete ECART audit records for the four published cases, including evidence location, auditability, epistemic status, localisation, and remedial disposition.

Competing interests:

The author has no relevant financial or non-financial interests to disclose.

Use of generative AI

ChatGPT (OpenAI GPT-6 Sol) was used under author supervision for literature-search support, comparative stress testing, and language and editorial refinement. The author reviewed and revised all AI-assisted material, independently verified factual claims, citations, and bibliographic information against publisher or other authoritative sources, and takes full responsibility for the framework, interpretations, and final manuscript. No confidential, proprietary, human-subject, or otherwise sensitive data were processed through the AI tool.

Funding:

No funding was received to assist with the preparation of this manuscript.

Author Contribution

G.B. conceived the TRACE framework, developed the methodology, conducted the conceptual synthesis and analytic demonstration, and wrote and revised the manuscript. G.B. approved the final manuscript and is solely responsible for its content.

Data availability:

No new datasets were generated or analysed for the conceptual framework. The analytic demonstration uses information reported in published articles cited in the manuscript.

References

1. Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, *11*(4), 959–975. https://doi.org/10.1016/j.joi.2017.08.007
2. Bornmann, L., & Daniel, H.-D. (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*, *64*(1), 45–80. https://doi.org/10.1108/00220410810844150
3. Chen, A., Luo, Z., & Zhang, J. (2025). Gut microbiota during pregnancy: A bibliometric analysis of global research trends and collaborative networks. *Frontiers in Microbiology*, *16*, 1708359. https://doi.org/10.3389/fmicb.2025.1708359
4. Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, *133*, 285–296. https://doi.org/10.1016/j.jbusres.2021.04.070
5. Ellegaard, O. (2026). The method of bibliometric analysis: A critical evaluation. *Scientometrics*, *131*, 4175–4196. https://doi.org/10.1007/s11192-026-05670-6
6. Farooq, U., Nasir, A., & Khan, K. I. (2023). An assessment of the quality of the search strategy: A case of bibliometric studies published in business and economics. *Scientometrics*, *128*, 4855–4874. https://doi.org/10.1007/s11192-023-04765-8
7. Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature*, *520*, 429–431. https://doi.org/10.1038/520429a
8. Hoang, A.-D. (2025). Evaluating bibliometrics reviews: A practical guide for peer review and critical reading. *Evaluation Review*, *49*(6), 1074–1102. https://doi.org/10.1177/0193841X251336839
9. Kurmasheva, M., Zhorabekova, A., Barikulov, R., Duisenova, M., & Ashimova, D. (2026). Mapping research on digital adaptive learning in education: A bibliometric approach. *Cogent Education*, *13*(1), 2699507. https://doi.org/10.1080/2331186X.2026.2699507
10. Lim, W. M., & Kumar, S. (2024). Guidelines for interpreting the results of bibliometric analysis: A sensemaking approach. *Global Business and Organizational Excellence*, *43*(2), 17–26. https://doi.org/10.1002/joe.22229
11. Martín-Martín, A., Orduna-Malea, E., & Delgado López-Cózar, E. (2018). Coverage of highly-cited documents in Google Scholar, Web of Science, and Scopus: A multidisciplinary comparison. *Scientometrics*, *116*(3), 2175–2188. https://doi.org/10.1007/s11192-018-2820-9
12. Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, *106*, 213–228. https://doi.org/10.1007/s11192-015-1765-5

13. Moya-Anegón, F., Guerrero-Bote, V. P., López-Illescas, C., & Moed, H. F. (2018). Statistical relationships between corresponding authorship, international co-authorship and citation impact of national research systems. *Journal of Informetrics*, *12*(4), 1251–1262. https://doi.org/10.1016/j.joi.2018.10.004
14. Nazarovets, S. (2026). Poor bibliometric studies do not make bibliometrics poor science: A response to Ellegaard. *Scientometrics*, *131*, 7475–7477. https://doi.org/10.1007/s11192-026-05791-y
15. Ng, J. Y., Liu, H., Masood, M., Syed, N., Waltman, L., Haustein, S., Sabé, M., Solmi, M., Stephen, D., Schniedermann, A., Willemin, S., Gauffriau, M., Andersen, J. P., Bornmann, L., Costas, R., Glänzel, W., Hinze, S., Robinson-Garcia, N., Sivertsen, G., … Moher, D. (2025). Guidance List for Reporting Bibliometric Analyses (GLOBAL): A two-round modified Delphi study. In *20th International Conference on Scientometrics and Informetrics* (pp. 1–25). https://doi.org/10.51408/issi2025_069
16. Ng, J. Y., Syed, N., Zubashev, D., Masood, M., Liu, H., Sabé, M., Solmi, M., Stephen, D., Andersen, J. P., Bornmann, L., Costas, R., Gauffriau, M., Glänzel, W., Hinze, S., Neylon, C., Robinson-Garcia, N., Schniedermann, A., Sivertsen, G., Stahlschmidt, S., … Haustein, S. (2026). *Guidance List for repOrting Bibliometric AnaLyses (GLOBAL): Explanation and elaboration* [Preprint]. Research Square. https://doi.org/10.21203/rs.3.rs-9322986/v1
17. Öztürk, O., Kocaman, R., & Kanbach, D. K. (2024). How to design bibliometric research: An overview and a framework proposal. *Review of Managerial Science*, *18*, 3333–3361. https://doi.org/10.1007/s11846-024-00738-0
18. Szomszor, M., Adams, J., Fry, R., Gebert, C., Pendlebury, D. A., Potter, R. W. K., & Rogers, G. (2021). Interpreting bibliometric data. *Frontiers in Research Metrics and Analytics*, *5*, 628703. https://doi.org/10.3389/frma.2020.628703
19. Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, *10*(2), 365–391. https://doi.org/10.1016/j.joi.2016.02.007
20. Watson, R., & Hayter, M. (2025). Bibliometric analyses: Do they contribute to knowledge? *Journal of Clinical Nursing*, *34*(4), 1101–1102. https://doi.org/10.1111/jocn.17423
21. You, C., Awang, S. R., & Wu, Y. (2024). Bibliometric analysis of global research trends on higher education leadership development using Scopus database from 2013–2023. *Discover Sustainability*, *5*, 246. https://doi.org/10.1007/s43621-024-00432-x
22. Zhang, T., Zhang, P., & Ma, H. (2025). Tracing the topic evolution of AI literacy literature: From knowledge enlightenment to knowledge symbiosis. *Knowledge Organization*, *52*(7), 44114. https://doi.org/10.31083/KO44114
23. Zitt, M., & Bassecoulard, E. (2006). Delineating complex scientific fields by an hybrid lexical-citation method: An application to nanosciences. *Information Processing & Management*, *42*(6), 1513–1531. https://doi.org/10.1016/j.ipm.2006.03.016
24. Zupic, I., & Čater, T. (2015). Bibliometric methods in management and organization. *Organizational Research Methods*, *18*(3), 429–472. https://doi.org/10.1177/1094428114562629

Figures

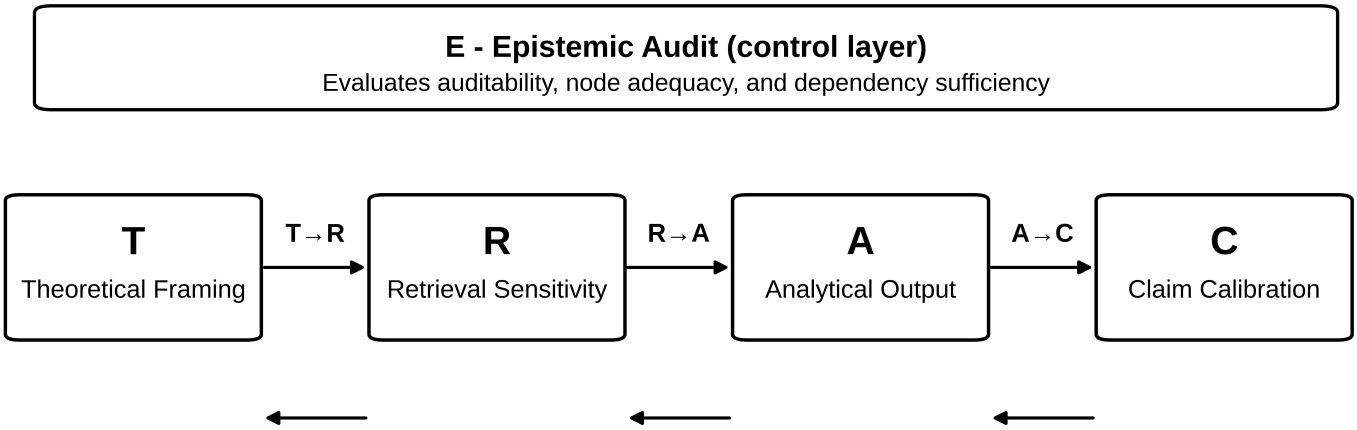


Figure 1

TRACE architecture: T, R, A, and C as epistemic stages, the inferential dependencies connecting them, E as the control layer, and ECART as the complete reverse audit

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ESM1TRACEECARTaudits.pdf