Eurasia Specialized Veterinary Publication

Entomological Research Letters

ISSN: 3108-639X

2026, Volume 6, Issue 1, Page No: 82-91
Copyright CC BY-NC-SA 4.0
Available online at: www.esvpub.com/

Taxonomic Revisions Should Function as Versioned and Reproducible Research Objects With Traceable Evidence Provenance

Santiago Morales[1]*, Laura Rojas[2], Camila Vega[1], Diego Molina[3]

[1]Department of Versioned Taxonomy and Reproducible Systematics, Faculty of Biosciences, University of Buenos Aires, Buenos Aires, Argentina.
[2]Department of Traceable Evidence and Research Objects, Faculty of Life Sciences, Pontifical Catholic University of Chile, Santiago, Chile.
[3]Department of Taxonomic Revision and Provenance, Faculty of Sciences, National University of La Plata, La Plata, Argentina.

***E-mail** ✉ santiago.morales@fce.uba.ar

ABSTRACT

Taxonomic revisions are usually published as stable scholarly outputs even though the conclusions they contain arise from changing constellations of specimens, character observations, sequence data, analytical procedures, species hypotheses, and nomenclatural decisions. This mismatch makes it difficult to reconstruct why a taxonomic conclusion was reached, determine which evidence belonged to a particular revision state, or distinguish later reinterpretation from silent replacement. We argue that insect taxonomic revisions should instead be treated as versioned and reproducible research objects. The proposed architecture separates physical specimens and types, observations, character and genomic derivatives, analytical outputs, interpretations, taxonomic decisions, and nomenclatural acts while preserving explicit relationships among them. Evidence provenance records where an observation or derivative originated; decision provenance records how heterogeneous and sometimes discordant evidence was interpreted; version provenance records how matrices, genomic datasets, analyses, and hypotheses changed through time. This separation is important because coalescent groups, barcode divergences, gene trees, algorithmic clusters, morphology, behavior, life-stage characters, and genomic signals do not have interchangeable inferential meanings. The approach does not promise automatic species resolution or nomenclatural correctness. Rather, it is intended to make revision histories reconstructable, auditable, and capable of retaining rejected, superseded, and uncertain hypotheses alongside accepted ones. Implementation will require persistent identifiers, specimen-to-derivative linkage, explicit dataset versions, machine-readable relationships, and validation across contrasting taxonomic workflows.

Keywords: Digital taxonomy, Evidence provenance, Integrative taxonomy, Persistent identifiers, Reproducible systematics, Taxonomic versioning.

Received: 17 Month 2026
Revised: 09 Month 2026
Accepted: 13 Month 2026

How to Cite This Article: Morales S, Rojas L, Vega C, Molina D. Taxonomic Revisions Should Function as Versioned and Reproducible Research Objects With Traceable Evidence Provenance. Entomol Appl Sci Lett. 2026;6(1):82-91. https://doi.org/10.51847/hycyhUzs5z

Introduction

A taxonomic revision is conventionally encountered as a publication, yet the scientific object being revised is distributed across physical specimens, type material, collection records, measurements, sequence accessions, character matrices, analytical files, and interpretive decisions. Persistent identifiers and extended- and digital-specimen initiatives already provide mechanisms for making such biological research objects durable and linkable

© 2026 Entomological Research Letters

across repositories [1–3]. The unresolved problem is that these mechanisms generally identify components of taxonomic work rather than the changing evidential state of the revision itself.

Computational reproducibility creates a parallel problem. A documented workflow must expose its version, dependencies, metadata, and reusable components if another researcher is to reconstruct what was actually executed [4]. Workflow registries demonstrate that version histories, creator information, persistent citation links, and relationships between successive workflow states can be represented explicitly rather than remaining implicit in prose [5]. Taxonomic revisions increasingly depend on comparable computational objects, but a final tree, matrix, delimitation output, or supplementary archive rarely expresses the complete dependency chain by which a nomenclatural conclusion emerged.

The central argument developed here is therefore methodological rather than taxonomic: a revision should be represented as a proposed version-addressable compound research object. Its physical evidence, digital derivatives, analytical states, interpretations, decisions, and nomenclatural consequences should remain distinguishable but connected. This architecture is not proposed as a mechanism for deciding what a species is. It instead addresses a different problem: whether another researcher can determine what evidence existed at a given revision state, how that evidence was transformed, why competing hypotheses were accepted or rejected, and what changed when the revision was later corrected, expanded, or superseded.

Revision as a research object

Treating the revision itself as a research object requires more than depositing data. Biodiversity genomics already demonstrates that data stewardship operates across an entire project lifecycle, from specimen acquisition and data generation to reuse, governance, and preservation [6]. At the same time, semantic resources used in environmental research differ substantially in interoperability and documentation, showing that machine-readable availability alone does not guarantee that relationships among objects will be interpreted consistently [7]. A taxonomic research object therefore needs both identifiable components and explicit semantics for how those components are related.

Recent insect revisions illustrate why the object is necessarily compound. Revision of Platydracus integrates specimen-based morphology, molecular information, historical taxonomy, and formal changes rather than producing a single homogeneous datum [8]. Work on the Mediterranean Temnothorax rottenbergii group likewise combines morphological and molecular evidence with revised taxonomic treatment [9]. In Pseudoloxops, molecularly generated candidate groups were evaluated against morphometrics, genitalia, host association, and geography, making the logical sequence from initial hypothesis through corroborating or conflicting observations recoverable in principle [10].

These studies do not establish a universal digital-revision standard, and their taxonomic outcomes are specific to their sampled clades and materials. They nevertheless expose the inadequacy of representing a revision as an undifferentiated PDF-plus-supplement package. The proposed research-object view instead distinguishes at least four kinds of state: the evidence state available to the reviser, the analytical state produced by transformations of that evidence, the interpretive state assigning taxonomic meaning to analytical outputs, and the formal state created when an interpretation results in nomenclatural action. Those states may be related, but none should be assumed to be equivalent to the next.

Evidence provenance

Evidence provenance is needed because the evidential path from observation to species hypothesis is rarely one-dimensional. Integrative delimitation has shown that additional data modalities can reveal taxonomic structure that is not evident from another source alone [11]. The methodological implication is not that adding data necessarily produces a better species boundary; it is that the origin, type, and transformation of each evidence class should remain recoverable when an inference is evaluated.

The distinction is particularly important for model-based delimitation. Comparisons involving Drosophila show substantial but incomplete correspondence between coalescent-delimited entities and evidence of reproductive isolation, while geographically structured populations can be over-split under multispecies-coalescent analyses in some conditions [12, 13]. A coalescent entity should therefore be provenance-labelled as an analytical result generated under particular assumptions and sampling conditions, rather than silently converted into a biological or nomenclatural species assertion.

Empirical insect cases further show that disagreement and null evidence are informative. In Swiss Myrmica, most morphologically recognized species corresponded to genetically differentiated lineages, yet one nominal species lacked equivalent support and the examined common taxa did not reveal the predicted reservoir of additional cryptic species [14]. In the beet-leafhopper complex, molecular operational units received corroborating information from acoustic and morphometric evidence [15]. Conversely, phylogenomic and multi-character examination of Phloeosinus demonstrates that previously defined typological hypotheses can be tested rather than simply inherited as fixed classifications [16]. In Cheguevaria, conspicuous color polymorphism was not evidence of equivalent genetic species differentiation [17]. High-altitude anthocorid bugs additionally show that reticulate histories can complicate an inference based on a single genealogy [18].

These cases argue against an evidence hierarchy in which one technique automatically adjudicates every conflict. Instead, the proposed provenance model assigns each evidence item an identifiable origin and scope: specimen or observation source, sex or developmental stage when relevant, geographic and temporal context, measurement or sequence derivative, processing history, and the analytical operation through which it entered a taxonomic argument. Concordant evidence may support a hypothesis, but concordance is not recoded as proof. Discordant or null evidence is retained rather than deleted when the eventual decision favors another interpretation.

This distinction also prevents retrospectively rewriting the logic of a revision. If a molecular grouping is proposed first and later rejected after morphology, behavior, broader sampling, or genomic evidence is examined, the earlier grouping remains part of the revision history. Provenance therefore records not only what supported the accepted conclusion but what alternatives were genuinely tested. Such a record permits later researchers to ask whether a disagreement originated in biological conflict, specimen coverage, character definition, model assumptions, or an analytical transformation rather than treating every changed name as a simple replacement of one answer by another.

Specimen and type linkage

Any reproducible revision ultimately depends on physical evidence. Sample-tracking research outside taxonomy has documented how identity and attribution can fragment as physical material moves among collections, laboratories, derived datasets, and publications [19]. The principle transfers directly to systematic entomology: a sequence, image, geometric-morphometric landmark set, or extracted character state should remain connected to the voucher or type-associated material from which it was derived whenever such linkage is possible.

Historical specimens make this relationship especially consequential. UCE phylogenomics of long-horned bees demonstrates that even extremely degraded specimen DNA can contribute useful phylogenomic information under appropriate laboratory and analytical conditions [20]. That result does not make degraded and fresh DNA evidentially equivalent. It instead makes DNA condition, extraction history, library preparation, locus recovery, and specimen identity part of the provenance required to interpret the resulting genomic object.

Linkage can also resolve otherwise difficult biological associations. High-throughput sequence data were used to associate and describe the female of Calicnemia haksik, illustrating how molecular information may connect a phenotype or sex to an existing species hypothesis [21]. Such an association is scientifically stronger when the relationship among the sequenced individual, its voucher, the comparison material, and the inferred identity remains machine-recoverable rather than being represented only by a sentence in the revision.

Specimen membership is also itself version-sensitive. Expanded geographic sampling of Appalachian Dasycerus exposed cryptic structure within the previous concept of D. carolinensis [22]. The appropriate generalization is not that additional sampling necessarily produces splitting. Rather, a taxonomic conclusion is conditional on the set of examined specimens. A reproducible revision should therefore preserve the membership and identity of that set so that a later investigator can determine whether a changed conclusion followed from new specimens, reinterpreted characters, different analyses, or some combination of these.

Type material requires an additional distinction. A type anchors nomenclatural application, whereas a voucher or sequenced specimen can contribute evidential information without acquiring the same nomenclatural role. The proposed architecture therefore links types, vouchers, and derivatives but does not flatten their functions. Where a genomic sequence cannot be generated from a name-bearing type, the evidential chain should make explicit how other specimens were associated with that taxonomic concept and where inferential distance from the type remains.

Versioning of character matrices and genomic data

Versioning becomes scientifically meaningful when a change alters the evidential state presented to an analysis. Phylogenetic work on Nomadinae combined UCE-rich samples with taxa represented primarily by barcode data, demonstrating that heterogeneous sequence completeness can expand taxonomic coverage while leaving shallow inferences sensitive to the amount and kind of information available [23]. A matrix identifier should therefore resolve not merely to "the data," but to a particular taxon-by-character or taxon-by-locus state.

The same principle applies when data classes are integrated. UCEs, COI barcodes, and morphology jointly informed analysis of the Hypoponera pruinosa group [24]. Their integration does not erase their different origins or inferential properties. A revised morphology coding, a replaced sequence, an expanded UCE matrix, and a newly added taxon are distinct changes even when all ultimately feed a single combined analysis.

Genome-scale evidence can also change the interpretation of an earlier molecular signal. In selected alpine Lepidoptera, genomic data were used to evaluate substantial allopatric COI divergence and informed revision of the Symmoca signella complex [25]. Such a history should not be represented as if the genome-wide inference had existed at the initial barcode stage. The barcode-defined hypothesis is a prior version of the evidential argument; later genomic evidence may support, modify, or reject it.

Even apparently technical preprocessing decisions can be consequential. Large-scale testing of barcode length shows that longer sequences are not uniformly superior for species discovery or identification [26]. Sequence trimming, locus inclusion, alignment, masking, filtering, missing-data treatment, and character recoding therefore belong to provenance when they change the input on which inference depends. No universal threshold follows from this observation. What matters for reproducibility is that the transformation is identifiable and that the transformed object can be distinguished from its predecessor.

We therefore propose three nonexclusive reasons for declaring a new taxonomic-data version: membership change, when specimens, taxa, characters, or loci enter or leave the dataset; state change, when an observation, coding, sequence, alignment, or annotation is corrected or replaced; and transformation change, when processing or filtering creates a substantively different analytical input. Minor metadata repair can be recorded without implying a biological change, whereas modifications that alter inference-relevant content should generate a citable version relationship. **Table 1** separates these situations without treating them as a numerical hierarchy.

Table 1. Distinguishing inference-relevant versions of taxonomic character matrices and genomic datasets

Versioning state or scenario — evidential organization		**Interpretive reach, restraint, and verification**	
Taxonomic context with observed basis	**Versioning state with evidential standing**	**Interpretation within its limit**	**Discriminating check**
Revision expanded to newly examined material or geographic coverage **Added, removed, or reassigned specimens/taxa**	Specimen- or taxon-membership change Evidence-supported distinction; version rule proposed	The evidential universe supporting the revision has changed Additional sampling may reinforce, merge, split, or leave hypotheses unchanged	Re-run analyses whose inference depends on taxon or specimen composition
Morphological or other coded characters are corrected or reinterpreted **Changed coding traceable to a specimen, stage, sex, image, or observation**	Character-state revision Proposed version rule	Character support for competing hypotheses may change Reinterpretation may reflect observer, definition, preservation, or stage differences	Retain both previous and revised coding with reason for change
A sequence is corrected, regenerated, or substituted; locus coverage changes **New sequence derivative, quality information, or locus membership**	Sequence or locus replacement Evidence-supported distinction; version rule proposed	Molecular support may differ without any change in the physical specimen Data quality and completeness remain method dependent	Preserve specimen-to-sequence identity and supersession relationship
Some taxa are UCE-rich while others are represented by reduced sequence classes **Unequal locus or barcode representation among taxa**	Mixed genomic completeness Evidence-supported	Placement may remain possible although shallow inference can become data-sensitive Sparse and dense representations are not evidentially interchangeable	Report matrix completeness at the version level

Initial COI divergence is later evaluated with genome-scale evidence **Earlier barcode grouping and later genomic dataset**	Barcode hypothesis followed by genomic reassessment Evidence-supported pattern; provenance response proposed	Species interpretation can be retained, modified, or rejected Allopatry and taxon-specific history constrain interpretation	Preserve the earlier hypothesis rather than overwriting it
Alignment, trimming, masking, filtering, locus selection, or barcode length changes **Different analytical derivative from the same underlying observations**	Processing-transformation change Evidence-supported premise; version rule proposed	Downstream inference may change because the input object changed Sensitivity is dataset and method specific	Cite transformation parameters and make derivative version-addressable
Identifier, annotation, or descriptive metadata is repaired without changing inference-relevant content **Administrative or descriptive update**	Metadata-only correction Proposed distinction	Scientific input need not be treated as biologically different A metadata correction can become substantive if it changes specimen identity or evidence interpretation	Record correction while distinguishing it from evidential version change

The same versioning logic is difficult to recover from a simple sequence of supplementary files because dependencies are nested and later evidence can feed back on earlier interpretation. **Figure 1** therefore represents the versioned dataset as a set of connected but separable evidence states, with explicit boundaries between physical evidence, derived data, analytical inputs, and taxonomic interpretation.

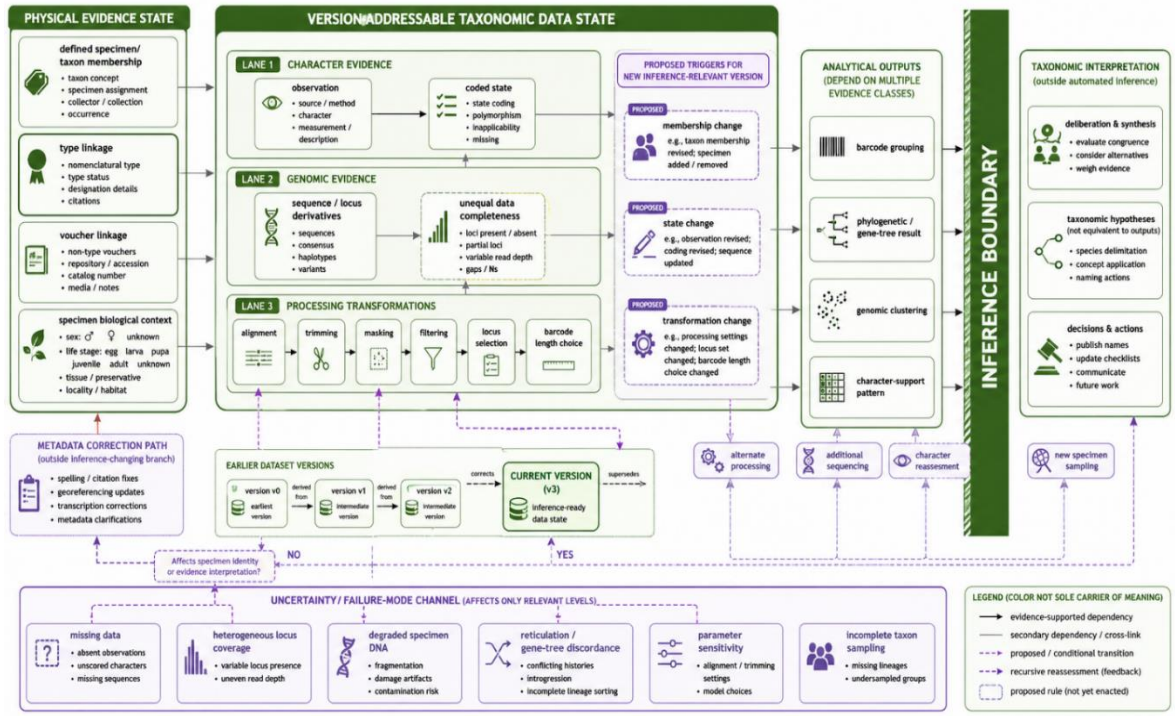


Figure 1. Where a Taxonomic Dataset Changes Identity: Nested Version States, Dependencies, and Inference Boundaries

Decision provenance

Data provenance answers where evidence came from; it does not by itself answer why a taxonomist accepted one interpretation over another. This distinction becomes more important as analytical procedures generate candidate groups automatically. Unsupervised machine learning applied to eurytomid parasitoids recovered plausible groupings and additional hypotheses, illustrating that computational classification can function as a discovery step without becoming a nomenclatural decision by itself [27]. An algorithmic cluster therefore requires provenance

for its input data and settings, but it also requires a separate record of what the reviser concluded that cluster meant.

Discordance makes such decision provenance indispensable. In Neotropical Nylanderia, UCE phylogenomics and male genital morphospace supported some species hypotheses while failing to support several groupings suggested by earlier COI evidence [28]. A conventional final revision can communicate the accepted result, but it may be much harder to reconstruct which hypothesis failed at which evidential comparison. The proposed decision record retains each candidate hypothesis, the evidence considered for and against it, the interpretation applied to conflicts, the resulting status, and any residual uncertainty. Rejection is therefore a recorded state rather than disappearance.

The evidence considered can also be biologically conditional. Comparative work on Anastatus and Mesocomys shows that reproductive systems and immature stages can reveal distinctions that are difficult to recover from superficially similar adult morphology [29]. Consequently, a decision record should not label an item simply as "morphological evidence" when its relevance depends on sex, developmental stage, anatomical system, or availability of particular material. Evidence that was unavailable for one taxon should likewise not be encoded as negative evidence.

We therefore distinguish four proposed decision-provenance stages: candidate hypothesis, identifying the taxonomic proposition being evaluated; evidence adjudication, recording which evidence classes support, contradict, fail to test, or remain ambiguous with respect to that hypothesis; interpretive decision, documenting why the available pattern was considered sufficient or insufficient for a taxonomic change; and decision state, identifying the hypothesis as retained, modified, rejected, unresolved, or superseded. These states do not constitute a universal species-delimitation algorithm. They are an audit structure for preserving reasoning across methods that may legitimately differ in evidential relevance among taxa.

Decision provenance also prevents a common category error in digital taxonomy: treating transparency as automatic adjudication. Recording every input does not determine whether morphology should outweigh a discordant genomic signal, whether reticulation explains conflicting trees, whether allopatric sequence divergence warrants species status, or whether an apparently diagnostic character is stable across sex, geography, or development. Those remain scientific judgments constrained by the organism, evidence, and species concept. The proposed contribution is narrower but consequential: when such a judgment is made, its evidential basis and unresolved alternatives should remain recoverable in the next version of the revision rather than being reconstructed retrospectively from the final name alone.

Nomenclatural acts

Nomenclatural acts occupy a distinct layer because they formalize consequences of taxonomic interpretation without being equivalent to the evidence that motivated it. Integrative revision of Cerurina and Afrocerura illustrates how type material, historical identifications, morphology, and molecular evidence can jointly lead to new descriptions and altered taxonomic status [30]. A reproducible revision should therefore connect each act to the immediately preceding taxonomic hypothesis and evidential state while preserving its separately governed nomenclatural status.

This distinction prevents two opposite errors. A nomenclatural act should not appear detached from the specimens and reasoning that motivated it, but provenance metadata should not be represented as establishing nomenclatural availability, priority, or validity. The proposed architecture consequently records an act as a downstream formal event with links to relevant types, prior names or concepts, decision state, and revision version, while leaving Code-governed nomenclatural requirements conceptually separate.

Machine-readable revision histories

A revision history becomes scientifically useful when it records dependencies rather than merely dates. Reproducible computational research requires metadata spanning inputs, software, analytical procedures, workflows, and outputs [31]. Applied to taxonomy, that principle means that a later user should be able to recover which specimen set, character matrix, genomic derivative, analytical configuration, and decision state belonged to a particular revision version.

We therefore propose that machine-readable histories preserve explicit relations such as derived from, corrects, supersedes, reinterprets, and retains with uncertainty. Earlier versions should remain citable rather than being silently overwritten. Persistent identifiers can stabilize references to those states [1], while digital extended-

specimen concepts provide a model for connecting distributed specimen-derived objects [3]. Semantic interoperability remains a constraint because identical labels do not guarantee identical meaning across infrastructures [7].

The central requirement is preservation of discordance. A machine-readable history should retain rejected hypotheses, unresolved conflicts, corrected specimen identities, and alternative analytical branches when they materially affected the decision path. **Table 2** distinguishes the revision states that require different provenance responses rather than treating every update as equivalent.

Table 2. Machine-Readable Revision States and the Provenance Responses They Require

Observable trigger	Interpretive consequence within its limit	Required provenance response
	Evidence added	
	Evidence-supported pattern; response proposed	
New specimens, characters, sequences, or loci enter the revision	Evidential basis expands New evidence may confirm or alter interpretation	Link additions to the preceding version and identify their origin
	Evidence corrected	
	Proposed provenance response	
Specimen identity, coding, sequence, or annotation changes	Earlier inference may require reassessment Not every correction is inference-relevant	Preserve previous state and record correction relation
	Hypothesis rejected	
	Proposed state	
Later evidence fails to support an earlier taxonomic proposition	Decision history remains reconstructable Rejection may be method- or sampling-dependent	Retain the rejected hypothesis and reason for rejection
	Hypothesis unresolved	
	Proposed state	
Evidence remains conflicting or insufficient	Uncertainty remains explicit Future evidence may alter status	Preserve competing interpretations without forced resolution
	Revision superseded	
	Proposed state	
A later revision replaces part or all of an earlier treatment	Historical taxonomic reasoning remains recoverable Supersession does not imply every earlier observation was wrong	Link versions through supersession rather than deletion

These states are difficult to communicate with a simple chronology because disagreement can branch, persist, and later re-enter the decision process**. Figure 2** therefore treats revision history as a discordance-triage architecture in which agreement, conflict, correction, and unresolved evidence produce different machine-readable responses.

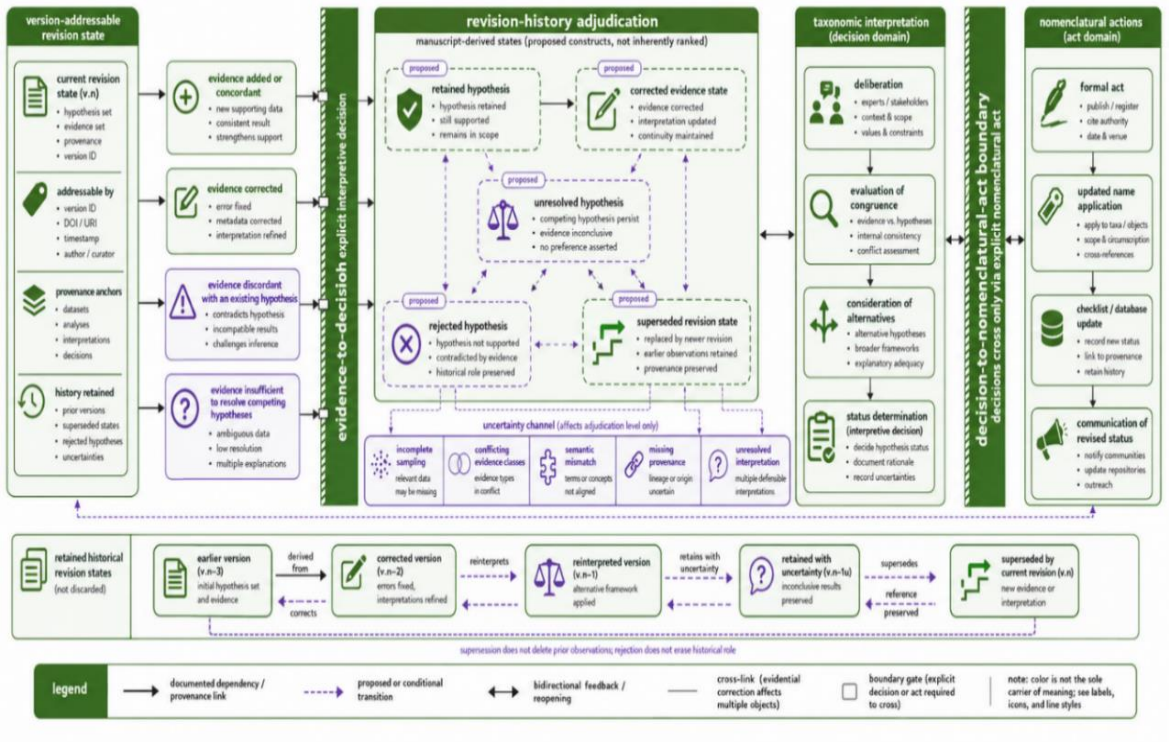


Figure 2. When a Revision Changes: Machine-Readable Triage of Agreement, Conflict, Correction, and Supersession

Reproducible digital-taxonomy workflow

A reproducible digital-taxonomy workflow must connect versioned computational procedures, registry-level history, and metadata across inputs and outputs [4, 5, 31]. In the proposed architecture, a revision version therefore resolves not only to its manuscript but also to the specimen universe, evidence derivatives, analytical configuration, decision record, and formal acts that belonged to that state.

Reproducibility here means reconstructability rather than guaranteed replication of a taxonomic conclusion. A later investigator may reproduce the same analytical route and reasonably interpret the evidence differently. The workflow should therefore preserve executable or reconstructable dependencies where feasible while keeping scientific adjudication explicit rather than embedding it invisibly in software or file replacement.

Implementation challenges

Retrospective conversion will be difficult where identifiers are unstable, semantic conventions differ, or specimen–derivative relationships were never recorded [1]. Semantic heterogeneity can persist even within ostensibly interoperable infrastructures [7], and biodiversity-genomics stewardship depends on institutional capacity and governance [6]. Physical-sample tracking can likewise fail when material changes repositories or identifiers across laboratory stages [19].

More fundamentally, provenance does not eliminate taxonomic uncertainty. Population structure can affect delimitation [13], reticulation can complicate genealogical interpretation [18], unequal data completeness can influence shallow inference [23], and analytical parameters can change identification or discovery performance [26]. The proposed system should expose these dependencies rather than claim to resolve them. Adoption will also require durable workflow maintenance [5] and compatibility with nomenclatural practice [30]. Comparative implementation across taxonomic groups, collection infrastructures, and revision scales is therefore needed before minimum conformance requirements or field-wide standards can be justified.

Conclusion

Taxonomic revisions are better represented as evolving scholarly objects than as terminal publications. Their reproducibility depends on preserving how specimens, evidence derivatives, analytical states, interpretations, decisions, and nomenclatural acts relate across versions. The architecture proposed here does not determine species boundaries or guarantee correct nomenclature. Its purpose is narrower: to make taxonomic reasoning

reconstructable, disagreement preservable, corrections explicit, and supersession auditable. Validation should now test whether these principles can be implemented without imposing disproportionate technical burdens across contrasting insect-systematics workflows.

Acknowledgments: None

Conflict of Interest: None

Financial Support: None

Ethics Statement: None

References

1. McMurry JA, Juty N, Blomberg N, Burdett T, Conlin T, Conte N, et al. Identifiers for the 21st century: How to design, provision, and reuse persistent identifiers to maximize utility and impact of life science data. PLoS Biol. 2017;15(6):e2001414. doi:10.1371/journal.pbio.2001414
2. Lendemer J, Thiers B, Monfils AK, Zaspel J, Ellwood ER, Bentley A, et al. The Extended Specimen Network: A strategy to enhance US biodiversity collections, promote research and education. Bioscience. 2020;70(1):23-30. doi:10.1093/biosci/biz140
3. Hardisty AR, Ellwood ER, Nelson G, Zimkus B, Buschbom J, Addink W, et al. Digital Extended Specimens: Enabling an extensible network of biodiversity data records as integrated digital objects on the Internet. Bioscience. 2022;72(10):978-87. doi:10.1093/biosci/biac060
4. de Visser C, Johansson LF, Kulkarni P, Mei H, Neerincx P, van der Velde KJ, et al. Ten quick tips for building FAIR workflows. PLoS Comput Biol. 2023;19(9):e1011369. doi:10.1371/journal.pcbi.1011369
5. Gustafsson OJR, Wilkinson SR, Bacall F, Soiland-Reyes S, Leo S, Pireddu L, et al. WorkflowHub: A registry for computational workflows. Sci Data. 2025;12(1):837. doi:10.1038/s41597-025-04786-3
6. Forsdick NJ, Wold J, Angelo A, Bissey F, Hart J, Head M, et al. Journeying towards best practice data management in biodiversity genomics. Mol Ecol Resour. 2025;25(2):e13880. doi:10.1111/1755-0998.13880
7. Di Muri C, Pulieri M, Raho D, Muresan AN, Tarallo A, Titocci J, et al. Assessing semantic interoperability in environmental sciences: Variety of approaches and semantic artefacts. Sci Data. 2024;11(1):1055. doi:10.1038/s41597-024-03669-3
8. Zhao QH, Hansen AK, Brunke A, Tang L, Solodovnikov A. Integrative taxonomic revision of the rove beetle genus Platydracus of the Chinese fauna (Coleoptera, Staphylinidae, Staphylininae). Insect Syst Divers. 2026;10(3):ixag003. doi:10.1093/isd/ixag003
9. Csősz S, Taheri A, Schifani E, Reyes-López JL, Alicata A, Báthori F, et al. Taxonomic revision of the Mediterranean Temnothorax rottenbergii species group via integrated morphological and molecular approaches. Insect Syst Divers. 2025;9(6):ixaf045. doi:10.1093/isd/ixaf045
10. Balukjian B, Van Dam MH. Using integrative taxonomy to delimit species in the plant bug genus Pseudoloxops (Hemiptera: Miridae) in French Polynesia. Insect Syst Divers. 2024;8(5):1. doi:10.1093/isd/ixae012
11. Chaplin K, Sumner J, Hipsley CA, Melville J. An integrative approach using phylogenomics and high-resolution X-ray computed tomography for species delimitation in cryptic taxa. Syst Biol. 2020;69(2):294-307. doi:10.1093/sysbio/syz048
12. Campillo LC, Barley AJ, Thomson RC. Model-based species delimitation: Are coalescent species reproductively isolated? Syst Biol. 2020;69(4):708-21. doi:10.1093/sysbio/syz072
13. Chambers EA, Hillis DM. The multispecies coalescent over-splits species in the case of geographically widespread taxa. Syst Biol. 2020;69(1):184-93. doi:10.1093/sysbio/syz042
14. Lavanchy G, Galkowski C, Jecha K, Freitag A, Avril A, Dépraz A, et al. Do morphologically defined groups correspond to reproductively isolated species? A case study in Myrmica ants from Switzerland. Insect Syst Divers. 2025;9(3):3. doi:10.1093/isd/ixaf021
15. Sinaiko G, Dietrich CH. Untangling the beet leafhopper (Hemiptera: Cicadellidae) species complex: A case study in integrative taxonomy. Insect Syst Divers. 2025;9(3):4. doi:10.1093/isd/ixaf017

16. Fernández-Campos AN, Ramírez-Reyes T, Cuellar-Rodríguez G, Atkinson TH, Armendáriz-Toledano F. Testing the classical typological taxonomy: An approach to evaluate the species hypotheses in Phloeosinus spp. (Coleoptera: Curculionidae), from the Mexican Transition Zone, based on phylogenomic data and multiple character sources. Insect Syst Divers. 2024;8(6):10. doi:10.1093/isd/ixae040
17. Ferreira VS, Keller O, Barbosa FF, Ivie MA. Integrative systematics of Cheguevaria Kazantsev, 2007 (Coleoptera, Lampyridae, Cheguevariinae) identifies genetic stability in color-polymorphic individuals and a disjoint relationship with Amydetinae. Insect Syst Divers. 2024;8(6):3. doi:10.1093/isd/ixae033
18. Tang ZC, Dong X, Zhu XX, Wang Y, Zhang DL, Ye Z, et al. Exploring species delimitation of high-altitude flower bugs (Heteroptera: Anthocoridae) in the presence of complex reticulate evolution: The case of a biodiversity hotspot in the Qinghai–Tibet Plateau and surrounding mountains. Insect Syst Divers. 2025;9(3):9. doi:10.1093/isd/ixaf019
19. Damerow JE, Raia NH, Stanley V, Choe S, Borton MA, Byers N, et al. Opening doors to physical sample tracking and attribution in Earth and environmental sciences. Sci Data. 2025;12(1):1047. doi:10.1038/s41597-025-05295-z
20. Freitas FV, Branstetter MG, Franceschini-Santos VH, Dorchin A, Wright KW, López-Uribe MM, et al. UCE phylogenomics, biogeography, and classification of long-horned bees (Hymenoptera: Apidae: Eucerini), with insights on using specimens with extremely degraded DNA. Insect Syst Divers. 2023;7(4):3. doi:10.1093/isd/ixad012
21. Wade CD, Sutherland LN, Lupiyaningdyah P, Bennett SP, Pham TH, Powell GS, et al. High throughput sequence data for association and description of female Calicnemia haksik (Odonata: Platycnemididae). Insect Syst Divers. 2025;9(3):2. doi:10.1093/isd/ixaf014
22. Caterino MS, Harden CW. A review of Appalachian Dasycerus Brongniart and the recognition of cryptic diversity within Dasycerus carolinensis (Coleoptera: Staphylinidae: Dasycerinae). Insect Syst Divers. 2024;8(2):5. doi:10.1093/isd/ixae009
23. Sless TJL, Odanaka KA, Alva-Caballero LX, Searle JB, Danforth BN, Rehan SM. Expanded phylogeny of Nomadinae (Hymenoptera: Apidae) with integration of UCE and DNA barcode sequence data. Insect Syst Divers. 2025;9(3):1. doi:10.1093/isd/ixaf011
24. Wang WY, Blaimer BB, Jaitrong W, Hosoishi S, Yamane S. Ultraconserved elements, DNA barcoding and morphology unravel the diversity and evolution of the Hypoponera pruinosa group (Hymenoptera: Formicidae) in Indochina and the Indo-Australian region. Insect Syst Divers. 2026;10(2):ixag013. doi:10.1093/isd/ixag013
25. Joshi M, Huemer P, Espeland M, Mutanen M. Genomics helps resolve allopatric COI (Cytochrome c oxidase subunit I) divergences in selected alpine Lepidoptera and allows for a taxonomic revision of the Symmoca signella species-complex (Autostichidae). Insect Syst Divers. 2026;10(2):ixag015. doi:10.1093/isd/ixag015
26. Yeo D, Srivathsan A, Meier R. Longer is not always better: Optimizing barcode length for large-scale species discovery and identification. Syst Biol. 2020;69(5):999-1015. doi:10.1093/sysbio/syaa014
27. Weinrich CL, Bubeníková K, Sheikh SI, Steffensen MJ, Ward AKG, Zhang YM, et al. Unsupervised machine learning for species discovery in Eurytoma and Phylloxeroxenus (Hymenoptera: Eurytomidae) parasitoids of oak gall wasps. Insect Syst Divers. 2026;10(3):ixag021. doi:10.1093/isd/ixag021
28. Williams JL, Punnath A, Fernández MB, Calcaterra LA, LaPolla JS, Lucky A. “Picking up signals” in male genital morphospace and integrating phylogenomics to delimit Neotropical Nylanderia Emery species (Hymenoptera: Formicidae). Insect Syst Divers. 2024;8(5):6. doi:10.1093/isd/ixae028
29. Chen YM, Qu XR, Shang CJ, Peng LF, Zang LS. Identification of morphologically similar genera based on reproductive systems and immature stages: A case study of Anastatus and Mesocomys (Hymenoptera: Eupelmidae). Insect Syst Divers. 2024;8(5):2. doi:10.1093/isd/ixae031
30. Mulvaney LRJ, St Laurent RA, László GM, Schintlmeister A. Integrative taxonomic revision of the Cerurina-Afrocerura genus-pair (Lepidoptera, Noctuoidea, Notodontidae, Cerurinae). Insect Syst Divers. 2024;8(5):5. doi:10.1093/isd/ixae029
31. Leipzig J, Nüst D, Hoyt CT, Ram K, Greenberg J. The role of metadata in reproducible computational research. Patterns (N Y). 2021;2(9):100322. doi:10.1016/j.patter.2021.100322