

## Article

# What Is a Taxon? Identity, Persistence, and Operability in Taxonomy

Thierry Bourgoïn <sup>1,\*</sup> , Nicolas Bailly <sup>2</sup> , René Zaragüeta <sup>1</sup> and Régine Vignes-Lebbe <sup>1</sup>

- <sup>1</sup> Institut Systématique, Évolution, Biodiversité (ISYEB), UMR 7205 MNHN-CNRS-Sorbonne Université-EPHE-University Antilles, Muséum National d'Histoire Naturelle, 75005 Paris, France; rene.zaragueta\_bagils@sorbonne-universite.fr (R.Z.); regine.vignes\_lebbe@sorbonne-universite.fr (R.V.-L.)
- <sup>2</sup> Beaty Biodiversity Museum—Department of Zoology, University of British Columbia, 2212 Main Mall, Vancouver, BC V6T 1Z4, Canada; n.bailly@q-quadrics.org
- \* Correspondence: thierry.bourgoïn@mnhn.fr

## Abstract

Although central reference units in biology, taxa remain difficult to integrate coherently across scientific and digital frameworks because they are repeatedly created, destroyed, or redefined through taxonomic revisions. To address this problem, we develop a conceptual and ontological analysis of taxonomic practice that distinguishes taxa from biological lineages, phylogenetic clades, names, and individual taxonomic treatments. We show that a taxon is a historically continuous object of knowledge instituted by a formal taxonomic act, whose identity is carried by a taxonomic temporal string, whose changing content is its taxonomic substance, and whose published interpretative history is its taxonomic trajectory. This framework is constrained by empirical properties of taxonomic practice, including persistence, historical continuity, irreversibility, revisability without destruction, traceability, and variable operability through time. Within this model, splitting, lumping, synonymy, and redefinition are interpreted as redistributions or reconfigurations of taxonomic substance rather than the creation or destruction of taxon identities. The framework also clarifies the articulation between taxonomy, phylogeny, and biological data by separating stable reference objects from the hypotheses that inform them. Because taxa originate only through unique code-compliant formal acts, they are intrinsically suitable for persistent identifiers (PIDs), enabling unambiguous and interoperable digital integration of systematic knowledge for biodiversity.

**Keywords:** taxonomy; taxon ontology; classification; temporal continuity; persistent identifier; interoperability; digital infrastructure



Academic Editor: Juan Manuel López-García

Received: 27 February 2026

Revised: 26 March 2026

Accepted: 27 March 2026

Published: 31 March 2026

**Copyright:** © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\) license](https://creativecommons.org/licenses/by/4.0/).

## 1. Introduction

Across all the life sciences, systematics provides the general framework for describing, organising and communicating about biological diversity, within which taxonomy constitutes its foundational conceptual infrastructure. Drawing on phylogeny, which informs the structuring of biological groups, taxonomy provides the reference units necessary for the organisation, comparison, and transmission of knowledge about biological diversity. Yet, in the era of digital infrastructures, taxonomy is confronted with a growing paradox: while it is indispensable as a reference framework, it remains difficult to make operational in contemporary information systems.

This paradox is mainly attributed to the presumed instability of taxa. Changes in circumscription, synonymies, splits, mergers, and successive reinterpretations are invoked

to support the idea that taxa are objects that are too dynamic, too dependent on biological hypotheses and available data, to be persistently referenced. This perception has encouraged the development of frameworks based on potential taxa, or contextual instantiations of the taxon, in which taxonomic entities are continuously reconstructed from data, to the detriment of their conceptual stability.

But is this diagnosis well founded? What if the current operational difficulties of taxonomy did not result from an intrinsic instability of taxa, but from an inadequate conception of their ontological nature? The taxon is indeed still frequently confused with its mere name, or treated either as a classical semantic concept defined by intension and extension, or as a biological entity directly deduced from the evolution of individuals into natural biological groups assumed to pre-exist their scientific recognition. This long-standing confusion between genealogical individuals and phylogenetic entities has been rigorously analysed by Zaragüeta (2011) [1], who showed that the notion of an identifiable ancestral individual is both inaccessible and conceptually irrelevant in systematics, and that only hierarchical taxa constitute coherent entities in phylogenetic reasoning. All these assimilations rely on intellectual shortcuts, already identified and denounced in Bourgoïn et al. (2021) [2], which lead to ontologically truncated and epistemologically sterilizing conceptions of the taxon.

Within these simplified frameworks, the taxon is implicitly reduced to a punctual state of knowledge, i.e., to whatever is currently asserted about it at a given moment: its extension (circumscription), its intension (diagnosis and rank), and its then-accepted name and usages. Such a reduction obscures a fundamental and empirically observable property of taxonomic systems: once a taxon has been formally established as a taxonomic unit, its existence is never called into question, even when its biological interpretation, content, circumscription, diagnosis, classificatory placement and accepted name undergo profound revisions. Taxa persist through time as reference entities, independently of the successive epistemological accesses and interpretations through which they are apprehended. This persistence, although implicit in taxonomic practice and consistently observed in taxonomic history, remains largely unformalised at the conceptual level.

We have recently shown (Bourgoïn et al., 2021) [2] that the taxonomic triptych constitutes the minimal operational framework that is indispensable, and practically sufficient, for handling a taxon as a data object in digital environments: enabling unambiguous identification, comparison of usages, database structuring, and interoperability. The present work completes its conceptual scope by making explicit the ontological foundation of this object, and by formally distinguishing the persistent identity of the taxon from its interpretative states.

In this article, we propose to requalify the taxon as a historically continuous entity of knowledge, which we conceive as a taxonomic temporal string, and which we describe, in taxonomic practice, as a taxonomic trajectory. A taxon is neither a fixed concept nor a direct biological entity, but a scientific object whose identity unfolds through time, from its initial act of formalisation to its successive revisions. Its definitions, circumscriptions, and usages constitute interpretative states situated along this trajectory, without ever exhausting or calling into question its existence. This reconceptualisation operates a decisive ontological shift: the taxon ceases to be an unstable semantic object and becomes a historical object of knowledge. As a direct consequence, the taxon acquires a fundamental property that has until now been largely denied: it is intrinsically eligible for persistent identification by persistent identifiers (PIDs). The attribution of a persistent identifier does not freeze the taxon; it anchors its continuous existence, while leaving intact the interpretative dynamics that characterizes taxonomic research.

We show that this conception makes it possible to coherently resolve the difficulties traditionally associated with splits, mergers, and synonymies, by interpreting them as redistributions of knowledge among taxonomic trajectories, rather than as creations or destructions of taxa. It also makes possible a rigorous articulation between taxonomy, phylogeny, and information systems, by clearly distinguishing the stability of reference objects from the variability of their interpretations.

The objective of this article is threefold. First, to clarify the ontological nature of the taxon as a taxonomic temporal scientific entity from an explicitly operational and digital perspective, without addressing the biological nature of species or providing a precise technical implementation protocol. Second, to demonstrate that this clarification constitutes the minimal conceptual foundation for the attribution of persistent identifiers to taxa. Third, consequently, to enable the full and definitive integration of taxonomic, phylogenetic and occurrence data within contemporary biodiversity information infrastructures, in line with FAIR principles.

## 2. Results

### 2.1. Clarification of Taxonomic Objects

- **Primary data in taxonomy.** Specimens, their occurrences, their molecular sequences, and more generally all biological data associated with them constitute the empirical foundation of primary data in evolutionary biology. These are material or observational objects from which hypotheses are formulated about biological diversity, its structures and its relationships. These data belong to a distinct descriptive and analytical register (primary data level). They support morphological, genetic, ecological or phylogenetic hypotheses, but do not in themselves constitute units of classification. Assimilating primary data (in particular specimens) to taxa amounts to shifting taxonomic reasoning to a different analytical level: a specimen may serve as evidence for a species hypothesis, but it is not, in itself, a species, nor, more generally, a taxonomic object.
- **Taxa and the taxonomic triplet.** Taxa belong to a register distinct from that of primary data: that of formalised units of knowledge. A taxon is neither a raw empirical object nor a simple linguistic label reduced to its name. It exists as an operational scientific object only when it is formalised according to a minimal operable structure (Bourgoin et al., 2019) [3], defined later as the taxonomic triplet (Bourgoin et al., 2021) [2]:
  - (1) an extension, defining the content of the taxon (subordinate taxa, specimens, and their associated biological characteristics) as documented by primary data;
  - (2) an intension, corresponding to the set of characters, relations or criteria allowing its diagnosis and its insertion into a classificatory hierarchy.
  - (3) a name established in accordance with applicable nomenclatural conventions and codes, constituting the only formally regulated component;

This triplet does not define a fixed essence, but a formal structure that makes the taxon manipulable as a scientific object. It therefore becomes essential to clearly distinguish: (1) the taxon as a formalised entity, (2) particular taxonomic usages (taxonomic usages, *secundum*), and (3) the supposed biological entity that these usages seek to apprehend. Confusion between these levels constitutes one of the major sources of conceptual instability in taxonomy.

- **Taxonomy and operational classification.** Extension, intension and name define the object and the scope of taxonomy, which takes charge of grouping, structuring and naming taxa. A taxon's extension corresponds to the set of entities it groups, together with the associated evidential record by which those entities are documented:

specimens, occurrences, and character observations (morphological, molecular, ethological, ecological. Its intension selects from this record the characters interpreted as informative for diagnosis and placement, thereby linking the taxon to a parent taxon. In this sense, intension grounds classification and hierarchy: assigning a parent taxon is an act of classification, and fitting that link into a coherent system yields an explicit hierarchy. Naming then provides the handle by which the taxon is addressed and communicated.

Taxonomy thus translates biological diversity into communicable, cumulative and comparable units of knowledge. The resulting classification is not a passive representation of the natural world, but an operational tool (operational classification), intended to structure information, facilitate scientific communication and allow the accumulation of knowledge.

A recurrent confusion lies in the abusive aggregation of identification and description, two conceptually distinct operations. Identification links an empirical object to a previously described taxon, whereas description contributes to defining, refining or revising that taxon within a hierarchical classification. Assimilating these two levels leads to attributing to data or to acts of identification a classificatory role that they cannot assume, and fuels faulty intellectual shortcuts (Bourgoin et al., 2021) [2].

Taxonomic classification moreover has its own constraints: minimal stability, internal coherence, readability and transmissibility. These constraints cannot be reduced to evolutionary relationships alone; they result from a methodological compromise between biological information, historical inheritance and operational efficiency. Classification is thus a product of taxonomy, not an automatic by-product of other disciplines; it is not a given framework, but the cumulative product of successive intensions.

- ***Phylogeny as a distinct register.*** Phylogeny, for its part, constitutes a distinct register of knowledge, focused on the hypothetical reconstruction of evolutionary relationships among clades. It relies on its own methods and aims to produce explanatory hypotheses about the history of life, generally in the form of dichotomous tree-like structures. Today, it represents a major source of information for taxonomy, but it is neither conceptually nor functionally equivalent to taxonomic classification.

A phylogenetic reconstruction does not in itself constitute an operational taxonomic classification, but a phylogenetic hypothesis expressed in classificatory form. It proposes a phylogenetic classification based on evolutionary relationships among lineages, but it does not in itself determine the rank of groups, their names, or the stability of their usage. Assimilating phylogeny and classification amounts to superimposing two distinct registers and imposing on taxonomy constraints that do not pertain to its purpose.

Contemporary taxonomic conflicts largely emerge from the confusion between primary data, taxa, classification and phylogeny. Clarifying the nature of the objects being handled and the domain of validity of each register constitutes an indispensable prerequisite for any coherent reflection on the stability, dynamics and operability of taxa.

## 2.2. *The Taxon: Neither a Concept nor a Biological Entity*

- ***The taxon cannot be reduced to a concept.*** In a classical sense derived from logic and analytic philosophy, a concept relies on stable definitions of intension and extension to ensure its identification and use. This definition assumes that a concept can be grasped at a given moment through a set of determining properties and a clearly circumscribed domain of application.

The taxon does not satisfy these conditions. Its intension and extension are neither fixed, nor closed, nor cumulatively stabilized at a given moment. They evolve through

revisions, changes in diagnosis, adjustments of circumscription and successive biological interpretations. A taxon therefore cannot be defined by a “complete” extension and intension without being immediately contradicted by its own historical dynamics.

Reducing the taxon to a single classical logical concept, defined synchronically by intension and extension, amounts to applying to it an inadequate framework, which artificially freezes what in fact pertains to a cumulative and historical process. This reduction mechanically leads to considering the taxon as unstable, incomplete or provisional, since it does not meet the criteria of stability expected of a logical concept. The taxon is not a “bad concept”; it is not a concept in the logical sense of the term.

- ***The taxon cannot be assimilated to a biological entity.*** Taxa are often treated as if they were pre-existing natural lineages that taxonomy merely discovers, names, and reflects. This assimilation is problematic not because biological groups lack reality, but because taxonomic existence cannot be reduced to biological continuity alone. Biological entities are embedded in evolutionary processes whose continuity, by itself, does not provide taxonomic boundaries, names, or classificatory status. By contrast, a taxon comes into being through a formal act of recognition, is named in accordance with applicable nomenclatural conventions and codes, and persists as a unit of reference independently of ongoing biological transformations and reinterpretations.

This distinction becomes especially clear in cases of rank change. Such changes must be interpreted at the level of the taxonomic object, not at the level of the underlying biological referent or of phylogeny alone. As already shown [2], taxonomic intension is inseparable from classificatory contextualisation: the defining properties of a taxon depend not only on its content, but also on its position within a classificatory framework. Rank is therefore not treated as a secondary or dispensable classificatory label, but as a constitutive dimension of the taxon itself, inseparable from its formal recognition and classificatory contextualisation. A taxon recognized at one rank is therefore not strictly the same taxonomic object as a taxon recognized at another rank, even when strong continuity remains in biological content, nomenclatural history, or usage. Thus, a subspecies and a species may retain important continuities while still constituting distinct taxon-level reference objects; the same logic applies to the elevation of any supraspecific taxon to a higher rank. What may remain continuous through such changes is not necessarily the taxon itself, but elements of its content, usage, or nomenclatural history. Changing rank modifies not only the nomenclatural expression of the taxon, but also its intensional definition.

The relationship between biological entities and taxa is therefore always mediated by hypotheses, interpretative choices, and classificatory conventions. Recognizing this mediation does not deny the existence of natural evolutionary groups, but acknowledges that taxonomy operates at a distinct level, where such groups are stabilized, named, and rendered operational as reference units. Everyday taxonomic practice often conflates continuity of the biological referent, continuity of nomenclatural history, and the identity of the taxon as a formally recognized object of knowledge. Yet assimilating the taxon directly to a biological entity amounts to denying this mediation and to conflating biological continuity with taxonomic identity. Such assimilations have been shown, within a cladistic framework, to generate logical inconsistencies, including infinite regress and systematic paraphyly [1]. This distinction is not only conceptually necessary, but also practically decisive if taxonomy is to enter the digital world as a fully interoperable data domain: the biological and informational registers must be explicitly separated if taxonomic data are to interact coherently with other knowledge corpora, whether within biodiversity science or in domains beyond biology, such as spatial planning, conservation policy, or environmental regulation [4,5].

- **Historical continuity and irreversibility of the taxon.** Unlike concepts and biological entities, taxa exhibit a distinctive property: irreversible historical continuity. Once validly established, a taxon definitively enters the taxonomic corpus. It may be reinterpreted, renamed, displaced, placed in synonymy or subdivided, but its past existence is never annulled.

This irreversibility is an empirical fact of taxonomic practice, although it has rarely been formalised conceptually. No mechanism exists to retroactively “erase” a taxon from the history of taxonomy. Even when a taxon is regarded as invalid, erroneous or redundant, it remains a referenced entity, cited and integrated into subsequent taxonomic reasoning, and may be resurrected.

This persistence radically distinguishes the taxon from logical concepts, which can be abandoned, and from biological entities, which can disappear. These considerations lead to a first, straightforward conclusion: the taxon cannot be understood either as a semantic concept or as a biological entity, but as a historical scientific object.

The taxon thus ceases to be an unstable semantic object (e.g., a reference entity) dependent on provisional definitions, and becomes, ontologically, a unit of knowledge endowed with its own continuity, independent of its successive interpretative states. As such, it must necessarily be recognized as bearing the following fundamental properties, as they are mobilized in taxonomic practice:

1. **Identity persistence.** Once formalised by a valid taxonomic act, a taxon retains its identity throughout scientific time through successive revisions of its content, usage, and interpretation. This persistence does not apply, however, when a subsequent taxonomic act institutes a distinct taxonomic object, notably through a change of rank.
2. **Historical continuity.** A taxon is embedded in a cumulative history that begins with its act of formalisation and extends through all subsequent reinterpretations, without being recreated at each revision.
3. **Irreversibility.** No mechanism exists to retroactively erase a taxon from the history of taxonomy: even when judged inadequate or redundant, it remains a referenced entity that can be mobilized in later reasoning.
4. **Revisability without destruction.** The extension and intension of a taxon (the latter including its classificatory position) may be profoundly revised without calling into question its existence as a taxonomic object. A change of rank, by contrast, constitutes a distinct taxonomic act resulting in the institution of a new taxon, while the original taxon persists.
5. **Persistence across biological reinterpretation.** The biological hypotheses (morphological, genetic, ecological, phylogenetic) used to support, interpret, or place a taxon may be revised, replaced, or abandoned, without the taxon ceasing to exist as a taxonomic reference object.
6. **Traceability and operability.** The uses, redefinitions and transformations of a taxon, as well as its time-specific operability, are intrinsically traceable through scientific time, as they can be dated, attributed to authors, and linked to a single reference object.

Taken together, these properties are incompatible with any strictly synchronic conception of the taxon. They require it to be conceived not as a punctual state of knowledge, but as a scientific object embedded in time.

### 2.3. The Taxon as a Temporal Trajectory: Ontological Reconstruction

Zaragüeta (2011) [1] established that taxa persist as hierarchical entities independently of ancestral individuals; the present framework extends this result by explicitly requalifying this persistence as a temporal property, carried by a historically continuous reference entity.

Before developing this ontological reconstruction, it is therefore necessary to clarify the distinct registers involved and the meaning of the key terms used throughout this paper.

- **Biological lineage (BE-lineage):** a lineage is a non-formal grouping of biological entities identified within a phylogenetic framework as forming, in a first analytical approximation, a coherent evolutionary unit. Its content remains provisional and potentially revisable, as neither its monophyly nor its exact limits are assumed to be definitively established. Paraphyly is explicitly tolerated at this stage, whereas polyphyly is excluded by definition. A biological lineage has no taxonomic status, no rank, and no nomenclatural standing; it is used to organize and track hypotheses, and to guide subsequent taxonomic decisions. In some recent taxonomic revisions, such BE-lineages have been explicitly recognised to document evolutionary structure while deliberately postponing formal taxonomic acts, as a methodological choice aimed at disentangling biological inference from taxonomic formalisation [6–9].
- **Clade:** a clade is a phylogenetic construct corresponding to a set of biological entities inferred, on the basis of explicit characters and models, to descend from a common ancestor within a reconstructed phylogeny. A clade is defined by its position in a phylogenetic hypothesis (i.e., a node and its descendants), not by a name, a rank, or a taxonomic act. As such, a clade is hypothesis-dependent, model-dependent, and revisable; it has no intrinsic nomenclatural or taxonomic status and does not, by itself, constitute a taxon.
- **Taxon:** a taxon is a formally instituted taxonomic reference object, established through an explicit taxonomic act that assigns a name, a rank, and a position within a classification. Such founding acts are defined by the relevant nomenclatural codes (ICZN, ICN, ICNP), whose procedural differences govern the code-compliant formalisation of taxa and the standing of their names, but operate upstream of the persistent taxon identity considered here.

A taxon does not correspond directly to a biological lineage or to a phylogenetic construct; rather, it constitutes a historically continuous reference entity through which biological lineages and phylogenetic hypotheses are interpreted, stabilised, and communicated. Once instituted, a taxon persists as an object of reference independently of changes in its taxonomic substance (e.g., circumscription, diagnosis, accepted name, and associated taxonomic information) and independently of shifts in phylogenetic interpretation.

In this framework, objects inferred in phylogenetic analyses are treated as clades, not as taxa. Throughout, we use taxon strictly in this restricted sense as a historical object of knowledge.

- **New definitions.** In our framework, a taxon is treated as a historically continuous object of knowledge. Its persistent identity is a taxonomic temporal string; its changing content through scientific time is its taxonomic substance; and the documented sequence of taxonomic treatments through which this content is expressed constitutes its taxonomic trajectory. The taxon is no longer a punctual state of knowledge, but becomes a scientific object whose identity unfolds through scientific time from its initial act of formalisation.

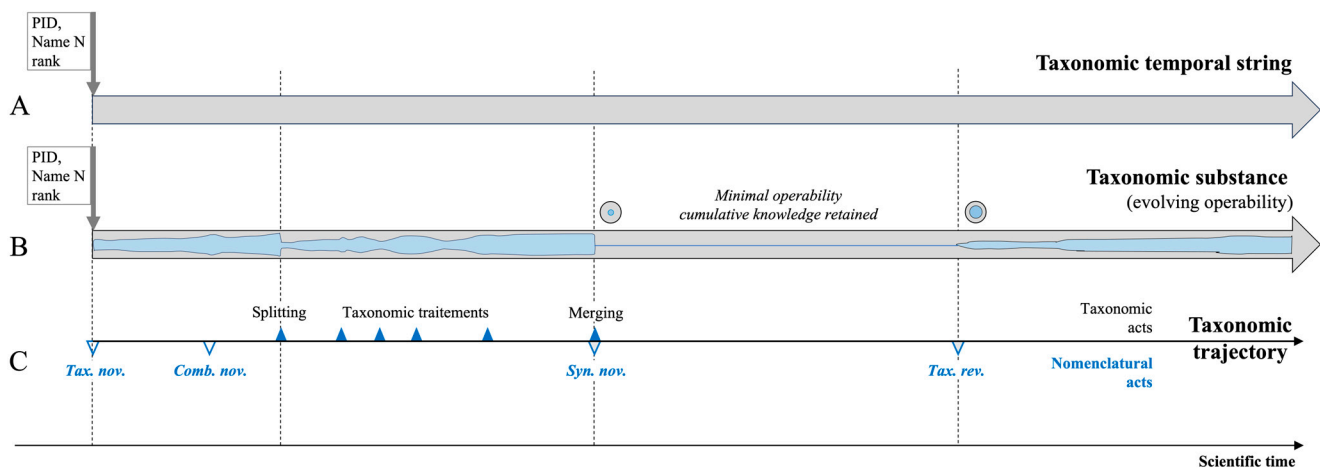
This approach imposes a strict conceptual distinction between the identity of the taxon and the body of knowledge associated with it over time. We propose the following definitions:

**Definition 1.** (Taxonomic temporal string). *A taxonomic temporal string is an ontological reference object: a historically continuous object of knowledge, endowed with a formal origin (the founding act of taxon formalisation) and a persistent identity. It underpins the taxon independently of the variability of its successive interpretations.*

**Definition 2.** (Taxonomic substance). *It is the corpus of taxonomic information associated with a taxonomic temporal string. It includes extension-level evidence (specimens, occurrences, and observational record), intension-level content (descriptive and diagnostic statements for delimitation and placement, including rank and classificatory position), and name-level, code-governed information (accepted name[s], synonymic relations, and other relevant nomenclatural acts and statuses).*

**Definition 3.** (Taxonomic trajectory). *The taxonomic trajectory is the ordered, dated, and attributed sequence of all published taxonomic treatments associated with a given taxonomic temporal string through scientific time.*

**Definition 4.** (Taxon). *A taxon is a historically continuous object of knowledge that comes into being through its formal establishment and persists through scientific time. As an ontological reference object, it is a taxonomic temporal string; its changing content is documented as taxonomic substance, and its interpretative history is documented as a taxonomic trajectory. Figure 1 provides a graphical summary of these definitions.*



**Figure 1.** Graphical summary of the new ontological model of the taxon. (A), the taxonomic temporal string (identity) persists through time. (B), the taxonomic substance is cumulatively retained as a historical record (grey), while its time-specific operability varies through time (blue). (C), the taxonomic trajectory corresponds to the dated sequence of published taxonomic treatments. Dotted circles indicate time intervals in which the temporal string persists but the associated substance has minimal operability, while remaining reactivatable by later treatments.

As a historical record, the taxonomic substance is cumulative: it is enriched by successive treatments and retains superseded, conflicting, or refuted states as part of the traceable record. Its organisation is nonetheless revisable: components can be reconfigured, and operative content may be redistributed between strings in splitting or merging, without affecting the existence or identity of the temporal string(s) involved. At any given time, taxonomic substance exhibits a level of operability (Figure 1B), i.e., the degree to which it supports effective and workable taxonomic communication. Likewise, successive treatments along a trajectory may be mutually inconsistent or coexist as alternative usages without implying any split in the underlying temporal string.

Accordingly, taxon stability does not lie in the fixity of substance, but in the continuity of the temporal string that carries identity. Taxonomic dynamics concern the redistribution and reinterpretation of knowledge, not the creation or destruction of taxonomic objects themselves. Confusing string and substance amounts to identifying the taxon with one of its interpretative states, and artificially produces ontological instability; conversely, recognising their dissociation makes it possible to understand how a taxon can be stable in identity yet dynamic in content.

- ***Suprafamilial taxa and nomenclatural regulation.*** The framework proposed here applies to all formally instituted taxa, independently of rank. The degree of nomenclatural regulation of taxa varies across Codes. In botanical nomenclature [10], names are regulated at all ranks, including suprafamilial ones, whereas in zoology [11] regulation formally stops at the superfamily level, and higher ranks rely on usage and convention. In prokaryotic nomenclature, fixed lists further constrain taxon availability [12]. These differences affect the normative conditions of taxon institution, but not the ontological status of taxa as historically continuous reference objects within taxonomic practice. Their taxonomic temporal string originates at the act of formal recognition, even when this act is not governed by the same degree of nomenclatural constraint as species- or lower familial-rank taxa in zoology.

Nomenclatural Codes thus do not create the ontological persistence of taxa, but provide standardized and socially enforced mechanisms for instituting it. Where Codes apply, they delimit the conditions under which a taxon can be formally instituted in accordance with the applicable Code and the correctly formed name under which it is recognized, but not the empirical or classificatory grounds on which it is established; where they do not, the persistence of the taxon still results from the same empirical properties of taxonomic practice: public institution, traceability, cumulative use and irreversibility. In zoology, suprafamilial taxa therefore fully participate in the fabric of taxonomic temporal strings, even when their formal anchoring relies on convention rather than codified rules.

#### *2.4. The Taxonomic Space as a Fabric of Taxonomic Temporal Strings*

A direct consequence of this distinction is a profound reconfiguration of taxonomic space. It should no longer be conceived as an indefinite set of potential taxa instantiated at each new state of knowledge, but as a structured fabric of distinct taxonomic temporal strings, each corresponding to a formally established and historically continuous taxon.

Within this framework, taxonomic space becomes conceptually finite and addressable. The number of strings corresponds to the number of taxa formally instituted under the rules of nomenclatural codes, while taxonomic dynamics are expressed through the transformation, redistribution and interrelation of their respective substances. Taxonomic revisions do not add new reference objects at each step; they modify the distribution of knowledge along already existing strings or, in the specific case of splitting, give rise to a newly formalised string.

This structuring radically transforms the operability of taxonomy. Instead of manipulating a multitude of unstable objects dependent on local states of knowledge, taxonomic systems can then rely on an explicit and stable set of reference objects (taxonomic temporal strings) to which dynamic properties are attached. Taxonomic complexity is not eliminated, but displaced: it no longer concerns the identity of objects, but the traceable and cumulative management of their states.

Thus understood, taxonomy ceases to be a conceptually open and operationally fragile space. It becomes a domain structured around persistent objects that are clearly identifiable, manipulable and interoperable—conditions necessary for its coherent integration into contemporary digital infrastructures.

#### *2.5. Epistemological Autonomy of Taxonomy as a Science of Taxa*

The reconceptualisation of the taxon as a historically continuous unit of knowledge entails a direct epistemological consequence: taxonomy can no longer be understood either as a simple descriptive auxiliary reduced to the inventory of groups of organisms, or as a technical by-product of the mechanical application of phylogenetic results. It constitutes an autonomous scientific discipline, defined by its own object, the taxon, and by the

specific operations it implements to formalise them, relate them and ensure their continuity through time.

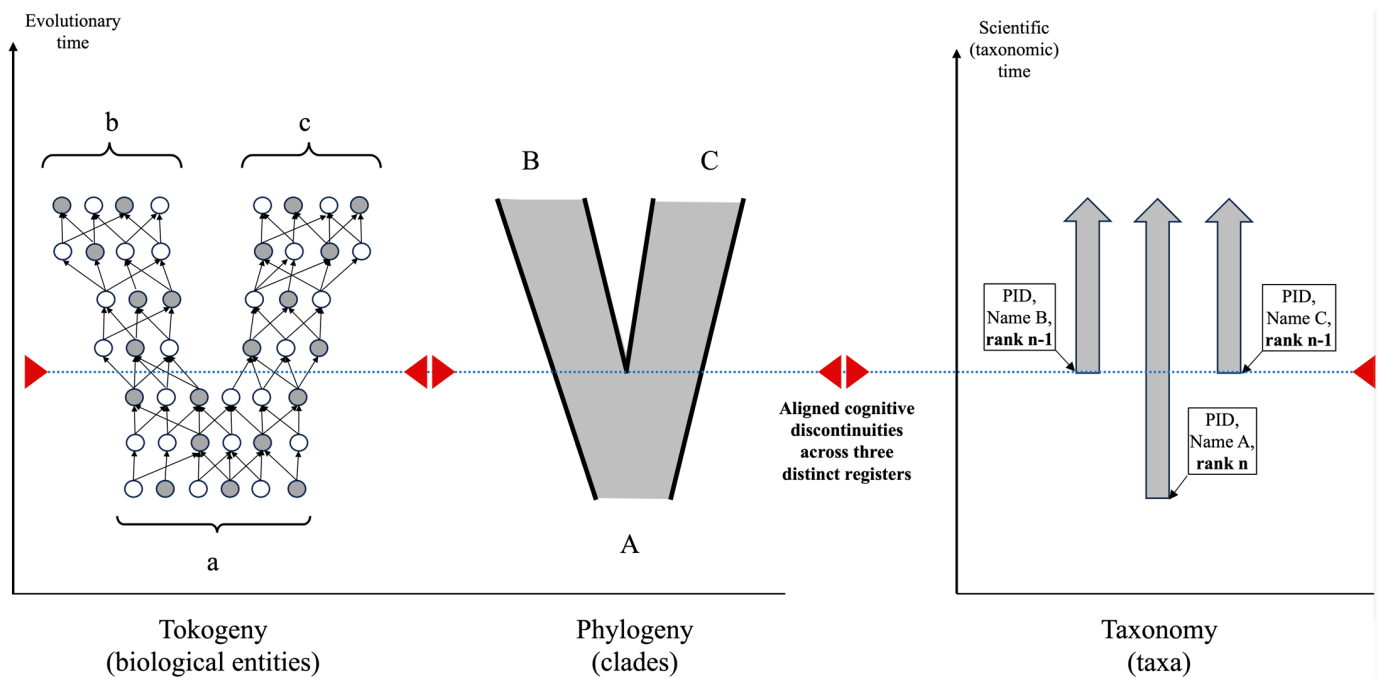
A different conception, however, widely disseminated, has contributed to obscure this status. It is based on a continuous and atomistic reading of classification, in which higher clades, species, populations, UTO, to even individuals, are conceived as units of a single continuum, simply nested at different scales. In this perspective, the biological relationships reconstructed by tokogeny and phylogeny are assumed to extend directly into taxonomic hierarchy, as if a single ontological continuum linked organisms, clades and taxa.

This reading relies on a doubly erroneous conceptual shift, partly originating from a simplified interpretation of Hennig's foundational figure [13]. His figure in fact allows a double reading. At a first level, it represents tokogenetic relationships between individuals, relating to biological transmission and material continuity. At a second level, the phylogenetic "Y" is superimposed on these individual relationships to reconstruct historical relationships between lineages. These two levels nevertheless belong to distinct registers: phylogeny does not extend tokogeny; it abstracts from it and conceptually superimposes itself upon it. Above all, neither tokogeny nor phylogeny produces taxa in themselves. The biological entities they describe are not taxonomic units. They have neither names, nor ranks, nor operative status as cumulative reference objects. They belong to the biological register, not to the taxonomic register.

A double confusion thus arises: on the one hand, from the erroneous assumption of an ontological continuity between tokogeny and phylogeny, although they belong to distinct and independent levels of reading; on the other hand, from the mistaken assimilation of reconstructed biological entities to taxa. Such an assimilation erases the register break between biology and taxonomy and reduces the latter to a simple classificatory projection of phylogenetic results, at the cost of losing the taxonomic status of units, including rank as a constitutive dimension of the taxon. The necessity of distinguishing tokogenetic, phylogenetic and taxonomic registers has been clearly established in cladistic theory, notably by Zaragüeta (2011) [1], who showed that phylogenetic reconstruction does not, by itself, produce taxonomic entities.

The framework proposed here clarifies this conceptual rupture by recognizing the constitution of a taxon as a distinct scientific act and the taxon itself as a named, hierarchized, historically constituted and persistent unit of knowledge, distinct both from biological individuals and from the punctual hypotheses mobilized to interpret them (Figure 2). Tokogeny and phylogeny provide indispensable biological hypotheses on the structure and history of life, but they do not, by themselves, result in taxa. Phylogeny informs taxonomy, but does not replace it. Recognizing, maintaining, subdividing or relating taxa relies on a specifically taxonomic reasoning, irreducible to automated phylogenetic inference. These registers must therefore remain explicitly distinct: a name may be available or unavailable under the Code; a biological or phylogenetic entity may be inferred or supported as a hypothesis; a taxonomic object may be operative or non-operative within an information system; and a taxon or its name may be accepted or not accepted in taxonomic practice.

Thus understood, taxonomy is a cumulative and operative science, whose purpose is the construction and management of a stable reference space for all life sciences. The persistence of its reference objects through scientific time does not contradict the revisability of their interpretations, but rather constitutes the condition for cumulative knowledge. Recognizing its epistemological autonomy is not an ideological stance, but the necessary consequence of a rigorous definition of its object, its practices and its role as a common foundation of biological sciences.



**Figure 2.** Conceptual separation and articulation of biological, phylogenetic and taxonomic registers. Tokogeny describes the relationships between unitary biological entities (a, b, c: individuals, colonies, ...) through evolutionary time; phylogeny reconstructs hypotheses of relationships among clades (A, B, C) within the same temporal framework. At the biological level, open and filled circles represent different character states of the unitary entities (respectively relatively plesiomorphic and relatively apomorphic states) (modified from Hennig, 1966 [13]). Taxonomy operates in scientific (taxonomic) time, instituting taxa as historically continuous reference objects through formal acts and associated identifiers (PIDs). The articulation of aligned cognitive discontinuities across three distinct registers underpins systematic reasoning without implying ontological continuity.

### 2.6. Toward an Operational Shift of Taxonomy in the Digital World: Taxa as Intrinsically PID-Able Objects

Taxonomy constitutes the common conceptual foundation of all life sciences. Every biological datum is ultimately attached to taxa, which serve as reference units to name, compare, organize and discuss biological diversity.

Paradoxically, this central discipline remains today structurally poorly integrated into contemporary digital infrastructures. Although taxa are omnipresent in scientific usage, they lack a stable and explicit referential anchor. Taxonomy thus evolves in a form of digital limbo, fragmented among names, contextual or local taxonomic concepts, competing usages and heterogeneous interpretations, without a clearly identified and persistent reference object, and within classificatory frameworks that are difficult to compare. Taxonomic revisions, although constitutive of the discipline, then appear as ruptures rather than as stages in a structured accumulation of knowledge. Taxonomy therefore struggles to fully fulfil its role as a cumulative and transversal reference system for the other biological sciences.

This situation does not result only from a technical or institutional failure, but from a shared conceptual lock, which affects even the major international biodiversity infrastructures. The Global Biodiversity Information Facility (GBIF) efficiently collects and links occurrences, but remains dependent on external taxonomies and pragmatic backbones due to the absence of taxa as persistent reference objects. The Catalogue of Life (COL) aims to establish a global reference list, but is constrained to reduce taxa to sets of stabilized names, limiting the integration of competing classifications from external Global Species Databases (GSDs), which it cannot manage simultaneously. ZooBank, finally, occupies a central posi-

tion as a nomenclatural registry, but remains focused on names and acts, precisely because the taxonomic object itself has never been formally defined as a persistent entity.

These limitations are in fact convergent manifestations of a single blind spot: the absence of a taxon explicitly defined as a historically continuous unit of knowledge, distinct from names, usages and biological interpretations.

The digital fragility of taxonomy thus does not result from an intrinsic instability of taxa, but from a persistent confusion about the very nature of the taxonomic object, too often assimilated to its punctual definitions, variable circumscriptions or local usages. In the absence of a clear ontological definition of the taxon as a historically continuous unit of knowledge, any attempt at coherent digital anchoring necessarily remains artificial. The problem faced by taxonomy is therefore not technical in nature; it is fundamentally conceptual.

- ***The PID as a referential anchor and operator of interoperability.*** A persistent identifier (PID) is a unique, stable and resolvable digital identifier designed to unambiguously reference an object of knowledge through time. Unlike a local identifier, a name or a simple web address, a PID is designed to remain valid independently of the technical, institutional or organisational evolutions of the systems that implement it. In contemporary scientific infrastructures, PIDs play a fundamental role by making objects of knowledge explicitly identifiable, enabling stable reference to them, linking them together and integrating them into distributed information graphs.

For a publication, a DOI does not refer the scientific content of an article, nor its value, nor its interpretation; it simply allows the same article to be persistently referenced despite changes in its context, uses or metadata. By functional analogy with the DOI, a PID does not replace the object; it makes its persistent existence operable. It does not encode the content of the object it references, does not guarantee its scientific validity, completeness or conceptual stability. It does not arbitrate between competing interpretations, does not resolve controversies and does not confer any ontological or biological truth. A PID does not freeze an object; it anchors it.

This distinction is central for taxonomy. The apparent failure of the “PID-isation” of taxa does not result from an inadequacy of PIDs, but from a persistent confusion between the object to be identified and its interpretative states. As long as the taxonomic object is not clearly defined as a reference entity distinct from its punctual definitions, no persistent identifier can correctly fulfil its role. The PID does not encode knowledge; it renders its persistence operable.

- ***Taxa are PID-able.*** When defined as a historically continuous knowledge entity, a taxon intrinsically satisfies all the properties required for persistent identification:
  1. it has a unique identity, acquired at the moment of its formal establishment and independent of its punctual definitions;
  2. it exhibits temporal persistence, a necessary condition for any durable reference;
  3. it is independent of its interpretative states, since its descriptions, diagnoses and circumscriptions may evolve without affecting the object itself;
  4. it is historically traceable, its usages and transformations being datable, attributable and cumulative;
  5. it is referentiable and mobilisable as a transversal anchoring point for data, classifications and biological reasoning;
  6. finally, it retains referential stability despite revisions, a sine qua non condition for any persistent identification.

It therefore becomes clear that the properties required for the attribution of a persistent identifier are not imposed externally on taxonomy: they are already constitutive of the taxon, provided that it is correctly defined. This leads to a simple but decisive conclusion:

it is not the PID that confers stability on the taxon; it is the recognition of the taxon as a historically continuous knowledge entity that makes the PID possible and operable. The PID exploits the stability of the taxon; it does not transform its nature, but renders its intrinsic properties fully operational within digital infrastructures.

- **Practical implementation of taxonomic PIDs.** The attribution of a persistent identifier to a taxon is based on a simple principle: a PID is associated with an object as soon as that object acquires a public, irreversible and referentiable scientific existence. In the case of taxa, this event corresponds to the valid act of taxonomic formalisation, in which an initial circumscription (a content) and a classificatory intension (a definition placing it within a classification) and a name are simultaneously assigned; this act renders the taxon public and its existence irreversible as a unit of knowledge. This founding act constitutes the origin of the taxonomic trajectory of the taxon and marks the initial point of the taxonomic temporal string that carries its identity.

The taxonomic PID is assigned only once, at this founding moment. It persists through any subsequent revision of the taxon, regardless of changes in name, modifications of circumscription, or particular biological interpretations. These events correspond to successive interpretative states of the taxon that evolve, not to the creation of new objects.

Accordingly, the PID is neither versioned nor replaced, ensuring a clear separation between:

- the existence of the taxon, anchored by the PID at the moment of its establishment;
- the evolution of its taxonomic substance, documented through successive taxonomic treatments that are unambiguously attached to it, each dated, attributed and contextualised, and associated with their own secondary identifiers.

The role of the nomenclatural Code is central in this framework. By strictly framing the act of naming, it transforms a taxonomic proposal into a public and shared object that cannot be retrospectively erased from scientific history. The Code does not define the biological truth of a taxon; it makes its objectivation possible by instituting it as a taxonomic unit. Persistent identifiers (PIDs) operate at this same level. They do not create taxa nor define their trajectories, but mark their formal origin and make their intrinsic persistent existence operational. A taxonomic PID anchors the identity of the taxon as a reference object, not the biological hypotheses or classificatory relations associated with it. Consequently, relations such as splitting or merging belong to taxonomic substance and documented interpretative history, not to taxonomic identity itself. What persists is therefore not biological truth, but the instituted existence of the taxon as a historically constituted reference object.

In practice, implementing taxonomic PIDs therefore requires only a minimal and explicit workflow. A registry records each taxon at the moment of its valid formal establishment and assigns it a unique persistent identifier. The PID identifies the taxon as a persistent object, while all taxonomic treatments are attached to it as versioned interpretative states. No PID is created for names, diagnoses or circumscriptions taken in isolation. This minimal architecture is sufficient to ensure stability, traceability and interoperability. This simple separation between a persistent taxonomic object and its evolving interpretations is sufficient to make taxonomic PIDs immediately operational.

- **Taxa, PIDs and taxonomic activities.** Recognising the taxon as a historically continuous knowledge entity makes it possible to reread the whole set of classical taxonomic activities while deeply clarifying the meaning of taxonomic practice. These everyday operations of the taxonomist (describing, revising, naming, synonymising, splitting or merging) do not act on the existence of taxa, but on their taxonomic substance, that is, on the body of knowledge historically associated with a given entity.

Table 1 synthesises the correspondence between traditional nomenclatural and taxonomic acts and their ontological interpretation within the taxonomic temporal string framework. It explicitly identifies which operations institute distinct taxon-level reference objects, which reorganize the substance of existing ones, and which affect only names or usages without ontological consequences for taxon identity.

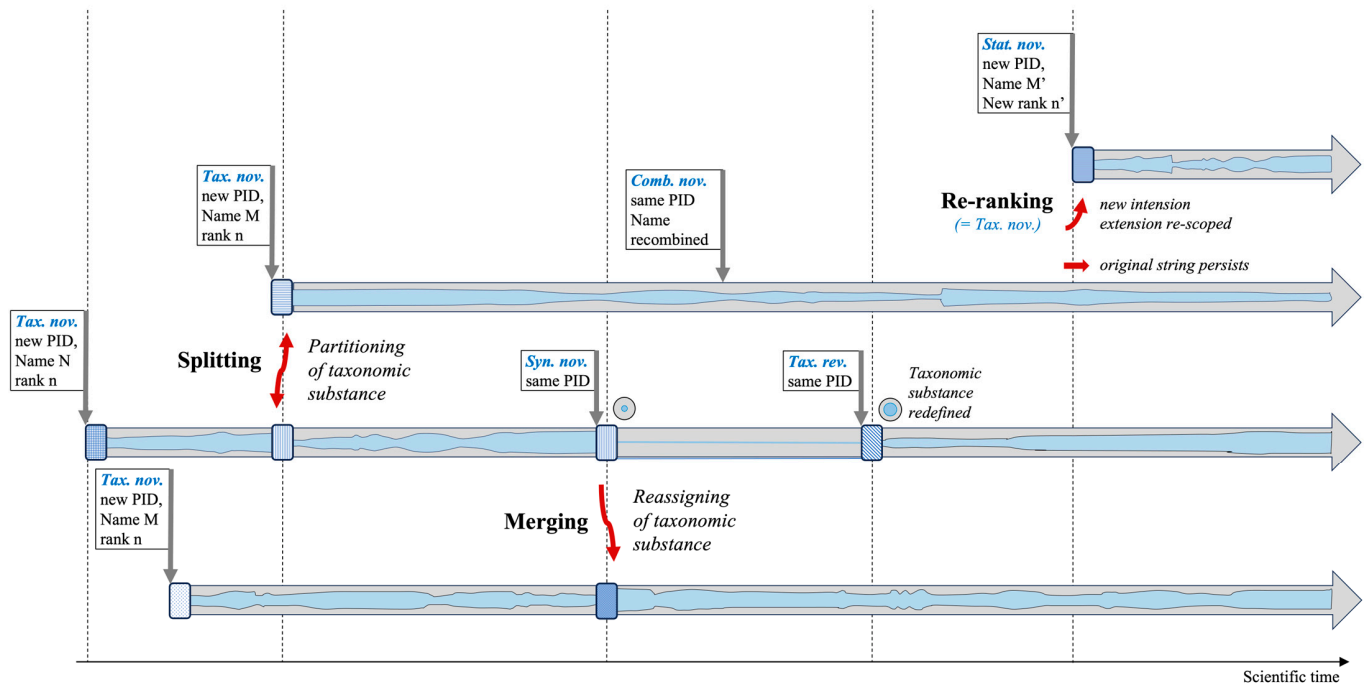
**Table 1.** Correspondence between traditional taxonomic terms and acts and their ontological interpretation. Within taxonomy, nomenclatural acts and name-status designations operate on names and their code-governed standing, whereas taxonomic acts operate on taxa by creating new taxon identities (taxonomic temporal strings) to which PIDs can be attached or by modifying, redistributing, and relating their taxonomic substance. In this framework, each taxon is a historically continuous object of knowledge (a taxonomic temporal string); its changing content through scientific time is its taxonomic substance, and the ordered record of published treatments that instantiate this content constitutes its taxonomic trajectory. Persistent identifiers (PIDs) attach to taxa (i.e., to taxonomic temporal strings) once a taxon has been formally established under a nomenclatural code; they do not identify names as such, nor individual treatments, but the taxonomic object instituted by the code-compliant founding act. Subsequent nomenclatural changes and interpretative revisions correspond to events or states along a trajectory, affecting substance and/or relations among strings. Some terms correspond to single actions with effects at both nomenclatural and taxonomic levels.

Taxonomic Designation or Act	Acting Domains	Ontological Effect on the Taxon	Interpretation Within the Taxonomic Temporal String Framework
<i>Taxon novum</i>	Taxon identity /substance	New string (+PID)	Creation of a new temporal string through a valid taxonomic act; attribution of a new PID; founding treatment as original trajectory event
Taxonomic revision ( <i>stat. nov./rev.</i> ) (no rank change)	Taxon identity /substance	Existing string (PID retained)	New instantiation or reactivation of the taxonomic substance of an existing temporal string; PID retained
Taxonomic revision ( <i>stat. nov./rev.</i> ) (with change of rank)	Taxon identity /substance	New string (+PID)	Creation of a distinct taxon-level reference object through recognition at a different rank; new temporal string and new PID; original string persists.
<i>Taxon inquirendum</i>	Taxon identity /substance	Existing string (PID retained)	Temporal string and PID valid; taxonomic substance insufficiently resolved
<i>Taxon incertae sedis</i>	Taxon identity /substance	Existing string (PID retained)	Temporal string and PID valid; classificatory position undetermined
Splitting	Taxon identity /substance	New string + original retained	Creation of a new temporal string and PID for the segregated taxon; original string and PID persist.
Merging (resulting from lumping)	Taxon identity /substance	Substance redistribution	Redistribution of taxonomic substance among existing temporal strings; strings and PID retained

Table 1. Cont.

Taxonomic Designation or Act	Acting Domains	Ontological Effect on the Taxon	Interpretation Within the Taxonomic Temporal String Framework
Synonymy ( <i>syn. nov.</i> )	Nomenclature and taxonomy	Name relation across strings	Nomenclatural act establishing a priority or usage relation between names attached to distinct temporal strings; may accompany a reinterpretation and redistribution of taxonomic substance; PIDs retained
Homonymy	Nomenclature and taxonomy	Name conflict across strings	Nomenclatural conflict involving identical names attached to distinct temporal strings; ontologically distinct taxa; PIDs retained
Basionym	Nomenclature (founding act)	New string (+PID)	Original name designation when a taxon is formally established; temporal string and PID fixed at this act.
<i>Combinatio nova</i>	Nomenclature (subsequent act)	Existing string (PID retained)	Temporal string and PID unchanged; may accompany a reinterpretation or reactivation of the taxonomic substance; Trajectory event: nomenclatural update
<i>Nomen dubium</i>	Nomenclature (status)	Existing string (PID retained)	Available name with uncertain application; temporal string and PID remain valid. Trajectory state: uncertain application; substance indeterminate.
<i>Nomen ambiguum</i>	Nomenclature (status)	Existing string (PID retained)	Nomenclatural ambiguity without ontological consequence; temporal string and PID unchanged
<i>Nomen nudum</i>	Nomenclature (status)	No string (no PID)	Absence of a valid founding act; neither temporal string nor PID exists. Name associated with another secondary name identifier without taxonomic status
<i>Nomen oblitum</i>	Nomenclature (status)	Existing string (PID retained)	Deactivation of an unused name; taxon retained; temporal string and PID unchanged

Within this framework, the attribution of a persistent identifier naturally integrates into the life cycle of the taxon. The PID is attached to the taxonomic temporal string at the founding act and remains invariant throughout all subsequent treatments. It allows usages, revisions, classifications and associated data to be unambiguously linked to the same taxon, or to different taxa, while fully respecting the intrinsic dynamics of taxonomic knowledge. The resulting dynamics of these operations (string persistence versus substance redistribution) are summarised in Figure 3.



**Figure 3.** The taxon as a temporal string and the dynamics of taxonomic substance. Taxa are represented as taxonomic temporal strings, that is, persistent taxon-level reference objects evolving through scientific time, while taxonomic substance changes in content and operability. Splitting (segregation) partitions taxonomic substance and institutes a new taxon (*Tax. nov.*): a new temporal string (taxon M) is created and receives a new PID, whereas the original string (taxon N) persists with modified operative substance. Merging (lumping) reorganizes, consolidates, or transfers components of taxonomic substance among already existing temporal strings (e.g., N and X), without suppressing their historical continuity. Nomenclatural acts modify name usage and code-governed status without altering taxon identity: a new combination (*Comb. nov.*) changes the accepted name while the taxon string retains the same PID, and synonymy (*Syn. nov.*) may accompany transfers or reinterpretations of substance but does not suppress temporal strings or their PIDs. Later redefinition (*Tax. rev.*) reshapes the substance of an existing string without recreating the taxon or assigning a new PID. Re-ranking (*Stat. nov.*) institutes a distinct taxon-level reference object at a different rank: a new temporal string (M') is created with a new PID and a rank-appropriate name, while the original taxon M persists historically. What may remain continuous through such a change is not necessarily the taxon itself, but the underlying biological referent, parts of its taxonomic substance, or its nomenclatural trajectory. Dotted circles mark periods in which the taxonomic temporal string persists with minimal operability while remaining reactivatable through subsequent taxonomic treatments.

In conclusion, the PID is not a technical solution artificially applied to a taxonomic problem. It is the formal, explicit and operational recognition of what the taxon has always been: a historically constituted knowledge entity, persistent in its identity and dynamic in its content, now fully integrated into the digital world.

### 3. Discussion

#### 3.1. Taxa as Reference Objects

In many contemporary practices, taxonomy remains affected by a persistent confusion between distinct registers: that of primary biological data (characters, sequences, occurrences, specimens) and the biological entities they allow to infer, and that of integrated, cumulative and historically constituted units of knowledge, namely taxa. This confusion

rests on a faulty assimilation of the taxon to a biological entity, implicitly treated as the endpoint of a continuous ontological chain:

population → lineage → clade → taxon

In contrast, the present framework explicitly recognises a register break:

population → lineage // clade (as a phylogenetic inference) // taxon

Biological information derived from morphological, molecular or etho-ecological analyses belong to the biological register and allows biological entities, such as BE-lineages, to be studied and characterised. By contrast, phylogenetic analyses operate in a distinct explanatory register and produce hypothesis-dependent constructs, aimed at recognizing clades and their evolutionary relationships. Both kinds of constructs may inform taxonomic decisions and contribute to taxonomic substance, but neither BE-lineages nor clades are taxa.

Such a confusion is exemplified in Pyle et al. (2021) [14,15], who conceptualise the taxon primarily through its names in line with the long-recognized limitations of names as identifiers in biological informatics [16], and an inferred circumscription aggregating biological data such as occurrences, characters, or population-level evolutionary lineages (e.g., [14] Figure 2). Within this framework, Taxonomic Name Usages (TNUs) are treated as competing representations of an underlying biological object. However, this approach overlooks two fundamental dimensions of the taxon. First, it neglects its intensional dimension, namely that a taxon is explicitly established to occupy a defined position within a hierarchical classification at a given rank. This is a constitutive feature of taxonomic practice rather than a secondary attribute: fully characterising a taxon requires recognition of the complete taxonomic triplet [2]. Second, and more crucially, it fails to account for the taxon as a historically continuous scientific object, whose identity persists through time by virtue of its “perdurant” nature.

The taxon belongs to another ontological order. It is a formalised unit of knowledge, constructed from this information but distinct from what it describes. This integrative construction results from an explicit taxonomic act, which assigns the taxon its own identity and historical continuity. The taxon is not produced by biological continuity; it is instituted as an object of knowledge on the basis of it.

In particular, the apparent “hinge” at the species-group level is not ontological but nomenclatural and classificatory: the obligation to refer a species to a genus in its name reflects a formal naming constraint, while also forcing the taxon to be placed in a classification. It therefore does not confer any privileged ontological status on the species level itself. Recognising that species-taxa are not themselves biological entities, but taxonomic representations of biological diversity, does not amount to denying the ontological reality of these species as biological entities. Rather, it rests on a necessary distinction between biological processes and the taxonomic register through which they are formalised, stabilised and rendered communicable. Species may well exist as real biological entities, according to various theoretical frameworks [17–19], . . .), however, the species-taxon constitutes a formally instituted reference object, operating within a distinct register of knowledge. Maintaining this distinction is not a theoretical luxury, but a prerequisite for conceptual clarity, cumulative reasoning and operational consistency in taxonomic practice.

By clearly distinguishing the temporal trajectory of the taxon from the continuous dynamics of the biological entities it encompasses, the taxonomic temporal string model makes it possible to fully integrate contemporary approaches without assigning them excessive ontological status. It thus avoids dissolving taxonomy into a continuum of

potential taxa, while preserving the richness and precision of the biological information on which it relies.

By disentangling biological continuity from taxonomic identity, this framework makes it possible to accommodate indefinitely expanding biological datasets without forcing taxonomic objects themselves to become provisional or unstable. Biological information can thus accumulate, diversify and even conflict, while remaining anchored to taxonomic entities that persist as stable reference points over time. This distinction becomes particularly consequential when taxonomic knowledge is translated into large-scale digital infrastructures.

Without this explicit separation of registers and the recognition of taxonomy as an autonomous domain producing its own reference objects, biological diversity cannot be coherently integrated into digital infrastructures. Biological continuity, by itself, does not yield addressable, persistent and interoperable objects; only formally instituted taxa fulfil this role. In the absence of such explicit anchoring, taxa remain digitally floating entities, trapped in a digital limbo, instantiated only locally through their names or contextual usages, and continuously recreated across datasets and workflows, the potential taxa described by Berendsohn (1995) [20]. By contrast, conceiving taxa as historically continuous reference objects makes it possible to delimit a finite, stable and technically achievable referential space (on the order of a few tens of millions of formally established taxa across all ranks), compatible with the requirements of persistent identification and large-scale interoperability. This shift directly conditions the possibility of integrating taxonomy into the contemporary digital ecosystem.

### 3.2. From Synchronic Concepts to Perduring Taxa

The framework developed here extends earlier logical and semiotic accounts of taxonomic reference by explicitly introducing a temporal dimension. Classical treatments of concepts and signs, from the Port-Royal logic [21] to the structural linguistics of de Saussure (1916) [22], rely on a fundamentally synchronic view, in which names, intensions and extensions (signifiers and signifieds) are analysed at a given moment, forming the minimal triplet required for the formal description of a taxon [2]. While essential for understanding the taxon as an object of knowledge, this framework remains insufficient to account for the cumulative and historically structured nature of taxonomic practice.

The conception of taxa developed here therefore relies on a perdurantist view of identity over time, as articulated in four-dimensionalist accounts of persistence [23,24]. Rather than enduring entities wholly present at each moment, taxa are understood as temporally extended reference objects, composed of successive temporal parts [24] corresponding to distinct taxonomic acts, usages and circumscriptions. Accordingly, taxon identity is not grounded in an invariant essence, but in the continuity of a historically documented trajectory.

This perdurantist interpretation does not commit to any metaphysical claim about biological entities themselves, but constitutes an operational stance on taxonomic reference objects. It provides an ontological grounding for the notion of taxonomic temporal strings and clarifies how taxa can remain identifiable despite changes in interpretation or scope, while allowing persistent identifiers to anchor their temporal continuity within cumulative digital infrastructures.

Beyond its internal coherence for taxonomic reasoning, this clarification of taxon persistence also has consequences for domains that rely on taxonomic stability as a normative reference. In particular, it helps to address the perceived instability that often hampers the integration of taxa into conservation and regulatory frameworks, such as the IUCN Red List or legal instruments for species protection [5,25]. By treating taxa as historically persistent reference objects, independent of changes in their biological interpretation, and

building on earlier clarifications regarding the non-individual nature of taxa as biological entities [1], this new conception of the taxon as a historically persistent reference object provides a stable point of reference that can be traced through time. When applied to such reference objects, persistent identifiers do not fix biological meaning but render this continuity operational [2,26]. This approach directly responds to concerns frequently raised under the label of “taxonomic anarchy”, by showing that taxonomic changes do not undermine the durability, traceability, or reliability of scientific references used in conservation policy and legal frameworks.

### 3.3. *PID-Able Taxa for Integration into the Contemporary Digital World*

Contemporary digital infrastructures rely on a simple and non-negotiable principle: in order to be integrated, connected and made interoperable, information must be attached to explicitly identified, persistent and referable objects. Publications, datasets, specimens, sequences or authors now meet this requirement through persistent identifiers, which ensure referential continuity independently of technical or conceptual changes. Operational guidelines for the use of persistent identifiers in taxonomic publishing have been widely proposed [27,28], yet they remain largely focused on editorial artefacts rather than on taxa themselves.

Taxonomy constitutes a paradoxical exception. Although taxa form the common conceptual backbone of all life sciences, they remain difficult to integrate as stable objects within digital infrastructures. These difficulties do not result from excessive taxonomic complexity, but from a more fundamental deficit: the absence of a taxonomic object clearly defined as a persistent entity.

Stabilised name lists, registers of nomenclatural acts, local taxonomic concepts dependent on usage contexts, or taxonomic backbones based on management classifications allow pragmatic functioning in the digital world, but at the cost of information fragmentation, loss of cumulability and limited interoperability. Taxa as recognised within these systems thus appear unstable, ambiguous and difficult to integrate, not because they are intrinsically so, but because they are confused with their names, their usages or their successive interpretations.

From this perspective, most current limitations of taxonomic infrastructures appear not as failures of implementation, but as symptoms of an unresolved ontological ambiguity regarding the nature of their core objects.

The framework proposed here removes this conceptual lock. By recognising the taxon as a historically continuous entity of knowledge, distinct from its successive interpretative states and instituted through a unique taxonomic act, it becomes possible to identify it explicitly as a full-fledged FAIR digital object [29]. In this context, the assignment of a persistent identifier is no longer a technical patch intended to artificially stabilise unstable objects, but the minimal and necessary expression of the ontological reality of the taxon.

PIDs do not create the persistence of taxa; they make it operational. They provide the explicit anchoring point that allows digital infrastructures to unambiguously link biological data, taxonomic usages, competing classifications and successive revisions to a single reference object. The convergence between taxonomy and digital infrastructures therefore does not rely on adapting taxonomy to computational constraints, but on explicitly recognising what taxa have always been: persistent scientific objects, now fully compatible with the requirements of digital interoperability.

Clarifying taxonomic objects as historically continuous entities creates the conditions under which persistent identifiers can function as genuine reference anchors, rather than as compensatory devices for conceptual instability. In this sense, the present framework does

not propose a new technical standard, but delineates the conceptual space within which future taxonomic systems can remain both dynamic and cumulative.

### *3.4. Operational Implementation: Articulating Taxonomic Temporal Strings with Existing Taxonomic Infrastructures*

The operational implementation of taxonomic temporal strings can be built on three complementary mechanisms: (1) assigning a persistent identifier (PID) at the code-compliant founding act, thereby anchoring the taxon's identity independently of subsequent revisions and applying only to taxa formally instituted through valid publication, not to provisional entities such as informal clades, operational taxonomic units, or candidate species; (2) representing taxonomic treatments (e.g., diagnoses, circumscriptions, synonymisations, and classificatory repositionings) as dated and attributed assertions or events, building on existing treatment-centric practices and workflows [30] and compatible with DiSSCo priorities [31]; and (3) maintaining an explicit separation between string persistence, as the reference object, and substance dynamics, as the revisable interpretive corpus, thereby enabling cumulative knowledge management without duplicating taxon identities. For routine taxonomic practice, the framework therefore distinguishes between acts that institute a new taxon-level reference object and require a new PID, and acts that modify nomenclatural usage or the interpretation of taxonomic substance without changing taxon identity.

Within this framework, taxonomic substance is not a homogeneous block mechanically transferred between taxonomic temporal strings. A taxonomic temporal string serves as a persistent reference line to which instantiated taxonomic entities and sufficiently independent taxonomic information units can be attached, while remaining cumulative as a historical record. These include extension-level evidence, intension-level content, including rank and classificatory properties [2], and name-level, code-governed information. At the implementation level, two principles are central: first, the attachment of instantiated taxonomic entities and associated information units to taxon-level PIDs; second, the traceability of their possible reassignment between PIDs through later taxonomic revisions. In cases such as synonymy or lumping, what is reorganized is therefore not the identity of the strings themselves, but the attachment and interpretation of some of their associated components. Shared specimens, diagnoses, or usages do not erase the historical distinctness of the strings; rather, they become part of a traceable history of reinterpretation. A more detailed technical implementation remains a separate task. It would require explicit solutions for issues such as the current status of a taxon when the temporal string persists but its operability varies, the storage of historical taxonomic substance and the reconstruction of current substance from it, the handling of incomplete and heterogeneous updates, the reassignment during revision of taxonomic components and treatment-level knowledge units attached to persistent identifiers, and the concrete insertion of such events into the taxonomic trajectory.

This approach does not replace existing infrastructures; it organises and extends them. Nomenclatural registries such as ZooBank [32] and IPNI [33] already provide a backbone for formally established names and associated identifiers, which can be mapped to, or used to seed, taxon-level PIDs within the taxonomic temporal string framework. Treatments remain first-class, citable scholarly artefacts, but are interpreted here as successive, attributable states along a single taxonomic temporal string. The taxon PID provides the stable anchor linking biological data, usages, and revisions to one historically continuous reference object while remaining compatible with FAIR principles and existing Taxonomic Name Usage infrastructures.

- **Scale of the taxonomic reference space.** The orders of magnitude provided by existing taxonomic and nomenclatural infrastructures illustrate the scale of the taxonomic reference space. The Catalogue of Life currently integrates more than 2.2 million accepted extant species [34]. In botany, the International Plant Names Index contains more than 1.6 million published plant names [33], while the World Checklist of Vascular Plants recognises more than 343,000 accepted plant species and more than one million synonyms [35]. In prokaryotic nomenclature, the List of Prokaryotic names with Standing in Nomenclature documents more than 59,000 taxon names [36]. At higher ranks, global registries such as the Interim Register of Marine and Nonmarine Genera record almost 400,000 published genus-level taxa across all organisms, including fossils [37]. These figures concern mainly taxa currently recognised and curated in global databases. Persistent identifiers could be progressively assigned to newly established taxa at the moment of valid publication and subsequently extended retroactively through curated taxonomic registries and synthesis infrastructures.

The total number of taxa historically instituted through valid taxonomic acts is necessarily larger, as many entities have subsequently been synonymised or reinterpreted, losing their taxonomic substance while remaining part of the historical record of taxonomic knowledge. Crucially, the framework proposed here concerns the identity of taxa established through valid publication under the relevant nomenclatural codes, not the much larger universe of name strings circulating in biodiversity data systems. These strings (including canonical forms, authorship variants, abbreviations, orthographic variants, and data-processing artefacts) likely reach several tens of millions across biodiversity infrastructures. By contrast, each taxon identity originates from a specific taxonomic act and persists through time as a taxonomic temporal string, within which successive reinterpretations redistribute taxonomic substance without creating new identities.

The taxonomic reference space therefore reflects the cumulative history of taxonomic acts rather than the fluctuating universe of name usages. In this perspective, it constitutes a finite set of historically instituted identities (PID-able taxa) that remains technically manageable with contemporary computational infrastructures, thereby grounding taxonomy as a stable digital reference system.

- **Governance and institutional integration.** Governance of taxon-level persistent identifiers would most realistically follow a federated model aligned with existing nomenclatural infrastructures. Registries such as ZooBank for zoological nomenclature [32], IPNI for plants [33], and LPSN for prokaryotes [36] already serve as reference systems within their respective Codes and could anchor identifiers at the moment of valid publication. Higher-level biodiversity infrastructures such as the Catalogue of Life or GBIF would then provide aggregation, resolution, and interoperability layers across disciplines. In this architecture, identifier governance remains distributed across the existing taxonomic ecosystem while enabling a coherent global layer of persistent taxon identities.

In practical terms, implementation could proceed through a staged approach: (1) assigning persistent identifiers to newly established taxa at the moment of valid publication; (2) progressively extending identifiers retroactively through curated taxonomic databases and registries; and (3) ensuring interoperability across biodiversity infrastructures through shared identifier resolution services. Within this framework, the nomenclatural Codes are treated as foundational because they govern the formal institution of taxa and the regulation of names. They therefore provide the basis for nomenclatural anchoring and PID assignment, while taxon identity is analysed here at a distinct ontological level.

#### 4. Conclusions: The Taxon as a Temporal Scientific Object

In this framework, a taxon is a temporal scientific object: its persistent identity is a taxonomic temporal string, its changing content is its taxonomic substance, and its documented interpretative history is its taxonomic trajectory. The dynamics of taxonomy do not lie in the repeated creation or destruction of taxa, but in the succession of interpretative states (descriptions, circumscriptions, diagnoses, usages and classificatory positions) through which taxonomic substance is revised, challenged and reorganised over scientific time.

Conceiving the taxon as a taxonomic trajectory, ontologically recognised as a taxonomic temporal string, makes it possible to dissociate the identity of the object from the variability of its content. This distinction resolves persistent confusions between biological entities, explanatory hypotheses and taxonomic units, and clarifies the register break that separates populations and lineages from phylogenetic constructs and from taxa. Within this framework, biological data (morphological, molecular, ecological or phylogenetic) do not generate taxa by continuity, but feed the substance from which taxonomic decisions are formulated, discussed and revised.

This ontological clarification renders the integration of taxonomy into contemporary digital infrastructures both intelligible and operational: taxa, as historically continuous entities, intrinsically satisfy the conditions required for persistent identification. Persistent identifiers do not confer artificial stability upon taxa; they recognise and exploit their inherent persistence. By anchoring taxa as reference objects, taxon-level PIDs add a missing referential layer alongside existing identifiers for names and taxonomic treatments, thereby reconciling conceptual stability with cumulative dynamics and providing a coherent framework in which empirical practices, biological theories and information systems can be articulated.

The ontology proposed here may best be presented through a process of exclusion, identifying what a taxon is not: neither a logical concept in the classical sense, nor a direct biological entity (population, species, lineage, or clade), nor a punctual state of knowledge (a taxonomic treatment), nor a purely nomenclatural artefact (a taxonomic name). What remains is a historically continuous entity of knowledge whose properties are observed in practice rather than postulated: identity persistence, historical continuity, irreversibility, revisability without destruction, relative independence from biological hypotheses, and traceability of uses and transformations. The ontological framework proposed here merely renders these properties explicit and coherent.

Thus defined, the taxon is recognised for what it must be above all: a scientific reference object, designed to enable communication, comparison, cumulative reasoning and collective revision of knowledge on biological diversity. It is precisely at this level that the paradoxes of stability are resolved, and that persistent identification becomes evident rather than artificial: indeed, the minimal, necessary and faithful translation of what the taxon has always been: a temporal scientific object. By grounding taxa as historically continuous reference objects, this framework restores taxonomy as a cumulative and autonomous scientific practice, capable of articulating biological knowledge, phylogenetic hypotheses and digital infrastructures within a single coherent ontology.

**Author Contributions:** Conceptualisation, T.B.; writing—original draft and visualisation preparation, T.B.; Formal analysis, T.B., N.B., R.Z. and R.V.-L.; Writing—review and editing, T.B., N.B., R.Z. and R.V.-L.; Visualization T.B., N.B., R.Z. and R.V.-L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** No new data were created or analyzed in this study. Data sharing is not applicable to this article.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Zaragüeta, R. Ancêtres. In *L'arbre du Vivant Existe-t-il?* Société Française de Système: Paris, France, 2011; Volume 28, pp. 41–62.
2. Bourgoïn, T.; Bailly, N.; Zaragüeta, R.; Vignes-Lebbe, R. Complete formalization of taxa with their names, contents and descriptions improves taxonomic databases and access to the taxonomic knowledge they support. *Syst. Biodivers.* **2021**, *19*, 359–378. [[CrossRef](#)]
3. Bourgoïn, T.; Vignes-Lebbe, R.; Bailly, N. Visualisation of taxonomic knowledge: Exploring and reporting taxonomic data, training students in taxonomy. *Biodivers. Inf. Sci. Stand.* **2019**, *3*, e37730. [[CrossRef](#)]
4. IPBES. *Scoping Report for a Methodological Assessment of Integrated Biodiversity-Inclusive Spatial Planning and Ecological Connectivity*; Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services: Bonn, Germany, 2024.
5. Garnett, S.T.; Christidis, L. Taxonomy anarchy hampers conservation. *Nature* **2017**, *546*, 25–27. [[CrossRef](#)]
6. Bourgoïn, T.; Szwedó, J. Toward a new classification of planthoppers (Hemiptera: Fulgoromorpha): 2. Higher taxa, their names and their composition. *Zootaxa* **2023**, *5297*, 562–568. [[CrossRef](#)] [[PubMed](#)]
7. Luo, Y.; Bucher, M.; Bourgoïn, T.; Löcker, B.; Feng, J.N. Phylogeny and classification of Cixiidae (Hemiptera, Fulgoromorpha): A new evolutionary scenario for the most diverse planthopper family. *Syst. Entomol.* **2025**, *50*, 428–447. [[CrossRef](#)]
8. Ai, D.; Bourgoïn, T.; Bai, W.; Zhang, Y. Toward a new phylogeny-based classification of Flatidae (Hemiptera: Fulgoromorpha): A diversification shaped by angiosperm expansion. *Mol. Phylogenet. Evol.* **2026**, *216*, 108508. [[CrossRef](#)]
9. Mozaffarian, F.; Bourgoïn, T. A new comprehensive generic framework for *Tettigometra* Latreille, 1804 s.l.: A taxonomic and nomenclatural revision of the tribe Tettigometrini (Hemiptera: Fulgoromorpha). *Insects* **2026**, *17*, 30. [[CrossRef](#)]
10. Turland, N.J.; Wiersema, J.H.; Barrie, F.R.; Greuter, W.; Hawksworth, D.L.; Herendeen, P.S.; Knapp, S.; Kusber, W.-H.; Li, D.-Z.; Marhold, K.; et al. *International Code of Nomenclature for Algae, Fungi, and Plants (Shenzhen Code)*; Regnum Vegetabile 159; Koeltz Botanical Books: Glashütten, Germany, 2018.
11. International Commission on Zoological Nomenclature. *International Code of Zoological Nomenclature*, 4th ed.; International Trust for Zoological Nomenclature: London, UK, 1999.
12. Parker, C.T.; Tindall, B.J.; Garrity, G.M. International Code of Nomenclature of Prokaryotes. *Int. J. Syst. Evol. Microbiol.* **2019**, *69*, S1–S111. [[PubMed](#)]
13. Hennig, W. *Phylogenetic Systematics*; University of Illinois Press: Urbana, IL, USA, 1966.
14. Pyle, R.L.; Barik, S.K.; Christidis, L.; Conix, S.; Costello, M.J.; van Dijk, P.P.; Garnett, S.T.; Hobern, D.; Kirk, P.M.; Lien, A.M.; et al. Towards a global list of accepted species V. The devil is in the detail. *Org. Divers. Evol.* **2021**, *21*, 657–675. [[CrossRef](#)]
15. Pyle, R.L.; Bailly, N.; Remsen, J.V. Modeling taxon concepts: A new approach to an old problem. *Biodivers. Inf. Sci. Stand.* **2022**, *6*, e93904. [[CrossRef](#)]
16. Remsen, J.V. The use and limits of scientific names in biological informatics. *ZooKeys* **2016**, *550*, 207–223. [[CrossRef](#)] [[PubMed](#)]
17. Mayr, E. *Systematics and the Origin of Species from the Viewpoint of a Zoologist*; Columbia University Press: New York, NY, USA, 1942.
18. Simpson, G.G. *Principles of Animal Taxonomy*; Columbia University Press: New York, NY, USA, 1961.
19. de Queiroz, K. Species concepts and species delimitation. *Syst. Biol.* **2007**, *56*, 879–886. [[CrossRef](#)]
20. Berendsohn, W.G. The concept of “potential taxa” in databases. *Taxon* **1995**, *44*, 207–212. [[CrossRef](#)]
21. Arnauld, A.; Nicole, P. *La Logique ou l'art de Penser*; Jean Guignart, Charles Savreux and Jean de Lavnay: Paris, France, 1662.
22. de Saussure, F. *Cours de Linguistique générale*; Payot: Paris, France, 1916.
23. Lewis, D.K. *On the Plurality of Worlds*; Blackwell: Oxford, UK, 1986.
24. Sider, T. *Four-Dimensionalism: An Ontology of Persistence and Time*; Oxford University Press: Oxford, UK, 2001.
25. Thomson, S.A.; Pyle, R.L.; Ahyong, S.T.; Alonso-Zarazaga, M.; Ammirati, J.; Araya, J.F.; Ascher, J.S.; Audisio, T.L.; Azevedo-Santos, V.M.; Bailly, N.; et al. Taxonomy based on science is necessary for global conservation. *PLoS Biol.* **2018**, *16*, e2005075. [[CrossRef](#)]
26. McMurry, J.A.; Juty, N.; Blomberg, N.; Burdett, T.; Conlin, T.; Conte, N.; Courtot, M.; Deck, J.; Dumontier, M.; Fellows, D.K.; et al. Identifiers for the 21st century: How to design, provision, and reuse persistent identifiers to maximize utility and impact. *PLoS Biol.* **2017**, *15*, e2001414. [[CrossRef](#)]
27. Geoffroy, M.; Berendsohn, W.G. The concept problem in taxonomy: Importance, components, approaches. In *MoReTax—Handling Factual Information Linked to Taxonomic Concepts in Biology*; Berendsohn, W.G., Ed.; Schriftenreihe für Vegetationskunde; Federal Agency for Nature Conservation: Bonn, Germany, 2003; Volume 39, pp. 5–14.
28. Agosti, D.; Benichou, L.; Addink, W.; Catapano, T.; Chakrabarty, P.; Dikow, T.; Groom, Q.; Penev, L.; Sautter, G.; Smith, V.; et al. Recommendations for the use of annotations and persistent identifiers in taxonomy and biodiversity publishing. *Res. Ideas Outcomes* **2022**, *8*, e97374. [[CrossRef](#)]
29. Sterner, B.W.; Upham, N.S.; Gupta, R.; Powell, J.A.; Franz, N.M. Standards for FAIR taxonomic concept representations and relationships. *Biodivers. Inf. Sci. Stand.* **2021**, *5*, e67348. [[CrossRef](#)] [[PubMed](#)]
30. Agosti, D.; Ioannidis-Pantopikos, A. Taxonomic treatments as FAIR Digital Objects. *Res. Ideas Outcomes* **2022**, *8*, e93709. [[CrossRef](#)]

31. Hardisty, A.; Saarenmaa, H.; Casino, A.; Dillen, M.; Gödderz, K.; Groom, Q.; Hardy, H.; Koureas, D.; Nieva de la Hidalga, A.; Paul, D.L.; et al. Conceptual design blueprint for the DiSSCo digitization infrastructure—Deliverable D8.1. *Res. Ideas Outcomes* **2020**, *6*, e54280. [[CrossRef](#)]
32. Pyle, R.L.; Michel, E. ZooBank: Developing a nomenclatural tool for unifying 250 years of biological information. *Zootaxa* **2008**, *1950*, 39–50. [[CrossRef](#)]
33. IPNI—International Plant Names Index. Royal Botanic Gardens, Kew: Richmond, UK; Harvard University Herbaria and Libraries: Cambridge, MA, USA; Australian National Herbarium: Canberra, Australia. Available online: <https://www.ipni.org/> (accessed on 2 February 2026).
34. Bánki, O.; Roskov, Y.; Döring, M.; Ower, G.; Hernández Robles, D.R.; Plata Corredor, C.A.; Stjernegaard Jeppesen, T.; Örn, A.; Pape, T.; Hobern, D.; et al. (Eds.) *Catalogue of Life (2026-02-13 XR)*; Catalogue of Life Foundation: Amsterdam, The Netherlands, 2026. [[CrossRef](#)]
35. Govaerts, R.; Nic Lughadha, E.; Black, N.; Turner, R.; Paton, A. The World Checklist of Vascular Plants, a continuously updated resource for exploring global plant diversity. *Sci. Data* **2021**, *8*, 215. [[CrossRef](#)] [[PubMed](#)]
36. Parte, A.C.; Sardà Carbasse, J.; Meier-Kolthoff, J.P.; Reimer, L.C.; Göker, M. List of Prokaryotic names with standing in nomenclature (LPSN) moves to the DSMZ. *Int. J. Syst. Evol. Microbiol.* **2020**, *70*, 5607–5612. [[CrossRef](#)] [[PubMed](#)]
37. Rees, T. (Comp.) *The Interim Register of Marine and Nonmarine Genera (IRMNG)*; Checklist Dataset. Available online: <https://www.irmng.org> (accessed on 11 July 2025).

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.