Zootaxa 1950: 39–50 (2008)
www.mapress.com/zootaxa/

ISSN 1175-5326 (print edition)
ZOOTAXA
ISSN 1175-5334 (online edition)

ZooBank: Developing a nomenclatural tool for unifying 250 years of biological information*

RICHARD L. PYLE[1] & ELLINOR MICHEL[2]

[1]*Department of Natural Sciences, The Bishop Museum, 1525 Bernice St., Honolulu, Hawaii 96817, USA.*
E-mail: deepreef@bishopmuseum.org
[2]*International Commisssion on Zoological Nomenclature, c/o The Natural History Museum, Cromwell Road, SW7 5BD London, U.K.*

* *In:* Minelli, A., Bonato, L. & Fusco, G. (EDS) *Updating the Linnaean Heritage: Names as Tools for Thinking about Animals and Plants. Zootaxa*, 1950, 1–163.

Table of contents

Accepted by A. Minelli: 24 Oct. 2008; published: 5 Dec. 2008

Abstract

Nomenclature represents the backbone upon which virtually all biological information is organized. However, the practice of zoological nomenclature has changed relatively little since its start in 1758. As modern technology changes the paradigm under which modern scientists exchange information, there is increasing need to capitalize on these same technologies to fortify nomenclature. ZooBank has been proposed as the official registry of names and nomenclatural acts, in zoology, as well as associated published works and their authors, and type specimens. Having a coordinated registry of zoological names, integrated with the existing Code of Zoological Nomenclature, will allow increased efficiency of communication among biologists, and enhanced stability of names. Such a registry would encompass two distinct realms, each with their own set of challenges. Retrospective registration involves the monumental task of aggregating and validating two and a half centuries of existing names, whereas prospective registration must be tightly integrated with the future paradigm in which scientific names are created and managed under new models of publication. The prototype of ZooBank has been hosted at Bishop Museum during its initial development phase. Following the lead of standard-setting bodies in biodiversity informatics, Life Science Identifiers (LSIDs) have been selected for use as the globally unique identifiers for ZooBank registration entries. The first ZooBank LSIDs were issued on January 1st, 2008, and included five new fish species described in a work published that same day, as well as all 4,819 names established in the 10th Edition of Linnaeus' *Systema Naturae*. Three alternate scenarios for implementing mandatory registration in ZooBank have been articulated, each incorporating different degrees of coordination between published works and registration events. A robust discussion involving a broad spectrum of practicing zoological taxonomists is required over the next several years to define the specific implementation aspects of ZooBank.

Key words: Linnaeus, Prospective registration, Registration of names and acts, Retrospective registration, *Systema Naturae*

Introduction

Sense and stability in nomenclature

Stable nomenclature is at the heart of clear and unambiguous communication about biodiversity. Species names provide the most consistent anchor to which all taxonomic, ecological, molecular, conservation, and other biologically relevant data are attached. Legal protection and policy are also linked with names, on the assumption that the groups indicated by the names are consistent through time and among places. Scientific discussion relies on names having unequivocal, context-independent meanings. Medical and veterinary implementation requires communication about unambiguous identifications. Although discovery and delineation of species may receive the emphasis of high-profile press coverage, all taxonomic practice is crucially dependent on a stable nomenclature to provide a steady platform on which to build. The International Commission on Zoological Nomenclature (ICZN) has, for the past 113 years, set the rules by which scientific names for animals are established, as currently set forth in the ICZN Code of Nomenclature ('the Code', 4th Edition, 1999). The ICZN is the single professional organization devoted to ensuring that this work happens in a globally consistent way, providing continuity both for new species discoveries and for the correction of errors and inconsistencies in past works.

Development of a registry for new animal names (prospective registration) and a complete listing of existing names (retrospective registration) has long been a goal for biologists. The stakeholders for a gold-standard registry of animal names are diverse and central to the functioning of many biological sciences and to policy concerned with the living world. They include not only taxonomists, ecologists, and biodiversity informatics specialists but also conservationists, medical and veterinary workers, planners, policy makers, lawyers and even customs enforcers. Their requirements include ready access to a system of unambiguous answers to questions on the availability and validity of animal names that can be retrieved both by ordinary people and machines. The ICZN is meeting this need by developing ZooBank, a web-based registry of animal names

(Polaszek *et al*. 2005a). This will include nomenclatural acts (including new names), publications, authors and information on primary type specimens, and serve as a resolver for LSIDs (Life Science Identifiers), a tool for global communication among bioinformatics projects. It will be both a hub and authority for nomenclatural information.

Taxonomy, nomenclature and typification

Taxonomy and nomenclature are closely allied, but separate and complementary endeavors in developing the language of biodiversity. Discovering and delimiting species is the challenging job of alpha taxonomy; determining relationships and establishing higher taxa is referred to as beta taxonomy. Delimiting both alpha and beta taxa requires using a range of character data to test hypotheses about the inclusiveness of taxon definitions. This can naturally lead to strongly opposing alternative points of view, depending on character selection, method of analysis, and philosophical stance of the taxonomist. Definitions of taxa, from species to genera to higher taxa, can thus change significantly as the iterative process of improving the tests of taxonomic boundaries weighs alternative hypotheses and moves to new conclusions. Although it may be a source of frustration to end-users who simply want defined taxonomic entities, this process of change is a sign of the health of the science of taxonomy. Ultimately, if data accumulation were to saturate and if philosophical perspectives on species definitions were to converge, it is possible that taxonomy would stabilize and reach consensus definitions for taxa (changing only to accommodate ongoing organismal evolution). This situation is not on the horizon.

By contrast, the establishment of scientific names of animals is not a scientific process of testing alternatives; rather, it involves a bibliographic and quasi-legal process of presentation of a name with appropriate supporting documentation in a publication. Although a scientific name is generally established within the context of a published work on taxonomy, its link to actual organisms is through the primary type specimen (or specimens). This process of typification allows the name to be tied to a physical standard (and hence provides an objective basis for identifications), but leaves room for taxonomy to change; different names can be applied to taxa as is appropriate for their new boundaries. Figure 1 presents a tree-based example, in which alternative interpretations by different taxonomists result in different generic groupings, each of which could take a different name depending on the type species of the generic group. The same process could be visualized simply based on variation, with a more inclusive ('lumping') perspective requiring one type specimen for a species, thus receiving one name; whereas a more divisive ('splitting') perspective requires names derived from several type specimens for the perceived groups. Choosing between available names for types in a group is generally governed by the Principle of Priority, such that name first established should be used for that group (Figure 1). However, even if names are not in current use for a group, if they were originally validly published they are not permanently retired, as they may well be needed in the future. Taxonomic work may split an existing group, because less inclusive taxa are more consistent with data in hand. Having older names ready to apply provides an immediate tool for recovering past information on that taxon. This means that ZooBank must include both names in current use and all past, validly described names.

We want to underscore that the work of nomenclature aims for stability in names, but is completely independent of the process of flexibility in taxonomic interpretation. This philosophy is fundamental to the ICZN's role, as articulated in the Introduction to the 4th Edition of the ICZN Code which states:

There are certain underlying principles upon which the Code is based. These are as follows:

(1) The Code refrains from infringing upon taxonomic judgment, which must not be made subject to regulation or restraint.

(2) Nomenclature does not determine the inclusiveness or exclusiveness of any taxon, nor the rank to be

accorded to any assemblage of animals, but rather provides the name that is to be used for a taxon whatever taxonomic limits and rank are given to it.

(3) The device of name-bearing types allows names to be applied to taxa without infringing upon taxonomic judgment. [etc] (ICZN p. xix).

A cartoon graphic for the relationship of the trinity of nomenclature, taxonomy and type specimens is shown in Figure 2.

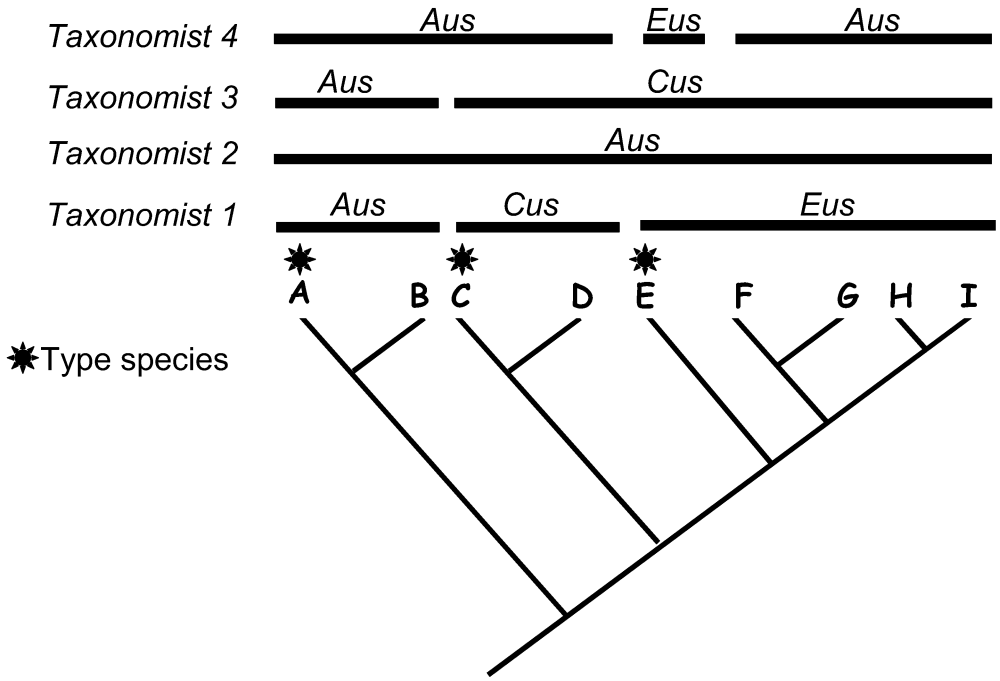

FIGURE 1. Diagrammatic representation of different nomenclatural interpretations for species A through I, mapped to a hypothesized phylogeny. Taxonomist 1 recognizes three new genus names; the genus *Aus* is typified by species A; *Cus* by species C, and *Eus* is by species E. Taxonomist 2 treats them as congeners (the diagram assumes that *Aus* has nomenclatural priority over *Cus*, and *Cus* over *Eus*). Taxonomist 3 recognizes two genera. Taxonomist 4 believes the underlying phylogeny is incorrect, and interprets the breakdown of two genera differently. Typification is necessary to establish which cluster each genus name is associated with, when taxonomic definitions of genera change.

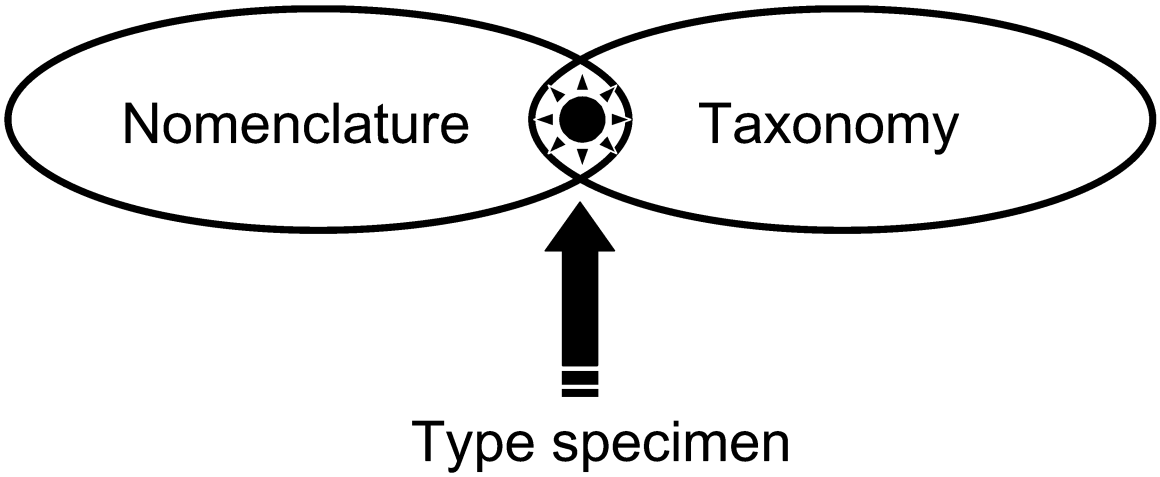

FIGURE 2. Diagrammatic representation of how type specimens represent the intersection of nomenclature and taxonomy.

Why do names need regulation?

The ICZN was established to address a situation in the mid 19th century described as 'increasing chaos in zoological nomenclature' (Melville 1995). Taxonomists were tackling the description and naming of an ever-expanding number of species they encountered from explorations of distant places, in addition to increased recognition of diversity from their home turf. The number of errors, inconsistencies and redundancies in application of names, either in synonymy or homonymy, began to create serious problems in the core objective of the Linnean enterprise of an unambiguous name for each species. Taxonomists working without access to the full literature inadvertently described taxa with the same name as another species in another taxonomic group, creating homonyms. Conversely, different workers in different parts of the world would unknowingly assign different names to what would otherwise be regarded as the same species, thereby creating synonyms.

At its inception, the ICZN acted on behalf of the zoological community at large, and with cooperation among taxonomists despite linguistic, political and taxonomic differences, even through world wars, to achieve a set of stabilizing rules for naming animals. The problems to be addressed then were, at their root, the result of lack of access to published information. The current revolution in information availability means that the information is now increasingly accessible, but in its shear volume the problems of disorganization from inappropriate names are becoming more starkly apparent. The consequences are serious, in that information is inappropriately presented and errors are propagated. Rigorous nomenclature must become dynamically integrated into the tools of cybertaxonomy.

The challenges of prospective and retrospective registration

In an ideal world, ZooBank would include full, verified information on all available names for animals. However, with 16,000–24,000 new additions yearly (N. Robinson, Zoological Record pers. comm., P. Bouchet, pers. comm.) to an estimated 1.7–1.8 million described animal species (Bouchet 2006), the numbers of names to be checked for homonymy and objective synonymy is enormous, so the logistics of populating the registry require strategic approaches. The need for reliable names for biodiversity work is urgent, so the tasks must be partitioned.

The first task for ZooBank is the establishment of a 'Black List' of unavailable names. This would serve as a foundation for on-line quality control that could be applied across projects, for example, by publishers to flag inappropriate usages of unavailable names. This is achievable with the recent digitization of the *Official Lists and Indexes of Names and Works in Zoology* by J.D.D. Smith, which could be implemented (through its 2007 version) within ZooBank.

The creation of a 'White List' of all available names is a much more difficult task, comprised of two parts. **Prospective registration** of new names, as they are published, will require active listing by taxonomists as they describe their thousands of new species each year, in a manner analogous to registering GenBank sequences today. This will rely on large-scale participant buy-in. Publishers are likely to support this, as the LSIDs resulting from ZooBank listing (described below) will allow greater exposure for their publications. The nomenclatural reliability of biodiversity informatics initiatives such as the Global Biodiversity Information Facility (GBIF), Encyclopedia of Life (EoL), and Catalogue of Life (CoL), which aim to present current taxonomic knowledge, will improve through dynamic integration with ZooBank. ZooBank will also increase the scope for linking with type specimen information held in museum databases to wider bioinformatics initiatives.

Working taxonomists appear eager to register names, and the incentives will increase with broader community use. Although the taxonomic community is strongly behind ZooBank, a straw poll at a large conference on the topic (EDIT 2008) indicated that registration mandated by the Code (see below) should wait until

the project has sufficient content, exposure and momentum. Ultimately, however, it is envisaged that as electronic (paperless) publication becomes more common, a mandatory registration system will become necessary. This point will be a central issue in discussions for publication of the fifth edition of the Code of Zoological Nomenclature (e.g. http://www.iczn.org/electronic_publication.html).

Retrospective registration of existing names is the greatest challenge for populating ZooBank. Initial assembly of published names can come from historical sources such as Sherborn (1902–1933) and Neave (1939–1996), and various taxon-specific nomenclatural databases. Names will then need to be flagged as to their level of nomenclatural vetting. One suggestion is a coding system. For example, a Gold (or green) flag would indicate the name has been checked to its original published source, a Silver (or yellow) flag indicates it has been checked to a reliable secondary source (e.g. a respected checklist) and a Bronze (or red) flag indicates it has been dumped from an unvetted source. Groups of names of particular interest could then be targeted to be worked-up *in toto*. For example, groups of concern to CITES could be prioritized with targeted funds. It might also be possible, with the completion of the Biodiversity Heritage Library, that 'citizen scientist' initiatives could be enlisted to populate ZooBank. Much of this depends on exactly how the scope of ZooBank will be established (see below), and what protocols for data quality assurance are put into place.

Although the ultimate goal is for ZooBank to act as a definitive source of names, it is the nature of historical work and taxonomic research that new information can cast doubt on past certainties. As a result, ZooBank will need to remain updatable, with allowance for external, expert input. We expect that the development of search tools will increase the possibility that conflicts within the ZooBank database, and also with external sources, could be flagged up for correction. We anticipate that ZooBank will develop into an indispensable tool, allowing unprecedented access to reliable nomenclatural information for biodiversity workers worldwide.

Technical implementation

ZooBank was launched as a functional prototype web site (http://zoobank.org) on January 1st, 2008, coinciding with the 250th anniversary of the official start of Zoological Nomenclature. At its launch the ZooBank registry included 4,819 names established in the 10th Edition of Linnaeus' Systema Naturae (Linnaeus, 1758), as well as five new fish species names established in an article published concurrently with the launch of ZooBank. Since that time, additional nomenclatural acts, published works, authors and type specimens have been both prospectively and retrospectively registered.

Server architecture and software platform

The initial prototype implementation of ZooBank is being developed at the Bishop Museum in Honolulu. The Bishop Museum was selected in part because of its association with the ICZN (former Commissioner and President of the ICZN Neal Evenhuis, and current Commissioner Richard Pyle), and in part because of the existing network facilities and technical support. In particular, the implementation of VMware Virtual Server Architecture allows for multiple distributed and redundant server platforms to be easily established, enabling improved performance and failover support.

In its initial implementation, ZooBank is split across two separate virtual servers, both running the Microsoft Windows 2003 Server operating system. One is a dedicated web server, hosting the user interface and web services (http://zoobank.org), which are being developed using Microsoft ASP.NET, and the VB.NET programming language. The other virtual server hosts the database content, which is implemented with Microsoft SQLServer 2005. The data model is derived from portions of the *Taxonomer* data model (Pyle

2004), with modifications specific to the needs of the ZooBank registry. For its initial implementation, there is only a single instance of the database (with suitable backup); but options for globally-distributed replicated instances are being explored for the long-term.

All of these (and other) technical implementation details are provisional, established only as an initial development environment, and do not necessarily have any bearing on the final specifications for the ZooBank implementation.

Scope of ZooBank registry

At present, four different kinds of data objects may be registered within ZooBank: Nomenclatural Acts, Publications, Authors, and Type Specimens. Each of these classes of data objects has direct implications for nomenclature (in accordance with the Code). Although some of these data objects fall within the overlapping scope of other data initiatives (e.g., the library community for publications and authors, and the natural history collections community for type specimens), their importance to Zoological Nomenclature mandates that ZooBank establish its own registry for these items in order to preserve autonomy, in the event that external data entities do not persist for as long as ZooBank persists. ZooBank was originally conceived as being a registry for zoological names as covered by the Code. However, as has been observed by the biodiversity informatics community, there are as many different notions of a name as there are database systems designed to record them. This is not only true across the different major Codes of nomenclature (Botanical, Bacteriological, Zoological, Viral, and Cultivated Plants) and within the broader biodiversity informatics community, but also within the realm of practicing zoologists. In some cases, names amount to little more than strings of text characters, sometimes inclusive of authorship and/or year, sometimes restricted to the individual name elements themselves. In some interpretations, alternate spellings constitute different names, whereas in other interpretations, such orthographic variations are regarded as alternative representations of the same name. Some database systems are designed to treat a name as only a singular element of a full taxonomic name (e.g. only the species epithet), treating binomials and trinomials more as concatenations of two or three separate names. Others only regard the complete set of name elements in a given combination as a single name.

To avoid ambiguity, the core data object as registered in ZooBank is the Nomenclatural Act. A Nomenclatural Act is a type of Taxon Name Usage instance, which is defined very generally as the usage or treatment of a particular taxon name within some form of documentation (see elaboration of documentation below). In the context of ZooBank, Nomenclatural Acts are those particular usage instances that have some direct or indirect bearing on nomenclatural details, as governed by the Code. The most common types of Nomenclatural Acts are those name-usage instances that constitute the original establishment of new zoological names (i.e., original descriptions) in the family-group, genus-group, and species group (for simplicity and clarity, the registration of such acts are often referred to as as the registration of names; but in fact it is the nomenclatural act establishing the name that is registered). Other Nomenclatural Acts include emendations, lectotypifications, neotypifications, First Reviser actions, and other nomenclatural assertions that have direct bearing on aspects of zoological nomenclature according to the Code. Some have suggested that Nomenclatural Acts may also include particular name usage instances such as species-group names used in combination with a genus-group name other than the original combination. Although such Acts are not directly governed by the Code, they may affect nomenclature indirectly, such as cases involving secondary homonymy. The complete spectrum of taxon name usages that may be registered as Nomenclatural Acts within ZooBank has not been formally established, and is the subject of ongoing discussion.

Inherent to any taxon name usage instance is a documentation instance in which the usage occurred. Such documentation may be interpreted very broadly, but in the context of ZooBank, it is more narrowly limited to works published in accordance with Article 8 of the ICZN Code. Because the Code officially regulates vari-

ous aspects of published works, they represent the second data object included within the scope of the ZooBank registry. In the context of ZooBank, objects that constitute published works are not limited to traditionally cited units of publications such as journal articles and book, but may also include individual taxon treatments within an article or book. The reason for allowing the inclusion of individual taxon treatments as units of publication within the scope of the ZooBank registry, is to accommodate circumstances where the authorship of the name (= the authorship of the taxonomic treatment of a name, representing a Nomenclatural Act) differs from the authorship of the traditionally-cited parent unit of publication (article, book, etc.).

The third type of object included within the scope of the ZooBank registry is Authors. Although not as extensively governed by the Code as Nomenclatural Acts or Publications, Authors have nevertheless been integral to zoological nomenclature since its inception, and therefore warrant individual registration. In addition to the Authors of registered ZooBank Publications, contributors to the ZooBank registry may also be registered as Authors within ZooBank.

The final object type included within the ZooBank scope is Type Specimens. Of particular importance are primary or name-bearing types, i.e. holotypes, syntypes, lectotypes, and neotypes. Although specimens certainly fall into the domain of natural history museums in terms of data management, they nevertheless play a critical role in nomenclature as they are the physical standard to which the name is tied, and as such fall within the scope of ZooBank. Whether or not secondary (non-name-bearing) types may also be entered into the ZooBank registry has not yet been determined.

ZooBank LSIDs

One of the primary functions of ZooBank is to issue Globally Unique Identifiers (GUIDs) to each registered data object. GUIDs are necessary because taxonomic names, author names, publication citations, and references to type specimens are not unique (e.g. homonymy of names), are subject to inconsistent representations in textual form (e.g. *Homo sapiens* vs. *H. sapiens*; Linnaeus vs. Linn vs. L.; etc.), and are generally fraught with ambiguity. Most trained taxonomists have no difficulty resolving these ambiguities; however, computers are not so adept at making informed inferences and subjective interpretations. They are much more effective when interpreting, linking, and resolving unambiguous GUIDs.

ZooBank follows the lead of the Biodiversity Informatics Standards (formerly the Taxonomic Databases Working Group; TDWG), and the Global Biodiversity Information Facility (GBIF) in adopting Life Science Identifiers (LSIDs) as the type of GUID assigned to registered data objects. LSIDs were originally developed by IBM, and are implemented, maintained and perpetuated primarily by the biodiversity informatics community. LSIDs do not require centralized issuance, and do not directly cost any money to issue. Moreover, there is a growing body of software in development and available for use in implementing and resolving LSIDs.

An LSID has minimally five parts, with an optional sixth part. Each part is delimited by a colon (:) character (Figure 3). The first two parts are always the same for all LSIDs: the lower-case characters urn:lsid. The first part identifies it as a Universal Resource Number (URN), and the second part identifies it as an LSID. The third part is called the Authority Identification, and is usually (but not always) an internet domain name registered to the LSID issuing entity. The authority identification part of all ZooBank-issued LSIDs is zoobank.org. The fourth part is the Namespace Identification, and is used to partition sets of identifiers within a particular authority. In the case of ZooBank, there are four such logical sets, represented in issued LSIDs by the text act (for Nomenclatural Acts), pub (for published works), author (for authors of published works, and for registered users of ZooBank), and specimen (for type specimens). Finally, every LSID must have an Object Identification part. This part must be unique within the Authority + Namespace combination. For ZooBank LSIDs, the object identification is a Universally Unique Identifier (UUID), a standard form of GUID common to many computer applications. There are several reasons why UUIDs were chosen for the

object identification part of ZooBank LSIDs (instead of, for example, an integer number or alphanumeric code). The main reason is that UUIDs are themselves globally unique, and thus retain their identity even when stripped of the rest of the LSID parts. UUIDs are not self-resolving (i.e., having only a UUID does not allow you to automatically find out what the UUID represents), but the LSID wrapper structure (particularly the first three parts) allows resolution of the information content associated with LSIDs (in this case, ZooBank registration entities). In the event that LSIDs eventually fall out of favor within the biodiversity informatics community, the UUID itself could be embedded within a different self-resolving GUID protocol (due to its global uniqueness).

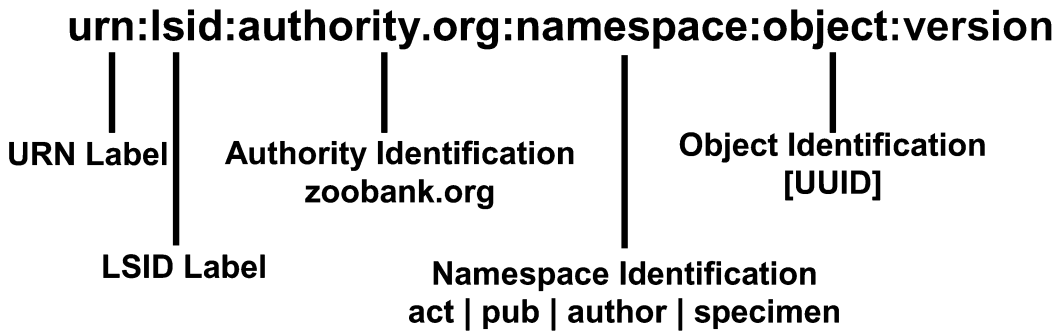

FIGURE 3. Components of an LSID. There does not appear to be a consistent implementation of the version part of the LSID, and it is not incorporated into ZooBank LSIDs.

Examples of existing ZooBank LSIDs include the following (the namespace part is emphasized in bold):

Nomenclatural Act: Original establishment of the species name, *Chromis abyssus*.
urn:lsid:zoobank.org:**act**:8BDC0735-FEA4-4298-83FA-D04F67C3FBEC

Publication: Published work in which *C. abyssus* was established (Pyle *et al*., 2008).
urn:lsid:zoobank.org:**pub**:68376390-7809-46FF-9EC4-1371B4AAD0FF

Author: First author of this published work (Richard L. Pyle).
urn:lsid:zoobank.org:**author**:8C466CBE-3F7D-4DC9-8CBD-26DD3F57E212

Type Specimen: Holotype of *C. abyssus* (BPBM 40861).
urn:lsid:zoobank.org:**specimen**:FDE70A5C-59C3-407B-B9A6-5A9A2DA14BD1

It should be emphasized that LSIDs (like UUIDs and all other GUIDs) are intended for use by computers, not by humans. LSIDs (again, like other GUIDs) are designed to operate behind the scenes, out of view of human eyes. They allow easy and unambiguous identification when linking electronic resources together, and it is in this context that they should be optimized.

Scenarios for mandatory registration

There has been a great deal of discussion concerning the implications of enforcing mandatory registration through amendments to or a new edition of the Code. In their announcement of ZooBank, Polaszek *et al*. (2005a) indicated that their intention was to make ZooBank a mandatory requirement (governed by the ICZN Code) for future names and nomenclatural acts. In a follow-up technical article, Polaszek *et al*. (2005b) outlined two proposed scenarios for how mandatory registration might be implemented in the Code, and a third scenario was proposed by Doug Yanega in a series of messages posted to the ZooBank email discussion list (Yanega, pers. comm.). A more cohesive description of these three alternative scenarios for mandatory registration (including diagrammatic flow charts) is included in Polaszek *et al*. (2008), and a brief summary is provided below.

It is important first to clarify the definition of three terms, which, for the purposes of this article, are as follows:

Registration: The process of entering a complete record in the ZooBank registry.

Publication: ICZN-compliant published works, as defined in Chapter 3 (Arts. 7–9) of the 4th Edition of the ICZN Code.

Availability: A nomenclatural act (such as a scientific name applied to an animal taxon) that meets the criteria of availability set forth in the Code.

Scenario #1: (Publication+Registration)=Availability

The first scenario posits that the act of registration would simply be added to the existing requirements of the Code, such that in order to be available under the Code, a name or nomenclatural act would need to be *both* published in accordance with existing Code rules, *and* separately registered in ZooBank. Registration could take place either before or after publication. If registration occurs before or within two years of publication, the date of availability is the publication date; but if registration is completed more than two years after publication, the date of availability is registration date (except in certain extenuating circumstances, as evaluated by the ICZN.)

Advantages of this scenario are that it would require a relatively small change to existing taxonomic practice, and it maintains implicit quality control via traditional publication venues. Moreover, many perceive this scenario as being the most likely to gain broad acceptance by the taxonomic community.

One disadvantage to this scenario is that it would require a somewhat complex procedure involving asynchronous publication and registration events, arbitrary time periods affecting date of availability, and petitions to the Commission in certain special circumstances. In particular, the temporal decoupling of publication and registration events establishes a somewhat ambiguous gray zone after publication and before registration when names & acts are assumed to be available, even though they are technically not available until registered. Also, this scenario still suffers from all the complexities and ambiguities associated with traditional paper publication entangled with nomenclatural availability. Finally, it may also require an increase in active role of ICZN staff (with associated costs) to process registration requests and verify Code compliance for issuance of GUIDs and exposing registration details to the public.

Scenario #2: Registration=Availability

In this scenario, the process of registration itself would be all that is required for availability of new names and nomenclatural acts. Prior or subsequent publication through traditional venues is encouraged, but would not be integral to nomenclatural availability.

The main advantage of this scenario is that the legalities of nomenclatural availability under the ICZN Code and the science of taxonomy are disentangled from each other. This philosophy is fundamental to the ICZN's role, as articulated in the Introduction to the 4th Edition of the ICZN Code (as quoted previously).

Other advantages of this scenario include the elimination of ambiguity concerning dates of availability, the rendering of existing complexities of nomenclatural availability of published works as moot, and the minor increases in the active role of ICZN staff.

The main disadvantage of this scenario is that it would represent a fundamental change to the way names and nomenclatural acts are established (i.e. altogether eliminating publication process as part of the requirements for availability). There is concern by many that by implication, taxonomists would lose their primary benchmark for establishing professional status (i.e. their CVs would have fewer publications listed). Even if

taxonomists followed through with proper taxonomic descriptions in published form, there is concern that journals might no longer publish taxonomic descriptions if the articles no longer carry the prestige of establishing new names and acts in accordance with ICZN rules. Moreover, although there are no existing requirements in the Code for peer-review or any other form of explicit quality control (for the taxonomy associated with the nomenclatural acts), the existing requirements for publication result in a de facto standard of peer-review and quality control. This would potentially be lost if nomenclatural acts were dissociated from the richer context of taxonomic work that often is included as part of published nomenclature. Finally, there is some concern that if the process of conferring availability of names under the Code were so simplified, lazy taxonomists might never get around to publishing the full description after the name is registered, potentially creating many names without robust taxonomic definitions. Even worse, bad taxonomists (and non-taxonomists) might abuse the system by registering hundreds of bogus and unneeded names, perhaps for unscrupulous reasons (e.g. selling names for money).

Scenario #3: Publication=Registration=Availability

In this scenario, the ZooBank web site would host a full-blown, edited, peer-reviewed online journal (like *ZooTaxa* or *Zookeys*) in which ***all*** names and nomenclatural acts must be published. In this scenario, the science of taxonomy becomes an explicit part of the nomenclatural process (by Code rules). Submitted manuscripts would be open to non-anonymous review by any interested or concerned taxonomist.

There are many potential advantages to this scenario. For example, all taxonomic publications would appear in a single venue (as is now done for bacteria), instead of scattered across thousands of journals. There would no longer be a potential for one author to steal another's work by trying to submit a plagiarized work to a journal that has a faster turnaround time. All manuscripts would be examined by a large contingent of reviewers, instead of just a handful, greatly improving the reviews as well as democratizing the process. These reviews would be public (instead of anonymous), so personal grudges or biases of the reviewers would be exposed to scrutiny by the whole community. Moreover, a dedicated nomenclatural journal of this sort would mean that the review criteria would explicitly address all necessary aspects of code-compliance and proper nomenclature. Indeed, this scenario would enjoy all of the other advantages of an online review process (fast, iterative, open to bidirectional feedback, etc.), and, perhaps most importantly, would not be subject to any copyright restrictions.

Equally significant are the potential disadvantages to this scenario. Foremost, it would represent a ***major*** and ***fundamental*** change to the way taxonomy is done, both in terms of legalities of nomenclature and for the science of taxonomy. The legalities of nomenclatural availability and the subjective science of taxonomy would, for the first time, be formally coupled under Code rules. Although the open review process proposed under this scenario is appealing, many taxonomic groups do not have many (or even any) experts who would serve as reviewers, and thus submitted manuscripts may never receive appropriate peer review (although this is no less true in the current publication paradigm). Such a system would impose a huge burden on the taxonomic community to provide peer reviews to 16,000–24,000 new names each year (again, in theory this would be no different from the current paradigm). It has also been pointed out that under this scenario, existing journals that depend on taxonomic descriptions and nomenclatural acts to fill their pages and maintain a subscriber base may be driven out of business. Also, the criteria for determining how, when and by whom a submitted manuscript should be deemed accepted will always be a subjective and contentious issue.

These are by no means the only possible scenarios for implementing mandatory registration in ZooBank. Many other possibilities exist, including various aspects of these three scenarios, as well as other factors not accounted by them. Much careful discussion and consideration will be required before a working scenario can be crafted, and the associated technical infrastructure developed. It is of vital importance that this discussion

be open to the broadest possible array of practicing taxonomists (not all of whom are able to participate in online discussion forums).

What is clear, however, is that the basic notion of online registration for nomenclatural acts (and associated publications) is generally desired by the majority of practicing taxonomists who have participated in discussions so far. As with so many aspects of science (and in particular issues concerning the Code) 'the devil is in the details'. Nevertheless, scientific names for animals are every bit as relevant and important to a wide variety of different fields in biology and medicine, as they were two hundred and fifty years ago in the time of Carl Linnaeus.

Acknowledgements

We wish to thank Sandro Minelli, Jon Todd, Ken Johnson, Andrew Polaszek, Yde de Jong, Charles Hussey, and the many ICZN Commissioners and other interested taxonomists who have contributed many of the ideas included herein.

References

Bouchet, P. (2006) The magnitude of marine biodiversity. *In*: Duarte, C. (Ed.) (2006). *The exploration of marine biodiversity: Scientific and technological challenges*. Fundacin BBVA, Bilbao, pp. 31–62.

EDIT (European Distributed Institute of Taxonomy) (2008). *Future Trends of Taxonomy - Annual General Meeting, Carvoeiro, Portugal, January, 2008*. EDIT, Paris.

Linnaeus, C. (1758) *Systema Naturae per regna tria naturae, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis*. Tomus I. Editio decima, reformata. Laurentus Salvius, Holmiae.

Melville, R.V. (1995) *Towards stability in the names of animals: A history of the International Commission on Zoological Nomenclature 1895–1995*. International Trust for Zoological Nomenclature, London.

Neave, S.A. (1939–1996) *Nomenclator Zoologicus*. Vols I–X. Zoological Society of London, London (available at http://uio.mbl.edu/NomenclatorZoologicus/browse.html)

Polaszek, A., Agosti, D., Alonso-Zarazaga, M., Beccaloni, G., de Place Bjrn, P., Bouchet, P., Brothers, D.J., Earl of Cranbrook, Evenhuis, N., Godfray, H.C.J., Johnson, N.F., Krell, F.-T., Lipscomb, D., Lyal, C.H.C., Mace, G.M., Mawatari, S., Miller, S.E., Minelli, A., Morris, S., Ng, P.K.L., Patterson, D.J., Pyle, R.L., Robinson, N., Rogo, L., Taverne, J., Thompson, F.C., van Tol, J., Wheeler, Q.D. & Wilson, E.O. (2005a) Commentary: A universal register for animal names. *Nature*, 437, 477. (http://www.iczn.org/Nature_Commentary.pdf)

Polaszek, A., Alonso-Zarazaga, M., Bouchet, P., Brothers, D.J., Evenhuis, N., Krell, F.-T., Lyal, C.H.C., Minelli, A., Pyle, R.L., Robinson, N.J., Thompson, F.C. & van Tol, J. (2005b) ZooBank: the open-access register for zoological taxonomy: Technical Discussion Paper. *Bulletin of Zoological Nomenclature*, 62, 210–220. (http://www.iczn.org/ZooBank_Paper.htm)

Polaszek, A., Pyle, R. & Yanega, D. (2008) Animal names for all: ICZN, ZooBank, and the New Taxonomy. *In*: Wheeler, Q.D. (Ed.). *The New Taxonomy*. CRC Press, Boca Raton, pp. 129–142.

Pyle, R.L. (2004) Taxonomer: A relational data model for managing information relevant to taxonomic research. *Phyloinformatics*. 1, 1–54. (http://systbio.org/files/phyloinformatics/1.pdf)

Pyle, R.L., Earle, J.L. & Greene, B.D. (2008) Five new species of the damselfish genus *Chromis* (Perciformes: Labroidei: Pomacentridae) from deep coral reefs in the tropical western Pacific. *Zootaxa*. 1671, 3–31. (http://www.mapress.com/zootaxa/2008/f/zt01671p031.pdf)

Sherborn, C.D. (1902–1933) *Index Animalium; sive, Index nominum quae ab A.D. MDCCLVIII generibus et speciebus animalium imposita sunt*. Trustees of the British Museum. London (available at http://www.sil.si.edu/digitalcollections/indexanimalium/)