

## **ZooBank: Guiding Principles and Policies**

### **PREAMBLE**

ZooBank, an envisioned electronic registry of zoological names, is intended to serve the taxonomic and broader biological communities by providing open-access data content and services pertaining to taxonomic names, nomenclatural acts, and associated information as governed by the International Commission for Zoological Nomenclature (ICZN). At the time this draft was circulated, there are many different views about what ZooBank should be (in terms of both scope and function), as well as how it should be implemented within the context of the ICZN Code. Though many questions regarding specific implementation remain to be answered, this document outlines the basic guiding principles that will shape the future development of ZooBank, and establishes a generic template regarding data content as it is submitted to, and incorporated into, the electronic registry.

### **DEFINITIONS OF TERMS**

For the purposes of this document, the following definitions of terms (subsequently highlighted in italics) are established:

***ICZN***: International Commission on Zoological Nomenclature<sup>1</sup>, the operation of which is funded through the International Trust for Zoological Nomenclature (ITZN) and is governed by a Constitution and By-Laws, under the auspices of the International Union of Biological Sciences (IUBS)<sup>2</sup>.

***Code***: International Code of Zoological Nomenclature (4th [= current] edition, 1999)<sup>3</sup>.

***ZooBank***: An electronic database serving as the official registry for the *ICZN*<sup>4,5</sup>.

***Names and Nomenclatural Acts***: Scientific names (as established for, or applied to, zoological taxa), as well as acts that have bearing on zoological nomenclature as governed by the existing (4<sup>th</sup> Edition) and/or future amendments to, or versions of the *Code*.

***Registration***: The incorporation of zoological *names and nomenclatural acts* into the *ZooBank* electronic database. The structure of the database is yet to be defined, but the act of registration will likely include a set of minimum required data fields, and the assignment of Globally Unique Identifiers (GUIDs)<sup>6</sup>.

***Mandatory Registration***: A reference to cases and situations where the *registration* of certain *names and nomenclatural acts* is required by the *Code*.

***Voluntary Registration***: A reference to cases and situations where the *registration* of certain *names and nomenclatural acts* is not explicitly required by the *Code*, but is encouraged and/or supported in order to expand the data content of *ZooBank*, and/or to allow initial development of *ZooBank* prior to the establishment of any form of *mandatory registration*.

***20XX***: A placeholder term for the year that registration is incorporated as a mandatory component for nomenclatural availability, as established in the *Code*.

**Retrospective Registration:** The registration of zoological *names and nomenclatural acts* that were established prior to 20XX.

**Verified Record:** Any record in the *ZooBank* database that has met minimal requirements for data quality, as specified by the ICZN Commission.

**Unverified Record:** Any records or entries in the *ZooBank* database that have **not** met minimal requirements, and/or are considered incomplete. An *unverified record* cannot be regarded as fulfilling the minimal requirements for cases involving *Mandatory Registration*.

**Data Source:** Any entity outside of the ICZN that may provide data content for the *ZooBank* database. Examples of possible *data sources* include broad-scope nomenclators (*e.g.*, ION/Zoological Record, ITIS, Species 2000, GBIF ECAT, uBio, AnimalBase, etc.); focused nomenclators (*e.g.*, BioSystematic Database of World Diptera, Catalog of Fishes, AntBase, CephBase, etc.); regional taxonomic databases (*e.g.*, museum taxonomic authorities, taxonomic catalogs and databases, electronic checklists, etc.); taxonomic data aggregators (*e.g.*, GBIF, OBIS); ancillary data providers (*e.g.*, BHL, ToL, EoL, GenBank, etc.); and individual taxonomists, biologists, other scientists, and amateur enthusiasts.

## **GUIDING PRINCIPLES**

The following represents some of the basic guiding principles to be applied to the future development of *ZooBank*:

- The ICZN shall have complete and autonomous control over processes, policies, and *verified records* associated with *ZooBank*;
- Authority for the development and implementation of *ZooBank*, including (but not limited to) decisions about when, where and how it is implemented; which *names and nomenclatural acts* are included within its content; policies concerning data access and usage; rules concerning *mandatory registration*; minimal data content and quality for constituting *verified records*; policies regarding content quality assurance; and any other decisions that affect the nature and function of *ZooBank*, rests with the ICZN Commissioners;
- In keeping with the tradition of the ICZN and its role in supporting zoological taxonomists, the future development and implementation of *ZooBank* will be conducted in an open and transparent way, with the ICZN Commissioners and Secretariat soliciting and encouraging input, discussion, opinions, and feedback from all interested individuals and parties;
- In further keeping with this tradition, data content associated with *verified records* in *ZooBank* will be made freely available to anyone for complete or partial downloading, and software tools will be developed in the spirit of open-source and open-access availability and transparency, subject to usage policies yet to be established (*e.g.*, Creative Commons License<sup>7</sup>);
- Data content in *ZooBank* will conform to prevailing data structure and accessibility standards, as set forth by Biodiversity Information Standards (TDWG)<sup>8</sup>, and data access

methods (either directly by human users, or through automated software agents) will be continuously updated to conform to emerging standards and technologies as they become available.

### **PROPOSED POLICY FOR DATA ACQUISITION**

As it currently exists, *ZooBank* has no actual data content. The primary purpose of the *ZooBank* database is to serve as a repository for *names and nomenclatural acts* under the envisioned amendments to or revisions of the *Code* concerning the *mandatory registration* after 20XX. A secondary purpose of *ZooBank* is to serve as a repository for the *retrospective registration* of *names and nomenclatural acts*, whether *voluntary* or *mandatory* (as determined by the ICZN Commissioners and incorporated into the *Code*). Because there is value in having significant data content relating to *names and nomenclatural acts* established or proposed prior to 20XX – both for purposes of database and software development within *ZooBank*, and to enhance the featureset and utility of the fully-implemented *ZooBank* system (e.g., automated homonym recognition, etc.) – there is a need to define a general policy for the acquisition of such data from sources external to *ZooBank*.

The following represents a draft policy template for establishing cooperative agreements with potential *Data Sources*, particularly with regard to how data might be contributed to *ZooBank* and how control over *unverified records* is established, as well as mechanisms for the long-term perpetuation of cross-references among multiple *data sources*. Specific policies are subject to approval by the ICZN Commissioners, along with prospective individual *data sources*, so the following should be regarded only as a general template.

- All *data sources* holding electronic taxonomic and/or nomenclatural data relevant to *ZooBank* will be encouraged to grant *ZooBank* access to those data;
- Except in specific cases where provided data have been pre-verified (as decided by the ICZN) and where both *ZooBank* and the specific data source have reached mutual agreement, all data provided to *ZooBank* from outside data sources will be initially treated as *unverified records*;
- Ownership and disposition of all *unverified records* within *ZooBank* shall be at the discretion of the *Data Source*, for as long as the data remain *unverified*. During this time, it is the prerogative of each *data source* to determine how and by whom those records may be accessed and/or distributed, including having the option of withdrawing the *unverified records* from *ZooBank* entirely.
- As *unverified records* are assimilated from various *data sources*, they will be cross-referenced within *ZooBank* among the one or more *data sources* that have contributed data for each *name or nomenclatural act*, using software tools developed by *ZooBank* and its associates for this purpose.
- When individual records – as assembled from one or many *data sources* – achieve the minimum criteria (as determined by ICZN) of data content and integrity, these records will be marked as *verified records*, and will become property of *ZooBank* and ICZN, and will thenceforth be subject to the usage and access policies established by ICZN for *ZooBank*. Once records have become *verified*, those records cannot be withdrawn or otherwise restricted in use or access by the *data source(s)* that contributed the data,

because at that point *ICZN* accepts responsibility for the perpetual availability and maintenance of the associated data. This does **not** mean that *data sources* forfeit any rights to the data they contributed; it only means that the content of the specific data fields constituting each *verified record* can no longer be retracted from *ZooBank* (except in cases of error or other appropriate reasons to remove or alter *verified records*).

- The actual data content owned by *ICZN* will be limited to only the specific data fields associated with *verified records*, which will in most cases represent only a subset of associated data maintained by the original *data source(s)*.
- At all times, from initial submission of *unverified records* by *data source(s)*, through the verification process and into perpetuity as *verified records*, *ZooBank* will maintain information as metadata, GUIDs, and URN links (as provided by the *data sources*) back to the corresponding original *data source* record(s). This will not only allow the establishment of a cross-referencing index among multiple *data source(s)*, but will also enable a mechanism to provide credit to contributing *data sources(s)*, as well as allowing corrections and other feedback to be provided back to the original *data source(s)* to enhance the value of their own data resources.

## **REFERENCES**

1. ICZN Home Page: <http://www.iczn.org/>
2. IUBS Home Page: <http://www.iubs.org/>
3. ICZN Code Online: <http://www.iczn.org/iczn/>
4. Polaszek, A., D. Agosti, M. Alonso-Zarazaga, G. Beccaloni, P. de Place Bjørn, P. Bouchet, D.J. Brothers, Earl of Cranbrook, N. Evenhuis, H.C.J. Godfray, N.F. Johnson, F.-T. Krell, D. Lipscomb, C.H.C. Lyal, G.M. Mace, S. Mawatari, S.E. Miller, A. Minelli, S. Morris, P.K.L. Ng, D.J. Patterson, R.L. Pyle, N. Robinson, L. Rogo, J. Taverne, F.C. Thompson, J. van Tol, Q.D. Wheeler & E.O. Wilson. 2005. Commentary: A universal register for animal names. *Nature*. 437:477. [http://www.nhm.ac.uk/hosted\\_sites/iczn/Nature\\_Commentary.pdf](http://www.nhm.ac.uk/hosted_sites/iczn/Nature_Commentary.pdf)
5. Polaszek, A., M. Alonso-Zarazaga, P. Bouchet, D.J. Brothers, N. Evenhuis, F.-T. Krell, C.H.C. Lyal, A. Minelli, R.L. Pyle, N.J. Robinson, F.C. Thompson, & J. van Tol. 2005. ZooBank: the open-access register for zoological taxonomy: Technical Discussion Paper. *Bulletin of Zoological Nomenclature*. 62(4):210-220. 16 December 2005. [http://www.nhm.ac.uk/hosted\\_sites/iczn/ZooBank\\_Paper.htm](http://www.nhm.ac.uk/hosted_sites/iczn/ZooBank_Paper.htm)
6. Global Biodiversity Information Facility (GBIF) & Biodiversity Information Standards (TDWG). TDWG Globally Unique Identifiers Wiki. <http://wiki.gbif.org/guidwiki/wikka.php?wakka=HomePage>
7. Creative Commons Home Page: <http://creativecommons.org/>
8. Biodiversity Information Standards (TDWG) Home Page: <http://www.tdwg.org/>