

Electronic publications need registration in ZooBank to be available

Frank-Thorsten Krell

Chair, ICZN ZooBank Committee and ICZN Commissioner; Department of Zoology, Denver Museum of Nature & Science, 2001 Colorado Boulevard, Denver, Colorado 80205–5798, U.S.A. (e-mail: frank.krell@dmns.org)

Thomas Pape

ICZN Commissioner; Natural History Museum of Denmark, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark (e-mail: tpape@snm.ku.dk)

Abstract. The 2012 Amendment to the International Code of Zoological Nomenclature allows electronic publication for nomenclatural purposes if certain requirements are fulfilled. We here explain these simple requirements, including what needs to be registered in ZooBank and when, which version of an electronic work is potentially Code-compliant (only the final, immutable version), and what is the correct date of publication (that of the final, immutable version; pre-final versions posted online are preliminary and as such always unavailable). We advise registration in ZooBank of publications that are issued in both electronic and paper version to secure the nomenclatural priority of the generally earlier electronic version. Failure to update a ZooBank record after publication will not have any impact on availability under the Code but will keep the registered information hidden from public view. Publishers may want to consider automated registration as an integral part of the publishing process; this has already been shown to be feasible and would ease the burden of manual registration for editors and authors.

Keywords. Availability; date of publication; electronic publishing; nomenclature; registration; works; ZooBank.

Scholarly communication has reached a point where most literature is distributed and consumed electronically (Tenopir et al., 2015). Paper copies may still be produced, but the electronic versions enjoy the vast majority of access and use. Most traditional print journals of major publishers make articles accessible on the internet long before the paper version is printed, and a growing number of journals are exclusively electronic.

In 2012, the *International Code for Zoological Nomenclature* (ICZN, 1999) was amended to allow for new names and nomenclatural acts to be established in works published electronically (ICZN, 2012). The amended rules included provisions specific to electronic forms of publication, in order to make such works available under the Code. In consideration of the ephemeral and sometimes malleable nature of electronic documents, the Commission introduced specific, but relatively simple requirements for electronic publications over and above those required for paper-printed works (ICZN, 2012; Krell, 2013, 2015; Cranston et al., 2015), which stipulate that electronically published works must:

- have fixed content and layout (Article 8.1.3.2)
- be issued after 2011 (Article 8.5.1)
- state the date of publication within the work itself (Article 8.5.2)
- be published in a work with an ISBN or ISSN (Article 8.5.3.2)
- be registered in the *Official Register of Zoological Nomenclature* (i.e. ZooBank) (Article 8.5.3)
- contain evidence in the work itself that such registration has occurred (Article 8.5.3)
- be registered with the name and Internet address of an organization other than the publisher that is intended to permanently archive the work (Article 8.5.3.1)

The first four requirements impose no significant burden, as they are either trivial or already an integral part of our contemporary scientific culture. Only the remaining three conditions are different, in that they represent a novel kind of requirement reaching beyond the routine publication of a scientific paper. In particular, the requirement to present evidence of registration within the work itself necessitates registration to happen *in advance* of publication. Also, neither the ISBN/ISSN nor the intended archive need to be stated in the work itself; currently approved archives include Bioline International, Biotaxa, British Library online archive, CLOCKSS, German National Library, Harvard Digital Repository Service, Hathitrust, LOCKSS, National Digital Heritage Archive of the National Library of New Zealand, National Library of Australia, Portico®, PubMedCentral®, State Library of Queensland, Virginia State Publications Depository Program, and Zenodo.

Unfortunately, registration does not always happen for electronically published works with new nomenclatural data, and authors, editors and publishers are not always aware of what details need to be registered (and when). Our intention is to clarify the requirements of the 2012 Amendment (ICZN, 2012) for making electronic works available, with special focus on registration.

What is Registration?

ZooBank (zoobank.org), the official registry of zoological nomenclature, was first proposed in 2005 (Polaszek et al., 2005). It was initially launched on 1 January 2008, and continues to grow and evolve (Polaszek et al., 2008; Pyle & Michel, 2008, 2009, 2010; Krell & Pyle, 2010). Anyone with a functional email address is free to establish a ZooBank login account, and register authors, works, and new names established in the family-, genus- and species-groups. Registration simply means that a record is established in ZooBank and a permanent and globally unique identifier has been assigned.

According to the amended Code (ICZN, 2012), only works produced electronically are required to be registered. Works produced on paper (and in accordance with other provisions of the Code) are not required to be registered, but the ICZN strongly encourages registration of such works. Also, the Code does not require registration of new names or nomenclatural acts for any publication (electronic or paper), but the registration of new names is strongly encouraged (Rosenberg et al., 2012). Registration facilitates automated indexing, linking and data extraction, and increases the visibility of new taxa. ZooBank does not currently support the registration of other kinds of nomenclatural acts (e.g. type designations, emendations, first reviser

actions), but support for these is planned. Finally, ZooBank is intended for the registration of works and nomenclatural acts governed by the ICZN Code; registration of taxonomic acts not governed by the Code (e.g. subjective synonyms and new combinations, and names above the family group ranks) is neither supported nor planned.

Why some ZooBank registrations are invisible: Updating the ZooBank record after publication

An important requirement of the amended Code is that evidence of registration (such as the ZooBank registration number) must be included within the work itself. This means that the work must be registered *before* it is published, but pre-publication registrations are hidden from public view on the ZooBank website to prevent premature release of unpublished information. To release the registered information to the public view, the ZooBank record must be updated *after* publication to include the publication date (as well as other optional publication metadata, such as volume number and pagination). This update should be done by the contributor who initially registered the work before publication. Failure in updating the ZooBank record with the date of publication will keep the registered information hidden from public view, but will not have any impact on the availability (under the Code) of the published work or any included new names or nomenclatural acts. It is just inconvenient and confusing to users consulting ZooBank who try to find a particular record in ZooBank, but are unable to do so because the publication date has not been entered.

What is an *electronic work*?

The Amendment requires works published electronically to be registered in ZooBank to be available, but there is some confusion about what *works* are and what counts as an *electronic work* that is subject to mandatory registration to be available. Most of this confusion derives from the fact that many publications currently produce articles in both electronic and paper form, often at different times and in different versions.

The Code defines *work* as ‘Any text or illustration, whether published, unpublished, or carrying a disclaimer’. ‘Any text or illustration’ is admittedly a useless definition to determine what needs to be registered in ZooBank. Works that need to be registered are the entities we cite in the references section: journal articles, book chapters, books, published abstracts. Registration of a journal, a journal volume, or a journal issue does not count as registration of an individual work and does not make the nomenclatural acts therein available.

If the authorship of a chapter of a book differs from the authorship of the book, it is useful to register the chapter separately. Because ZooBank currently does not offer this feature to users, details for such registrations should be sent to the ZooBank administrator by email (address under the ‘contact’ menu on the ZooBank webpage).

An *electronic work* is ‘A publication issued and distributed by means of electronic signals.’ Most people associate electronically published works with Portable Document Format (PDF) files. Although a PDF document is cited as an example of an electronically published work within the amended Code, the rules allow for any electronic document that is both ‘widely accessible’ and has ‘fixed content and layout’ (Article 8.1.3.2). In cases where a publication is distributed electronically prior to the

paper edition, the publication must be registered in ZooBank (and conform to the other requirements of the amended Code) if the new names and other nomenclatural acts are to be considered available from the date of the electronic edition.

Works published both on paper and electronically – priority issues

Many (if not most) modern journals that publish paper editions also produce electronic editions (e.g. as PDF) of the articles they contain. When publishing in a paper journal, it is easy to forget the implications of a parallel electronic edition that might be issued much earlier than the paper edition and potentially result in an earlier publication date (in the sense of the Code). If such an earlier electronic version is the final, immutable version, but is not registered in ZooBank (or is registered but fails to conform to the other requirements for electronic works, such as failure to include evidence of registration within the work itself, or no indicated online archive for the work in ZooBank), the article would be available from the later date of the paper edition. Article 21.9 of the Code states, ‘A name or nomenclatural act published in a work issued in both print and electronic editions takes its date of publication from the edition that first fulfilled the criteria of publication of Article 8 and is not excluded by Article 9.’ If the electronic version fulfils all requirements of the Code, its earlier publication date is to be used to establish nomenclatural priority. Authors may unnecessarily delay nomenclatural priority of their electronically published works and nomenclatural acts, if they fail to register and fulfil other requirements of the Code. Moreover, publishing a final version electronically that is unavailable for nomenclatural purposes potentially creates confusion, because it will be read, used and cited as if it were available anyway. Authors should avoid publishing works that appear to be final and nomenclaturally available, but are not. When in doubt, it is generally best to register all books and articles containing Code-governed nomenclatural acts before publication and comply with other requirements for electronic publication even if they are published on paper at some point.

So many versions of my paper: How do I determine the ‘final, immutable’ version?

Most scientific journals produce electronic copies of articles prior to the paper edition, or prior to the final electronic journal issue. These electronic versions come in all forms, including unedited accepted manuscripts, uncorrected proofs, and the version of record, often released before being integrated into a journal issue. The publishing world has been well aware of the potential confusion these different versions of what is intended to be one and the same publication can cause. Guidelines have been established defining and controlling these versions, published by the National Information Standards Organization in partnership with the Association of Learned and Professional Society Publishers (NISO/ALPSP, 2008). These NISO Recommendations are widely accepted and followed by publishers.

For nomenclatural purposes, only the final version with immutable layout and content is relevant (i.e. is available). If the content of an early electronic version is open to potential corrections, edits or similar changes in the work itself, this early electronic version is to be considered unavailable for nomenclatural purposes. In publishers’ terms (NISO/ALPSP, 2008), the final, immutable version is the *Version of Record*. In a separate article, one of us analysed which kinds of early online publications are available for nomenclatural purposes: Some publishers only publish

the version of record online, some post earlier versions (Krell, 2015). According to NISO/ALPSP guidelines, bibliographical metadata such as volume, issue and page numbers are not part of the content and do not change the version of record if added or changed when the article gets integrated into a journal issue. This would be a sensible way to interpret the requirement of ‘fixed content and layout’, but some authors disagree and consider the version of record integrated into a journal issue to be a different work from the version of record published in advance, because of changed page numbering and added volume and issue number and other metadata (Dubois et al., 2015). We do not support this view, but this issue will continue to be debated by the broader taxonomic community, and will be clarified in the next edition of the ICZN Code.

With all these versions, which is the correct date of publication?

Journals that publish early online versions of their papers in most cases use the date of the early online publication as the date of publication for the work. The rationale behind this practice is that intellectual priority of content is established with the first posting on the web, be it as an accepted, unformatted manuscript, as uncorrected or corrected proofs, or as the final version. Note that intellectual and nomenclatural priorities are different. Nomenclatural priority can only be established from the final, immutable version (version of record), not from a previous, preliminary version.

To minimize confusion, it is important that publishers ensure that the stated date of publication, as included within electronic publications (required by Article 8.5.2), accurately reflects the date on which the electronic work actually fulfilled the requirements of the Code. As a general rule, this should be the date of publication of the version of record, not of the first posting of a preliminary version. Although a misstated date does not necessarily render the work and any nomenclatural acts it contains unavailable under the ICZN Code, it does make it extremely difficult to determine the actual date of publication of the electronic work for purposes of nomenclatural priority. If the actual date of publication cannot be determined, then the publication date for nomenclatural purposes is 31 December of the year of publication unless other evidence for earlier publication can be provided (Article 21.3.2 of the Code). Not stating the exact date of publication of the version of record is a disservice to by delaying their nomenclatural priority.

This is a general problem of many publishers, whether their early electronic versions are versions of record or earlier, nomenclaturally unavailable versions. Publishers need to be aware of those complications, which are easy to fix but rather counter-intuitive to people primarily concerned with establishing intellectual priority.

What happens if not all requirements of the Code are fulfilled?

In some cases, the electronic edition of a work is obtainable before the ZooBank record is completed (e.g. missing indication of the intended archive or ISSN/ISBN). There are no clear ICZN-sanctioned rules for determining the date of publication for purposes of priority in such cases, but the general consensus among commissioners is that the date of publication (in the sense of the Code) should be interpreted as the date on which all requirements of the Code were satisfied. To help avoid potential confusion, it is always best if authors, editors and publishers of electronic works

ensure that all mandatory information is included in the registration before publishing the final, immutable version.

Advice for authors

Any author who is concerned about the nomenclatural priority of his or her work should ensure that the publisher has a firm routine for ZooBank registration. Alternatively, the author can register his or her work in ZooBank and enter required details (see above) during the revision or proofing stage. The work will then be available from the publication of the version of record (given that all other requirements of the Code are fulfilled).

Advice for publishers

Publishers should state the date of publication as the date on which all requirements of the Code are fulfilled, which is at the earliest when, in publishing terminology, the version of record is published. This date of publication should be stated instead of (or in addition to) the date of the first posting of a preliminary version. This will reduce confusion when determining nomenclatural priority.

Easing the burden: Automated registration as an integral part of the publishing process

Although the task of registering a work or a new name in ZooBank requires only a few minutes, it might be perceived as a burden. Moreover, manual registration carries a risk of errors. ZooBank was developed to include an automatic registration process to allow registration to be integrated into the publication workflow. This automatic registration process is based on the *TaxPub* standard for XML markup, an extension to the Journal Tag Publishing Suite of the National Library of Medicine (NLM; Catapano, 2010; Penev et al., 2012). Such NLM standards are increasingly becoming adopted by scientific publishers, and the journal *ZooKeys* (Pensoft Publishers) has demonstrated the feasibility of the implementation of *TaxPub* into the publication process of a largely taxonomic journal. Starting with issue 346 (published in November 2013), *ZooKeys* has very successfully incorporated automatic registration in ZooBank (Erwin et al., 2015). While the ICZN is not in a position to prescribe how publishers should produce their journals and books, we suggest that the expanding adoption of NLM standards (including *TaxPub*) among professional publishers in the field of taxonomy will greatly improve compliance with the ICZN rules for electronic publication. The advantages of incorporating NLM standards (including *TaxPub*) go well beyond automatic registration in ZooBank. By themselves, PDF documents represent little more than electronic versions of paper documents. While this certainly simplifies dissemination, the true power of electronic publication comes from structured data embedded within these electronic documents. Extensible Markup Language (XML) is the widely adopted standard for distributing structured data in electronic documents via the internet. While many journals produce XML editions of their published articles, the power of doing so is dramatically improved when the XML schema conforms to an international standard (such as NLM).

Acknowledgements

Rich Pyle (*ZooBank Developer and ICZN Commissioner, Bishop Museum, Honolulu, HI*) was instrumental in clarifying our arguments, helped us to avoid misunderstandings and informed us about the status of automated registration in ZooBank. We are

grateful for his substantial contributions to this paper. Svetlana Nikolaeva (*ICZN Secretariat London*) commented on the manuscript.

References

- Catapano, T.** 2010. TaxPub: an extension of the NLM/NCBI Journal Publishing DTD for taxonomic descriptions. *Journal Article Tag Suite Conference (JATS-Con) Proceedings 2010*. <http://www.ncbi.nlm.nih.gov/books/NBK47081/> (accessed 30 April 2015).
- Cranston, P.S., Krell, F.-T., Walker, K. & Hewes, D.** 2015. Wiley's EarlyView constitutes valid publication for date-sensitive nomenclature. *Systematic Entomology*, **40**: 2–4.
- Dubois, A., Bour, R. & Ohler, A.** 2015. What is an online 'preliminary version' of a publication in the meaning of Article 9.9 of the Code?—One more step on the trail of the Asian elephant. *Bulletin of Zoological Nomenclature*, **72**: 6–18.
- Erwin, T., Stoev, P., Georgiev, T. & Penev, L.** 2015. ZooKeys 500: traditions and innovations hand-in-hand servicing our taxonomic community. *ZooKeys*, **500**: 1–8.
- ICZN (International Commission on Zoological Nomenclature).** 1999. *International Code of Zoological Nomenclature. Fourth Edition*. xxix, 306 pp. International Trust for Zoological Nomenclature, London.
- ICZN (International Commission on Zoological Nomenclature).** 2012. Amendment of Articles 8, 9, 10, 21 and 78 of the International Code of Zoological Nomenclature to expand and refine methods of publication. *Bulletin of Zoological Nomenclature*, **69**: 161–169.
- Krell, F.-T.** 2013. Open-access journals are now open for discoveries of new plants and animals. *European Science Editing*, **39**: 11–12.
- Krell, F.-T.** 2015. A mixed bag: when are early online publications available for nomenclatural purposes. *Bulletin of Zoological Nomenclature*, **72**: 19–32.
- Krell, F.-T. & Pyle, R.L.** 2010. ZooBank Progress Report. *Bulletin of Zoological Nomenclature*, **67**: 271–277.
- NISO/ALPSP Journal Article Versions (JAV) Technical Working Group.** 2008. *Journal Article Versions (JAV): Recommendations of the NISO/ALPSP JAV Technical Working Group*. National Information Standards Organization, Baltimore, MD.
- Penev, L., Catapano, T., Agosti, D., Georgiev, T., Sautter, G. & Stoev, P.** 2012. Implementation of TaxPub, an NLM DTD extension for domain-specific markup in taxonomy, from the experience of a biodiversity publisher. *Journal Article Tag Suite Conference (JATS-Con) Proceedings 2012*. <http://www.ncbi.nlm.nih.gov/books/NBK100351/> (accessed 30 April 2015).
- Polaszek, A., Agosti, D., Alonso-Zarazaga, M., Beccaloni, G., Place Bjørn, P. de, 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., Tol, J. van, Wheeler, Q.D. & Wilson, E.O.** 2005. A universal register for animal names. *Nature*, **437**: 477.
- Polaszek, A., Pyle, R. & Yanega, D.** 2008. Animal names for all: ICZN, ZooBank, and the New Taxonomy. Pp. 129–142 in Wheeler, Q.D. (Ed.), *The New Taxonomy*. CRC Press, Boca Raton. 237 pp.
- Pyle, R.L. & Michel, E.** 2008. ZooBank: Developing a nomenclatural tool for unifying 250 years of biological information. Pp. 39–50 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.
- Pyle, R.L. & Michel, E.** 2009. Unifying nomenclature: ZooBank and Global Names Usage Bank. *Bulletin of Zoological Nomenclature*, **66**: 298.
- Pyle, R.L. & Michel, E.** 2010. ZooBank: Reviewing the first year and preparing for the next 250. Pp. 173–184 in Polaszek, A. (Ed.), *Systema Naturae 250. The Linnaean Ark*. CRC Press, Boca Raton.
- Rosenberg, G., Krell, F.-T. & Pyle, R.** 2012. Call to register species in ZooBank. *Nature*, **491**: 40.
- Tenopir, C., King, D.W., Christian, L. & Volentine, R.** 2015. Scholarly article seeking, reading, and use: a continuing evolution from print to electronic in the sciences and social sciences. *Learned Publishing*, **28**: 93–105.