

Is Copyright undermining Biodiversity Research and Conservation?

Donat Agosti

Research Associate, American Museum of Natural History, Dalmaziqual 45, 3005 Bern, Switzerland; agosti@amnh.org; +41 31 351 7152

Key Words: biodiversity; open access; conservation commons; copyright, Antbase.org; ants; Convention on Biological Diversity

Powerful PCs, the Internet, data exchange protocols, and more recently a cultural change among biodiversity scientists are rapidly developing into a huge international framework of biodiversity data. For example, the Global Biodiversity Information Facility (GBIF) is now providing access to well over 45 million specimen records based at over 100 institutions world wide. Analyses derived from biodiversity databases, collating data from literally all around the world, are now becoming integrated into the political decision-making processes at all levels in such places as Mexico and Madagascar. At the same time, an increasing number of scientific journals are being disseminated both in paper as well as in electronic form over the Internet. However, there is one serious problem: copyright.

Information on ants might serve as a good example. They are one of still few groups of organisms for which a complete, continually updated list of all the currently known 11826 species exists, a fast increasing access to high resolution type and other specimen imagery (e.g. antweb.org), and for which almost all the non-copyrighted articles, including the descriptions of new species, are online accessible, linked to by either the citations in the database, or the entire publications. This includes some 3,800 publications or an estimated 77,000 printed pages, roughly 7 pages per species (Agosti, 2003a). Assuming that there are currently 1.5 million species known, there are an estimated 10.5 pages published which could at a cost of one dollar a page be made accessible to the entire community. These would be by far the fastest and cheapest way to provide access to the entire corpus of our legacy of taxonomic publications

The problem of providing access though is not the cost, but the copyright, which is strictly enforced by modern publishers (Figure 1), notably the most important ones, such as Elsevier, Harvard University Press, Blackwell, Australia's CSIRO, some of the scientific associations, and luckily only few scientists themselves. Generally, all the copyright advocates are in the developed world. Loss of revenue is given as the main reason by the copyright advocates to enforce it. But how is it that others such as California University Press, Springer Publishers, or the American Entomological Society allow their ant-related publications to be freely available online? Perhaps the business model for publishing needs be changed and recalculated, considering that the distribution of information hardly costs anything anymore? Or are up-front paying models the right way to go, such as the fledgling BioMedCentral or the Public Library of Science who provide open access journals?

Copyright is regulated at the national level, and in, for example, the United States, copyright goes back 75 years. This requires the tedious task of getting copyright waivers for well over 300 recently deceased authors, as well as over 400 different, often obscure scientific journals in which articles have been published on ants. This is already an enormous effort, although

just a fraction of the other 1.5 million taxa similarly widely published; clearly resources that could be better spent on other aspects of science.

The really important issue in copyright is not transferring paper copies of journals into the digital realm, and thus having another 'silent' copy sitting around. Such electronic journals rather pose a problem to our science libraries. In earlier times each hard copy bought could be shelved for eternity. Access to electronic journals ceases at the moment, where the library chooses to abandon a title, and thus though even exorbitant amounts of money has been paid for access, nothing remains. In fact, there is the potential risk, that our public libraries being thus over time replaced by few huge corporate holders of scientific knowledge.

The far more relevant issue of the hard-copy-digital transition is incorporated in the vision of the World Wide Web's architect, Tim Berners-Lee: the Semantic Web (Berners-Lee et al. 2001). Berners-Lee considers the Internet not as a huge collection of static pages marked up to please our eyes, but a system where computers can read all the documents and databases and thus extract all the relevant information to answer our queries. This way, it will be possible to ask questions on the distribution of taxa, the identity of a given species, since somewhere in a document there might be hidden a "red ant in Costa Rica smaller than 2mm". The tools to make this happen are here and need only be applied, such as XML mark up language, and specific data standards are being developed (TDWG – the Taxonomic Database Working Group). We even have systems in place to advise on which names are valid (e.g. ICZN for Zoology) If copyright persists, the future is bleak, and we are in fact restricted to rewrite a huge part of modern taxonomy to circumvent this problem, or remain in the medieval ages of systematics.

Two additional points need to be made. Couldn't species descriptions be considered as part of our cultural heritage? Species descriptions and those of higher taxa can hardly be seen as a new intellectual creation. There are more than 1.5 Million species out there, all described, within a given framework as defined by the various Codes, and about 8 million to follow,. At the same time, most of the species described are those from biodiversity rich places, typically the tropics. With few exceptions such as Brazil, Argentina or Mexico (Carvalho et al., 2005), most of them are described and published in the developed world, and the scientific credit goes to the respective scientists and institutions, and the corporate world and some scientific institutions are getting the profit from selling those descriptions. This is hardly in agreement with articles 16 and 17 on exchange of the legally binding Convention on Biological Diversity (CBD, 1992). In fact a new North-South divide has opened, leading to a huge disparity of access to digital information since the Earth summit in Rio in 1992. That no presentation on this issue was scheduled at this conference is another indicator of unawareness among the biodiversity community, which is not the case generally (e.g. Berlin Declaration, 2003; Creative Commons; Conservation Commons).

Our scientific community needs to think on how to resolve this problem. Shall those publishers enforcing copyright being embargoed? Shall new Web-based ways of publishing species descriptions be developed replacing the current model? Why not endorse the principal of the 'Conservation Commons' or 'the Creative Commons' in the systematics community, or behave similar to the highly successful open source community, producing in many cases more reliable and stable than the corporate world, such as the operating system Linux? Why not reforming the Codes of Nomenclatures, whilst adopting them to online publishing, also make them providing the central name registry and a repository for species descriptions analogous to what GenBank does for DNA sequences (Agosti, 2003b)? After all, species descriptions are not ideas, but strings of words describing entire organisms. The inclusion of

the description of a particular taxon, together with high-resolution imagery of the specimen and distribution data would this way be accessible for anybody in all its scientific integrity.

Certainly, there are many more aspects needed to be considered – but there is one maxim: If we are serious about halting species loss, we have to know, what's out there. We can do this only if we pool openly all our information over the Internet. We can't let this effort be undermined by economic self-interests.

References

Agosti, D., 2003a. The last of its kind. *Nature* 424, 727

Agosti, D. 2003b. Encyclopedia of life: should species description equal gene sequence? *Trends in Ecology and Evolution* 18, 273.

Berlin Declaration, 2003. Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. <http://www.zim.mpg.de/openaccess-berlin/berlindeclaration.html>.

Carvalho, M.R, de et. al., 2005. Revisiting the taxonomic impediment. *Science* 307, 353

Conservation Commons. <http://conservationcommons.org>

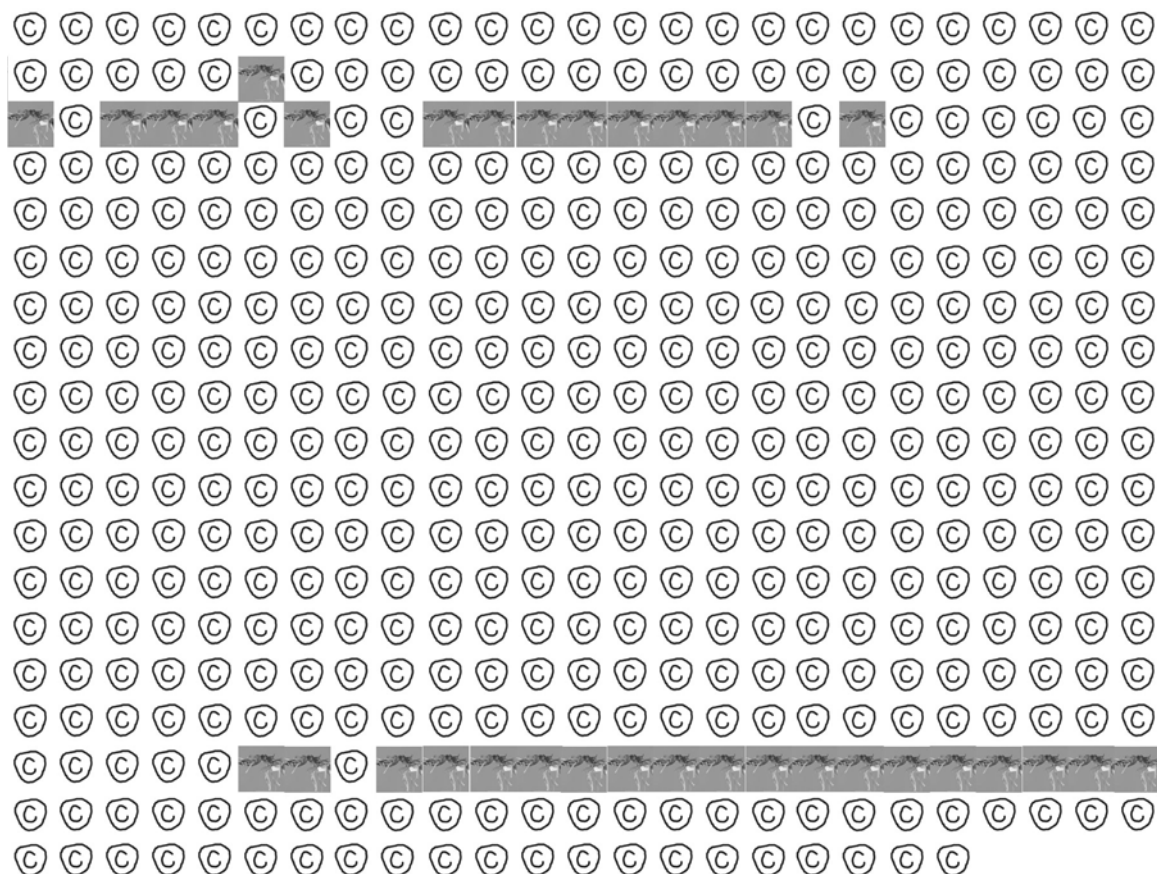
Creative Commons. <http://creativecommons.org>

CBD, 1992. Convention on Biological Diversity. <http://www.biodiv.org/convention/articles.asp>

Berners-Lee, T., Hendler, J., and Lassila, O. 2001. The semantic Web. *Scientific American* 285, 34 – 43

ICZN. International Code of Zoological Nomenclature. <http://www.iczn.org>

TDWG. Taxonomic Database Working Group. <http://www.tdwg.org>.



 open access; online

 copy-righted, no open access

2003

Figure 1. Access to ant systematics information. Each field represents one of the 476 new ant taxa described in the year 2003, including 472 species, for which descriptions were received and entered into the database by Feb. 18, 2005. The taxa are alphabetically listed and their identity can be seen at http://atbi.biosci.ohio-state.edu:210/hymenoptera/manage_lit.new_taxa_by_year?tnuid=152&the_year=2003 or <http://antbase.org>). Direct access to the original page of description as well as additional information for all the ant species, without the above mentioned cases, is available through antbase.org.