

LifeWatch Data Grant 2014

Filling the gaps in the World Register of Marine species (WoRMS)

Echinoidea

Final Report

Andreas Kroh



1. Data grant background

The aim of this grant project was to solve a number of important gaps identified for the Echinoidea group in WoRMS. Most pressing of these was the lack of links between the primary literature describing the extant species and the taxon pages. This is seen as a worthwhile addition to WoRMS in light of the fact that Echinoidea represent important members of benthic communities, due to which a wide range of researchers will have the need to access these data.

A first assessment had shown that c. 450 different publications account for the c. 1700 nominal extant species taxa that had been erected for Echinoidea during the last 250 years. These c. 1700 nominal taxa are represented by more than c. 3000 different name combinations in the WoRMS database of which c. 1000 are currently accepted. A small part of these taxa (those established after 1970) were already linked to the original descriptions by the editors, the majority, however, was still lacking.

The goal of this work was to provide verified references, including page numbers and references to illustrations for the original descriptions and to link those to full-text sources where possible (BHL, Archive.org, etc.).

2. Agreed deliverables (as specified in the Data Grant contract)

- verified references to original description (primary literature) for extant Echinoidea in WoRMS
- addition of links to online fulltext sources (e.g. BHL, Archive.org, Gallica,...) for these original description references, where available

3. Results of the project:

- original descriptions of 1465 Echinoidea basionyms published before 1970 were first traced in the literature and verified in the primary sources themselves
- in addition, as a consequence of the re-examining the primary sources a number of corrections of taxon pages and taxon additions were done – 339 pages were revised, 153 new taxa added
- 320 of 325 primary sources are linked to full text sources now

Additional: post-1970 original descriptions were already added prior to the project by me based on my *Index of Living & Fossil Echinoids 1970-2008* (Kroh 2008), which had been personally compiled from the primary sources originally. It was thus not deemed necessary to check these taxa again. Species published after 2008 are being continuously added to WoRMS by myself.

4. (Brief) description of the work/methodology

The following workflow was used:

- 1) pre-1970 basionyms of extant echinoid were identified and exported from the personal database of the taxonomic editor
- 2) citations of the original description of these taxa were searched for in the literature (using mainly Mortensen's *A Monograph of the Echinoidea*, Lambert & Thiéry's *Essai de Nomenclature Raisonnée des Échinides* and A. Agassiz's *Revision of the Echini*, plus other sources where necessary)
- 3) the citations were sorted according to the source papers, which in turn were grouped according to journal they were published in (to speed up tracing the actual papers in full-text sources and libraries)
- 4) BHL, archive.org, Gallica, Google Books, and a variety of other sources for digitized literature were searched for the actual papers – for this a colleague (J. Herler) working as a freelancer was employed to distribute the workload
- 5) Once verified the corrected citation of the primary source was entered in WoRMS (or updated if already present) and links to the full-text sources were added (mainly by J. Herler)
- 6) Primary sources not located in these digital repositories were first searched for in the extensive digital reprint collection of the editor and in Austrian libraries (pdf were uploaded whenever copyright limitations allowed it)
- 7) At the end of this process only about 20 primary sources had not been seen in the original yet – these were acquired using the help of the NHMW library team, the VLIZ library team and various international colleagues (M. Reich, S. Stöhr, A. Ziegler, and others).
- 8) Citation details for the basionyms and taxon authorities were verified using the primary sources
- 9) Errors in the WoRMS database identified during this process were corrected using the web interface.
- 10) The verified original description citation details were assembled in an Excel-table and handed to the Data Management Team for upload in WoRMS (File structure: ScientificName [as a control], AphiaID, SourceID, Pages [consisting of page range and reference to illustrations where applicable]) – file attached

5. Problems encountered and how it was solved (or expected solutions).

The workflow described above was relatively straightforward and few real problems were encountered. These related mainly to difficulties with locating rare works and the fact that a considerable number of citations of original descriptions taken from the literature were simply erroneous and seem to have been copied again and again without verification. In such cases tracing the first use of the name often proved to be a very time-consuming paper chase. Other difficulties arose from the common usage of pre-Linnaean authorities by echinoderm workers till about 1950 – here finding the first post-Linnaean usage that made a particular name available was sometimes difficult.

Another factor that had been seriously underestimated, however, was the amount of corrections necessary in WoRMS itself – mainly concerning publication dates, authorities, and spelling of taxa. In addition, a surprising number of basionyms were missing, since the original combinations were only used in the original descriptions, which apparently rarely were consulted by latter authors.

6. Other: remarks, suggestions, other information, bibliography, ...

Tracing the original descriptions in the primary sources did not only lead to a completion of WoRMS for this aspect, but also proved to be a highly efficient quality control mechanism. Numerous small errors were uncovered and corrected which had sneaked into WoRMS due to reliance on trusted, but secondary sources during initial data entry.

A big asset for similar projects in other taxon groups would 1) be the development of **a Webinterface that would allow to see all the names connected to a particular basionym** and 2) a website providing a chronological listing of correct citations of the major late 18th and early 19th century natural history books (e.g. the various *Systema nature* offshoots, and works like the *Encyclopedie methodique*, the *Dictionnaire naturelle*, etc.) together with links to fulltext sources. Many of these are indeed available, but being multi-volume works often originally issued in multiple deliveries (*livraisons*) it is often very time consuming to find the particular volume of interest and its correct publication details.