

Publishing Australian Marine Data to OBIS: Twenty Years of Lessons Learnt

Dave Watts[‡], Katherine Tattersall[‡], Sachit Rajbhandari[‡]

[‡] CSIRO, Hobart, Australia

Corresponding author: Dave Watts (dave.watts@csiro.au)

Abstract

In 2003, the [Australian Antarctic Data Centre](#) published the first Australian dataset of seabirds from the Southern Ocean to [OBIS](#) (Ocean Biodiversity Information System) via [DIGIR](#) (Distributed Generic Information Retrieval). The dataset initially had 17 fields with an emphasis on counts of individuals. Standards evolved and with the development of the [IPT](#) (Integrated Publishing Toolkit) by [GBIF](#) (Global Biodiversity Information Facility) around 2008, large datasets could be published. OBIS subsequently adopted the IPT as the preferred publishing tool for providers to use. In 2016, the Darwin Core Event core with the [OBIS Extended Measurements and Facts extension](#) was released (De Pooter et al. 2017), meaning that richer and more comprehensive datasets could be published via the IPT. It is only recently that the biological aggregators (e.g., OBIS, GBIF) are looking at enhancing functionality to report this data.

The [Australian OBIS Node](#) (OBIS-AU), hosted by [CSIRO NCMI](#) (the Commonwealth Science and Industrial Research Organisation National Collections and Marine Infrastructure Business Unit) now manages an [Australian region marine biodiversity IPT](#) with 30 million records from over 450 datasets. In the last 12 months, using the GBIF [DNA Derived Data Extension](#), the OBIS-AU Node has published extensive eDNA datasets to OBIS with sequences and DNA related metadata.

OBIS-AU has developed tools and procedures to ensure that data is of the best possible quality before it is published. Issues covered include preventing the duplication of data, preserving context, enhancing data once published with improvements in publication schemas, matching taxa, and identification of temporal or spatial errors.

Keywords

data publishing, marine biodiversity

Presenting author

Dave Watts

Presented at

TDWG 2023

Conflicts of interest

The authors have declared that no competing interests exist.

References

- De Pooter D, Appeltans W, Bailly N, Bristol S, Deneudt K, Eliezer M, Fujioka E, Giorgetti A, Goldstein P, Lewis M, Lipizer M, Mackay K, Marin M, Moncoiffé G, Nikolopoulou S, Provoost P, Rauch S, Roubicek A, Torres C, van de Putte A, Vandepitte L, Vanhoorne B, Vinci M, Wambiji N, Watts D, Klein Salas E, Hernandez F (2017) Toward a new data standard for combined marine biological and environmental datasets - expanding OBIS beyond species occurrences. Biodiversity Data Journal 5 <https://doi.org/10.3897/bdj.5.e10989>