



Collaborative Effort Towards a FAIR and OPEN Indo- Pacific Pollen Database (IPPD)

RESEARCH PAPER

ANNIKA V. HERBERT 

SIMON G. HABERLE 

HENRY MUNACK 

ALEXANDRU T. CODILEAN 

*Author affiliations can be found in the back matter of this article

]u[ubiquity press

ABSTRACT

The Indo-Pacific Pollen Database (IPPD) is a collection of modern and fossil Australian, Southeast Asian, and Oceanian pollen assemblages collected from Quaternary period depositional contexts. This compilation has been in the making since the late 1980s and is now being made available as a constituent database of Neotoma, thereby providing much needed palaeo-data coverage of the previously often underrepresented Indo-Pacific region. In addition, the IPPD is also being hosted by the OCTOPUS database, which integrates archaeology, sedimentology, geomorphology, charcoal, and non-human vertebrate fossil collections in a thoroughly Australasian context. These collections are all available in the same flattened table format, thereby providing regional researchers with multiple data types in an easily accessible manner. This will enable new avenues of cross-disciplinary research in the region. Here, we explore the collaborative efforts involved in compiling the IPPD and integrating it into OCTOPUS, as well as provide background information on both. Further, we consider future opportunities and the importance of FAIR, CARE, and OPEN data integration.

CORRESPONDING AUTHOR:

Annika V. Herbert

ARC Centre of Excellence for Australian Biodiversity and Heritage (CABAH), Australia; School of Culture, History and Language, ANU College of Asia and the Pacific, The Australian National University, Canberra, ACT 2600, Australia; ARC Centre of Excellence for Indigenous and Environmental Histories and Futures, Australia
annika.herbert@anu.edu.au

KEYWORDS:

Palaeoecology; pollen; database; FAIR; CARE; OPEN

TO CITE THIS ARTICLE:

Herbert, AV, Haberle, SG, Munack, H and Codilean, AT 2025 Collaborative Effort Towards a FAIR and OPEN Indo-Pacific Pollen Database (IPPD). *Data Science Journal*, 24: 5, pp. 1–9. DOI: <https://doi.org/10.5334/dsj-2025-005>

1. INTRODUCTION

Pollen analysis has been a key asset for understanding past vegetation dynamics in the Indo-Pacific region for decades (e.g., [Kershaw, 1970](#); [Hope, 1976](#)). As such, the number of pollen records from the region has been steadily increasing, eventually leading to the creation of a regional database, the Indo-Pacific Pollen database (IPPD; [Hope et al., 1999](#)). There have been several collaborative efforts in the region in the past to do with palaeoecology more broadly, a notable example being OZ-INTIMATE ([Turney et al., 2006](#); [Reeves et al., 2013](#)), a working group under the world-wide INTIMATE framework, with the aim to synthesise ice-core, marine and terrestrial proxy data for the period 30-8 ka BP ([Walker et al., 2001](#)). Another regional branch of a global initiative is the Aus2k working group ([Dixon et al., 2017](#)), a regional effort within the larger PAGES2k initiative ([PAGES2k Consortium, 2013](#)). These are large, international collaborations aimed at establishing a baseline for Earth's climate variability ([Masson-Delmotte et al., 2013](#)). Regional networks were established as part of this effort to fully utilise local expertise and highlight regional research ([Dixon et al., 2017](#)).

Despite these efforts, pollen data from the Indo-Pacific region have been scattered, poorly structured as a whole, and generally underrepresented in global databases in the past ([Gajewski, 2008](#)). This includes the online palaeoecology database Neotoma (<https://www.neotomadb.org/>), a global benchmark of palaeoecological databases, which up until recently only contained a handful of pollen records from the Indo-Pacific region. Due to Neotoma's quality and reputation, it was decided that we approach them regarding the hosting of this compilation. To that end, an international collaboration was established, with funding provided by the Australian Research Council Centre of Excellence for Australian Biodiversity and Heritage (CABAH). As a result, the IPPD is now a constituent database in Neotoma, with just under 180 associated sites at time of writing ([Herbert et al., 2024](#)).

To ensure palaeo-scientists in Australasia are aware of and have full access to the IPPD, we made it also available via OCTOPUS, a web-enabled database that includes several Sahul focused collections, and now incorporates a fresh collection of Australasian charcoal data as well as the IPPD. This provides the Australasian palaeo-community with a one-stop-shop of palaeo-data in a thoroughly Australasian context. OCTOPUS provides a way for these regional collections to be interoperable with each other, making it easy for scientists and non-scientists alike to access. Here, we will explore how this valuable resource has been developed, how it is managed and used, and what the future plans are.

2. THE IPPD ASSETS

The Indo-Pacific Pollen Database (now IPPD, but formerly known as INDOPAC, [Hope et al., 1999](#)) was originally developed and coordinated by Prof Geoff Hope (Australian National University) in the late 1980's as part of an emerging global effort to create a universally accessible database of pollen results (NEOTOMAdb, [Williams et al., 2018](#)). Employing newly developed pollen analytical tools such as TILIA ([Grimm, 1991](#)) and in collaboration with the programs author, Prof Eric Grimm (Illinois State Museum, [Graham et al., 2021](#)), the IPPD was envisaged to become a constituent part of the global NEOTOMAdb, with the potential to document the over 700 existing pollen sites from India, SE Asia, Australia, and the Pacific islands across to Easter Island. In cases where the raw data was no longer available, and after receiving permission from the researchers, the data was digitised from publications or theses. Updated age-depth models were produced upon integration, often through consultation with the associated researchers to preserve integrity. After data collection, we harmonised the IPPD content by screening, re-organising, and then uploading it to the Neotoma database ([Birks et al., 2023](#)). In this way, the data, now part of the IPPD, have been entered by a trained data steward (A.V.H.) through Neotoma's invariable TILIA software. Consequently, the IPPD became a constituent database of Neotoma's online palaeoecological data ecosystem, including pollen, diatoms, plant macrofossils, charcoal, insects, geochronological data, and much else besides ([Williams et al., 2018](#)).

To date, there are approximately 280 pollen sites, many containing multiple cores/units ([Figure 1](#)), with complete pollen records with related metadata available in the IPPD ([Herbert et al., 2024](#)), though efforts to integrate records from countries throughout the Indo-Pacific

region continue (Herbert *et al.*, 2024; Williams *et al.*, 2018). Each site is a locality associated with rich and detailed metadata, including coordinates, elevation, descriptions of sediments, radiocarbon dates, and other relevant information related to the site, including any published articles. Each core–or unit–within a site is associated with all available dating information so that each user can construct new age-depth models, should they so wish. Every sample in every core, or surface sample, is captured separately and is associated with relevant information, such as depth, age, sample type, and count type for the pollen data, recorded separately. The count type is either raw or percentage and may have been digitised from published pollen diagrams in instances where the original data was no longer available.

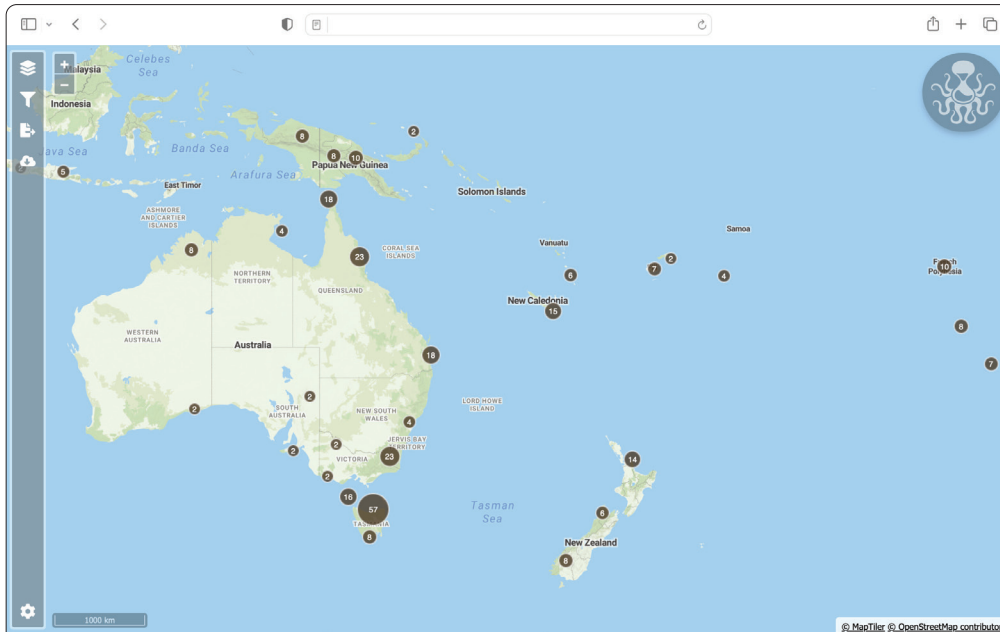


Figure 1: Geographical distribution of the IPPD’s data assets across the Indo-Pacific region. The circles (and numbers, if applicable) represent IPPD pollen datasets, i.e., the entity of pollen samples from one and the same collection unit (e.g. a core, excavation etc., cf. Munack *et al.*, 2023).

3. METHODOLOGY

When a new record is added to the IPPD, there are certain criteria it needs to meet, the most important of which relates to dating. The dating needs to be robust (multiple dates), and the raw dates available. The latter ensures new age models can be constructed for standardisation purposes and that any radiocarbon dates can be recalibrated when updated calibration curves are published. A principal researcher associated with each record is also important to ensure quality control and to enable appropriate acknowledgements. Each new site may also need to be reformatted upon entry into Neotoma. This involves a simple change in file format to Tilia, usually from Excel, as well as a more complicated taxon harmonisation process (Birks *et al.*, 2023). This process involves standardising the names of the pollen taxa found at the site. Some of the sites we have entered were analysed several decades ago and in the interim, the taxonomic placement of some plants have changed. In addition, with the IPPD representing a region with such few sites in Neotoma previously, and having such a unique flora, hundreds of taxa needed to be verified and entered in the Neotoma taxonomic structure as part of the integration project (Herbert *et al.*, 2024). These taxa are now a permanent part of Neotoma, thereby simplifying the process for researchers adding their records to the IPPD constituent database in the future. Upon upload, each record is assigned a unique DOI by the Neotoma IT team. This DOI links to the landing page for the record, detailing extended metadata, including relevant publications.

Upon entry into Neotoma, each site is also assigned a unique Neotoma Site ID, and each individual record (e.g., core) is assigned a unique Neotoma Dataset ID. All these identifiers ensure the data is Findable, as stated in the FAIR Guiding Principles (Wilkinson *et al.*, 2016). Having both a database ID and a DOI also enables the data to be highly Accessible in an open and free manner. In addition, all NEOTOMAdb protocols are freely available on their website, along with a clear data usage license, thereby making the data Interoperable and Reusable according to the FAIR principles (Wilkinson *et al.*, 2016).

Discussions are currently ongoing as to how best apply the CARE principles (Carroll *et al.*, 2020) to data from the IPPD. These discussions are being conducted with several Indigenous partners under the newly established Centre of Excellence for Indigenous and Environmental Histories and Futures (CIEHF). The discussions also involve Neotoma leadership, as any changes being made to data availability or presentation will necessarily involve the NEOTOMAdB. The starting off point for these discussions is the template currently used by NEOTOMAdB, where each record in the North American Pollen Database includes information on the Native lands each record was sourced from, thereby recognising Native rights and interests. In Australia, this process is in the early stages but have started with the understanding of the Collective benefit (Carroll *et al.*, 2020) of making data sourced from Indigenous lands available to everyone.

OCTOPUS features a simple, user-friendly interface in which the user can easily choose which datasets to include on an interactive map. NEOTOMAdB has a similar feature, but the simple layer display in OCTOPUS makes it easier to use. For both databases, sites can be selected from the interactive map, which will then show users what datasets are available for the site. These datasets can then be downloaded, with OCTOPUS providing a wider range of file formats than NEOTOMAdB.

4. THE IPPD IN OCTOPUS

OCTOPUS-v.2.3, since the 2024 release of the IPPD and SahulChar (Rehn *et al.*, 2024)– is a web-enabled database compliant with the Open Geospatial Consortium (OGC) standards. It allows users to visualize, query, and download data on cosmogenic radionuclide, luminescence, and radiocarbon ages, as well as denudation rates associated with erosional landscapes, Quaternary depositional landforms, and palaeoecological, palaeoenvironmental, and archaeological records. The database also features supplementary geospatial data layers in both vector and raster formats. Following the FAIR (Wilkinson *et al.*, 2016) data principles, and licensed CC-BY 4.0, it is built on open-source software and hosted on the Google Cloud Platform. Users can access the data through a custom web interface (<https://octopusdata.org/>), via desktop GIS applications (QGIS recommended), or any other software that support OGC data access protocols (e.g. R language). For a detailed OCTOPUS database overview, see Codilean *et al.*, 2022, and for details on how to access the database via R or QGIS, see Supplementary information.

Beyond the IPPD, OCTOPUS v.2.3 contains six major data collections: the Sahul focused SahulArch (archaeological records, Saktura *et al.*, 2023), SahulChar (palaeofire records, Rehn *et al.*, 2024), SahulSed (sedimentary records, Codilean *et al.*, 2022) and FosSahul (non-human vertebrate fossil records, Peters *et al.*, 2019); the latter being a not actively maintained partner collection. Further, two global collections – CRN Denudation (CRN denudation rates) and ExpAge (CRN exposure ages) – belong to OCTOPUS' data assets (Codilean *et al.*, 2022). Having all these collections available through one source and being able to access those data through multiple platforms maximises the usability and interoperability of OCTOPUS.

Prior to the technical integration of the IPPD with OCTOPUS, it was necessary to achieve integration at the semantic level. Due to the generally similar OCTOPUS and Neotoma data models, the migrated IPPD nicely fits into OCTOPUS' ordinal semantic database backbone–MetaSite, Site, Unit, AnalysisUnit, Sample, and Observation. However, some IPPD-specific tables had to be introduced to cope with the ingested migrated data structure as there were IPPD only tables (Figure 2, prefix `ippd_`), structurally unaltered – though some of them truncated– Neotoma tables (Figure 2, `neo_` prefix), and the regional 'cabah_AnalysisUnit' table.

Detailed table descriptions, their keys, constraints and relations can be found in OCTOPUS' thorough online database documentation at <https://octopus-db.github.io/documentation/> (Munack *et al.*, 2023). An interactive database schema is available at <https://octopus-db.github.io/schema/> (Munack and Codilean, 2023).

Despite data being stored in a relational database, to provide access to this data as geospatial layers via an interactive map interface, and to enable data manipulation through the WFS protocol, each data collection is served to GeoServer as a static, flat data table. The current version of GeoServer does not support dynamically generated PostgreSQL virtual tables



Figure 2 Outline of the IPPD's integration with OCTOPUS.

Each tile depicts a definable database partition (labelled bold) with relevant table names listed below. 'Global', 'Regional', 'Reference', and 'Neotoma' tiles feature only those tables that 'The IPPD' has relations to. Table names with an asterisk (*) refer to OCTOPUS' semantic database backbone, any other table names belong to parent or lookup tables that provide input for their hierarchical children.

(known as views). Therefore, static flat data tables are created to function as the equivalent of views. When users download IPPD data from OCTOPUS, they receive point geospatial data files accompanied by associated attribute tables. Direct connections to the PostgreSQL/PostGIS database are also available upon request. The above (and Figure 2) is therefore a description of the background table structure that is not visible to the user. The user sees a simple table containing basic site information such as what records are available for the site and in which collection they belong, as well as the age range of each record and any associated publications.

Providing an online repository of multiple palaeo-datasets in an easily accessible and consistent format for the Australasian region means OCTOPUS is a powerful tool for researchers in the region. Having the IPPD be part of this means it can be targeted at the researchers most likely to use it and benefit from it the most.

5. DATA MANAGEMENT

5.1 CURRENT PRACTICES

All data in the IPPD have been provided by the relevant researchers who gave their permission for them to be shared publicly. To ensure the efforts of these researchers are fully recognised, all associated publications are fully cited where possible. If the data are unpublished, this is noted, along with the name of the researchers responsible for the data. As part of the Neotoma workflow, all persons associated with the data are listed, including data collectors, data processors, and age model creators. All records are also provided with a unique DOI when they are uploaded, which can be cited along with relevant publications, thereby creating a citable product for unpublished data.

The IPPD workflow is fully transparent and outlined in Figure 3. It has been designed to ensure that the IPPD follows the FAIR principles, making the data Findable, Accessible, Interoperable and Reusable (see details above and Wilkinson *et al.*, 2016). With much of the data being gathered from Indigenous lands, it is also important to ensure that this project follows the CARE principles for Indigenous Data Governance—Collective benefit, Authority to control,

Responsibility, and Ethics (Carroll et al., 2020). This is an ongoing effort to be undertaken with Indigenous partners. The data in the IPPD come from the unceded lands of close to 90 Indigenous groups in Australia alone. As such, discussions regarding how best to display and store these data in everyone’s best interests are complex and will take time.

A key part of the IPPD workflow is interaction with the Australian palaeo-science community (Figure 3). Performing outreach activities such as workshops, seminars, and conference presentations raises the community awareness of the IPPD and how they can benefit from it and/or contribute to it. This approach has led to multiple community members reaching out for assistance in getting their own records integrated into the IPPD, as well as multiple new data stewards trained in Neotoma uploading and verification procedures.

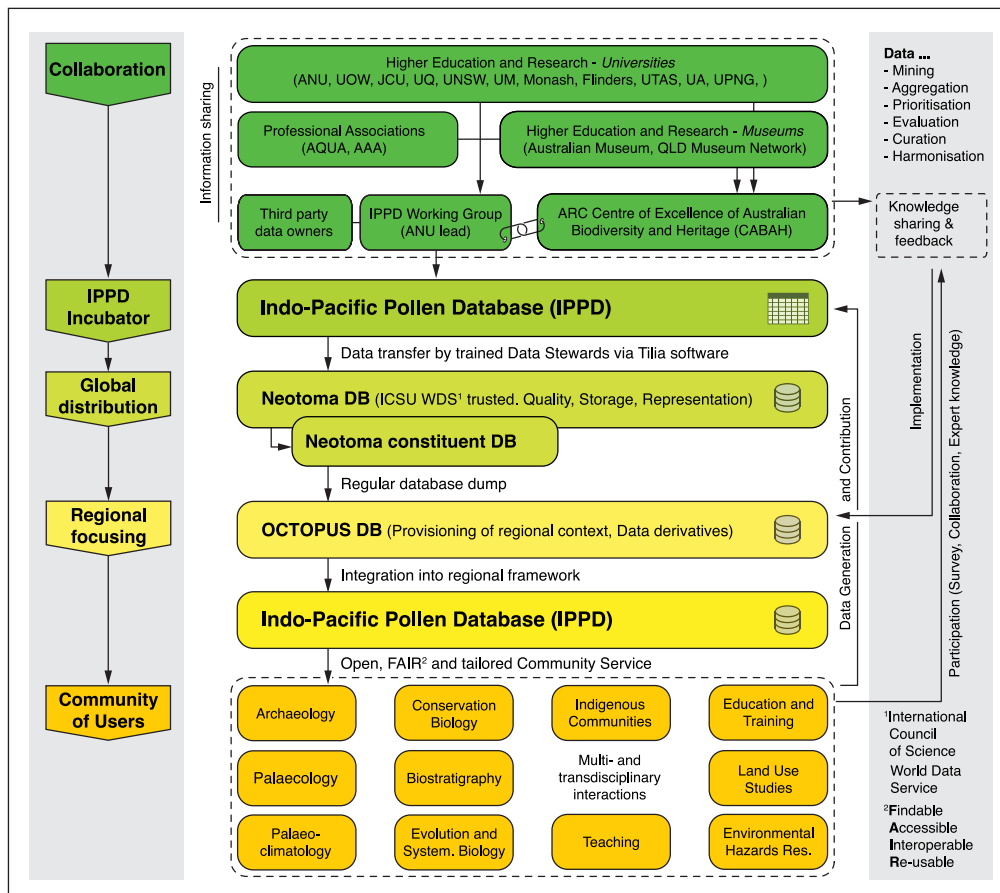


Figure 3 IPPD workflow. Description of the general IPPD workflow, from collaborators providing data to end users.

5.2 FUTURE PLANS/FUTUREPROOFING

Proponents of CARE principles for large databases, such as Carroll et al. (2020), Kukutai and Taylor (2016), and O’Brien et al. (2024), highlight that a key responsibility for researchers and curators of palaeoecological records (including pollen records) will be to critically assess the role played by these large datasets in making data about Indigenous Peoples and their lands more findable and accessible when held by colonial governments or organisations (Carroll et al., 2020). In addition, the adoption of explicit CARE principles that preference Indigenous Data Sovereignty in the governance of palaeoecological data, such as in the IPPD, will be essential for establishing the rights of Indigenous Peoples to determine how data about them and their lands will be collected and used. As stated in Carroll et al. (2020, p. 6), “Indigenous Peoples must have access to data that support Indigenous governance and self-determination. Indigenous Peoples must be the ones to determine data governance protocols, while being actively involved in stewardship decisions for Indigenous data that are held by other entities”.

Making the IPPD publicly available is part of a plan to ensure future researchers, as well as Indigenous groups, have access to these valuable data. Both Neotoma and OCTOPUS are based on open standards and open-source software, thereby ensuring the IPPD will remain accessible

in the future as it is currently transferrable into multiple formats and will continue to adapt to future format options. The previous iteration of the database was stored on a hard drive and not futureproofed. Taking it online and onto internationally recognised databases will ensure a lasting legacy of these data—a legacy that will have its own challenges in making sure Indigenous Data Sovereignty is at the forefront of research project efforts in the future.

6. DISCUSSION AND FUTURE WORK

The Indo-Pacific Pollen Database (IPPD) is a valuable collection of modern and fossil pollen records, representing decades of work by dozens of researchers across the Indo-Pacific region. The effort to make this collection publicly available has been many years in the making, through a large international collaboration of palaeo-scientists. Getting the Australian palaeo-science community involved in this effort will help ensure the IPPD will have a lasting legacy and secure future updates.

Making the IPPD publicly available opens up countless opportunities for new and exciting research, as well as new avenues for collaborative research. With this addition of data from the Indo-Pacific region, the ambitious aim of creating a Global Pollen Database is nearing completion (Gajewski, 2008). Being able to analyse global vegetation patterns going back millennia is an enticing prospect that may lead to many new and exciting research ideas. Indeed, a number of papers have already emerged from this effort and demonstrate the potential of the IPPD to “fill the gap” in global syntheses that can now include data from the IPPD region (e.g., Mottl *et al.*, 2021; Birks *et al.*, 2023), and regional syntheses that improve our understanding of regional palaeoclimate (Cadd *et al.*, 2021; Herbert and Fitchett, 2021), Indigenous fire and land management (Mariani *et al.*, 2022, 2024), and biodiversity (Adeleye *et al.*, 2021).

OCTOPUS plays an integral part in ensuring the legacy of the IPPD, by providing targeted access for researchers interested in the Australasian region. Having the IPPD accessible in the same flattened structure as other regional datasets enables cross-disciplinary research and will strengthen collaborative efforts in the region for years to come.

ADDITIONAL FILE

The additional file for this article can be found as follows:

- **Supplementary file 1.** Alternative ways of OCTOPUS db data access. DOI: <https://doi.org/10.5334/dsj-2025-005.s1>

ACKNOWLEDGEMENTS

OCTOPUS database is supported by the Australian Research Council Centre of Excellence for Australian Biodiversity and Heritage (CABAH) and received funding from the University of Wollongong (UOW) and the Australian National Data Service (ANDS), now part of the Australian Research Data Commons. AVH is currently supported by the Australian Research Council Centre of Excellence for Indigenous and Environmental Histories and Futures (CIEHF). Shaping and migrating the IPPD strongly benefited from the liberal knowledge transfer and advice of Jessica Blois, Simon Goring, Eric C. Grimm, and John W. (Jack) Williams. The authors gratefully acknowledge the helpful comments from two anonymous reviewers.

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS

SGH conceptualised and initiated the project, AVH wrote the first draft and is responsible for the IPPD, HM drafted the figures, HM and AC are responsible for OCTOPUS, and all authors reviewed and revised draft versions of the manuscript.

Annika V. Herbert  orcid.org/0000-0001-9322-043X

ARC Centre of Excellence for Australian Biodiversity and Heritage (CABAH), Australia; School of Culture, History and Language, ANU College of Asia and the Pacific, The Australian National University, Canberra, ACT 2600, Australia; ARC Centre of Excellence for Indigenous and Environmental Histories and Futures, Australia

Simon G. Haberle  orcid.org/0000-0001-5802-6535

ARC Centre of Excellence for Australian Biodiversity and Heritage (CABAH), Australia; School of Culture, History and Language, ANU College of Asia and the Pacific, The Australian National University, Canberra, ACT 2600, Australia; ARC Centre of Excellence for Indigenous and Environmental Histories and Futures, Australia

Henry Munack  orcid.org/0000-0002-3750-9674

ARC Centre of Excellence for Australian Biodiversity and Heritage (CABAH), Australia; ARC Centre of Excellence for Indigenous and Environmental Histories and Futures, Australia; School of Earth, Atmospheric and Life Sciences, University of Wollongong, Wollongong, Australia

Alexandru T. Codilean  orcid.org/0000-0003-1895-5633

ARC Centre of Excellence for Australian Biodiversity and Heritage (CABAH), Australia; ARC Centre of Excellence for Indigenous and Environmental Histories and Futures, Australia; School of Earth, Atmospheric and Life Sciences, University of Wollongong, Wollongong, Australia

REFERENCES

- Adeleye, M.A.** et al. (2021) 'Long-term drivers of vegetation turnover in Southern Hemisphere temperate ecosystems', *Global Ecology and Biogeography*, 30(2), pp. 557–571. Available at: <https://doi.org/10.1111/geb.13232>
- Birks, H.J.B.** et al. (2023) 'Approaches to pollen taxonomic harmonisation in Quaternary palynology', *Review of Palaeobotany and Palynology*, 319, pp. 1–6. Available at: <https://doi.org/10.1016/j.revpalbo.2023.104989>
- Cadd, H.** et al. (2021) 'A continental perspective on the timing of environmental change during the last glacial stage in Australia', *Quaternary Research*, 102, pp. 5–23. Available at: <https://doi.org/10.1017/qua.2021.16>
- Carroll, S.R.** et al. (2020) 'The CARE principles for indigenous data governance', *Data Science Journal*, 19(43), pp. 1–12. Available at: <https://doi.org/10.5334/dsj-2020-043>
- Codilean, A.T.** et al. (2022) 'OCTOPUS database (v.2)', *Earth System Science Data*, 14(8), pp. 3695–3713. Available at: <https://doi.org/10.5194/essd-14-3695-2022>
- Dixon, B.C.** et al. (2017) 'Low-resolution Australasian palaeoclimate records of the last 2000 years', *Climate of the Past*, 13, pp. 1403–1433. Available at: <https://doi.org/10.5194/cp-13-1403-2017>
- Gajewski, K.** (2008) 'The Global Pollen Database in biogeographical and palaeoclimatic studies', *Progress in Physical Geography: Earth and Environment*, 32(4), pp. 379–402. Available at: <https://doi.org/10.1177/0309133308096029>
- Graham, R.W.** et al. (2021) 'Memorial to Eric Christopher Grimm 1951–2020', *Memorials The Geological Society of America*, 50, pp. 31–35.
- Grimm, E.C.** (1991) TILIA and TILIA-GRAPH computer programs. Springfield: Illinois State Museum.
- Herbert, A.V.** and **Fitchett, J.M.** (2021) 'Quantifying late Quaternary Australian rainfall seasonality changes using the Poaceae: Asteraceae pollen ratio', *Quaternary Research*, 102, pp. 24–38. Available at: <https://doi.org/10.1017/qua.2020.18>
- Herbert, A.V.** et al. (2024) 'The Indo-Pacific Pollen Database – a Neotoma constituent database', *Climate of the Past*, 20, pp. 2473–2485. Available at: <https://doi.org/10.5194/cp-20-2473-2024>
- Hope, G.S.** 1976 The vegetational history of Mt Wilhelm, Papua New Guinea. *Journal of Ecology*, 64(2), pp. 627–663. Available at: <https://doi.org/10.2307/2258776>
- Hope, G.S., O'Dea, D.** and **Southern, W.** (1999) 'Holocene vegetation histories in the Western Pacific – alternative records of human impact', in I. Lilley, J.-C. Galipaud (eds.) *Le Pacifique de 5000 à 2000 avant le present: suppléments à l'histoire d'une colonisation*. Paris: IRD, pp. 387–404.
- Kershaw, A.P.** (1970) 'A pollen diagram from Lake Euramoo, north-east Queensland, Australia', *New Phytologist*, 69, pp. 785–805. Available at: <https://doi.org/10.1111/j.1469-8137.1970.tb02463.x>
- Kukutai, T.** and **Taylor, J.** (2016) 'Data sovereignty for indigenous peoples: Current practice and future needs', in T. Kukutai and J. Taylor (eds.) *Indigenous Data Sovereignty: Toward an agenda*, pp. 1–22. ANU Press, The Australian National University, Canberra, Australia. Available at: <https://doi.org/10.22459/CAEPR38.11.2016.01>
- Mariani, M.** et al. (2022) 'Disruption of cultural burning promotes shrub encroachment and unprecedented wildfires', *Frontiers in Ecology and the Environment*, 20, pp. 292–300. Available at: <https://doi.org/10.1002/fee.2395>
- Mariani, M.** et al. (2024) 'Shrub cover declined as Indigenous populations expanded across southeast Australia', *Science*, 386, pp. 567–573. Available at: <https://doi.org/10.1126/science.adn8668>

- Masson-Delmotte, V.** et al. (2013) *Information from paleoclimate archives*. Cambridge: Cambridge University Press.
- Mottl, O.** et al. (2021) 'Global acceleration in rates of vegetation change over the past 18,000 years', *Science*, 372(6544), pp. 860–864. Available at: <https://doi.org/10.1126/science.abg1685>
- Munack, H.** and **Codilean, A.T.** (2023) OCTOPUS database schema (online). <https://octopus-db.github.io/schema/> (accessed 26-06-2024).
- Munack, H., Rehn, E., Saktura, W.M.** and **Codilean, A.T.** (2023) OCTOPUS database documentation (online). <https://octopus-db.github.io/documentation/> (accessed 26-06-2024).
- O'Brien, M.** et al. (2024) 'Earth science data repositories: Implementing the CARE principles', *Data Science Journal*, 23(37), pp. 1–29. Available at: <https://doi.org/10.5334/dsj-2024-037>
- PAGES2k Consortium**, et al. (2013) 'Continental-scale temperature variability during the past two millennia', *Nature Geoscience*, 6, pp. 339–346. Available at: <https://doi.org/10.1038/ngeo1797>
- Peters, K.J.** et al. (2019) 'FosSahul 2.0, an updated database for the Late Quaternary fossil records of Sahul', *Scientific Data*, 6, pp. 272. Available at: <https://doi.org/10.1038/s41597-019-0267-3>
- Reeves, J.M.** et al. (2013) 'Climate variability over the last 35,000 years recorded in marine and terrestrial archives in the Australian region: An OZ-INTIMATE compilation', *Quaternary Science Reviews*, 74, pp. 21–34. Available at: <https://doi.org/10.1016/j.quascirev.2013.01.001>
- Rehn, E.** et al. (2024) 'The SahulCHAR Collection: A palaeofire database for Australia, New Guinea, and New Zealand', *Earth System Science Data Discussions* [preprint], in review. Available at: <https://doi.org/10.5194/essd-2024-328>
- Saktura, W.M.** et al. (2023) 'SahulArch: A geochronological database for the archaeology of Sahul', *Australian Archaeology*, 89, pp. 1–13. Available at: <https://doi.org/10.1080/03122417.2022.2159751>
- Turney, C.S.M.** et al. (2006), 'Integration of ice-core, marine and terrestrial records for the Australian Last Glacial Maximum and Termination: A contribution from the OZ INTIMATE group', *Journal of Quaternary Science*, 21(7), pp. 751–761. Available at: <https://doi.org/10.1002/jqs.1073>
- Walker, M.J.C., Björck, S.** and **Lowe, J.J.** (2001) 'Integration of ice core, marine and terrestrial records (INTIMATE) from around the North Atlantic region: An introduction', *Quaternary Science Reviews*, 20, pp. 1169–1174. Available at: [https://doi.org/10.1016/S0277-3791\(00\)00164-5](https://doi.org/10.1016/S0277-3791(00)00164-5)
- Wilkinson, M.D.** et al. (2016) 'The FAIR Guiding Principles for scientific data management and stewardship', *Scientific Data*, 3, 160018. Available at: <https://doi.org/10.1038/sdata.2016.18>
- Williams, J.W.**, et al. (2018) 'The Neotoma Paleoecology Database, a multiproxy, international, community-curated data resource', *Quaternary Research*, 89, pp. 156–177. Available at: <https://doi.org/10.1017/qua.2017.105>

Herbert et al.
Data Science Journal
 DOI: 10.5334/dsj-2025-005

TO CITE THIS ARTICLE:

Herbert, AV, Haberle, SG, Munack, H and Codilean, AT 2025 Collaborative Effort Towards a FAIR and OPEN Indo-Pacific Pollen Database (IPPD). *Data Science Journal*, 24: 5, pp. 1–9. DOI: <https://doi.org/10.5334/dsj-2025-005>

Submitted: 31 July 2024

Accepted: 13 January 2025

Published: 27 January 2025

COPYRIGHT:

© 2025 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Data Science Journal is a peer-reviewed open access journal published by Ubiquity Press.