

## Conference Abstract

# Challenges in the Development and Curation of Species-plot Datasets in South Africa: The National Vegetation Database of phytosociological plots as a case study

Brenda Daly <sup>‡</sup><sup>‡</sup> South African National Biodiversity Institute, Cape Town, South AfricaCorresponding author: Brenda Daly ([b.daly@sanbi.org.za](mailto:b.daly@sanbi.org.za))

Received: 31 Jul 2019 | Published: 08 Aug 2019

Citation: Daly B (2019) Challenges in the Development and Curation of Species-plot Datasets in South Africa: The National Vegetation Database of phytosociological plots as a case study. Biodiversity Information Science and Standards 3: e38675. <https://doi.org/10.3897/biss.3.38675>

## Abstract

The South African National Biodiversity Institute is the custodian of numerous national level botanical and zoological datasets that have been collated over several decades and is mandated to ensure that taxonomic and ecological data are made available to the public through responsible data sharing. This study describes the nature of, and presents/discusses relevant standards for, the case study of the National Vegetation Database; the process adopted in the development of a vegetation-plot database; and current data management practices being undertaken in relation to the various stages of research data management. Phytosociological data is a record of vegetation abundance, richness, density and the associated environmental variables within a specified area or plot which usually includes a record of locality. The study aims to review the diversity of approaches in storing species-plot information in databases and to provide minimum data standards for these datasets.

The surveying, classifying, and mapping of vegetation enables monitoring of ecosystems and ultimately can lead to improved conservation planning and land management. A coordinated and integrated approach is therefore needed to record, rectify, and manage

these data and capture accurate metadata. Preliminary findings indicate that a lack of version control can impact the authenticity of the data if records are altered or deleted. Data affluence/abundance (currently comprised of 53 500 plots within 384 sample projects, totalling 1 064 770 species occurrence records) is a challenge because these data often differ in formats, varying methodologies, and metadata within these research projects. The curation of plot data requires a standardised approach in the different steps from data acquisition to provision of results. Species names need to coincide with currently accepted taxonomy, and although certain details are specific to a species-plot project depending on their research interest, various other data should be made consistent in terms of field names and formats to improve the quality of the resulting aggregated set of botanical records. All decisions to modify data records to achieve data consistency should be clearly explained in the metadata record for the dataset.

## **Keywords**

phytosociology, vegetation-environment plot data, database, data standards

## **Presenting author**

Brenda Daly

## **Presented at**

Biodiversity\_Next 2019

## **Hosting institution**

South African National Biodiversity Institute