Introducing ‘The bdverse’: a family of R packages for biodiversity data

Tomer Gueta†, Vijay Barve§, Thiloshon Nagarajah¶, Povilas Gibas¶, Yohay Carmel‡

† Department of Civil and Environmental Engineering, The Technion – Israel Institute of Technology, Haifa, Israel
§ Florida Museum of Natural History, Gainesville, United States of America
¶ Informatics Institute of Technology, Colombo, Sri Lanka
‡ Vilnius University, Vilnius, Lithuania

Abstract

The bdverse is a collection of packages that form a general framework for facilitating biodiversity science in R. We build it to serve as a sustainable and agile infrastructure that enhances the value of biodiversity data by allowing users to conveniently employ R, for data exploration, quality assessment, data cleaning, and standardization. The bdverse supports users with and without programming capabilities. It includes six unique packages in a hierarchal structure — representing different functionality levels (Fig. 1). Major features of three core packages will be highlighted and demonstrated: (i) bdDwC provides an interactive Shiny app and a set of functions for standardizing field names in compliance with Darwin Core (DwC) format; (ii) bdchecks is an infrastructure for performing, filtering and managing various biodiversity data checks; (iii) bdclean is a user-friendly data cleaning Shiny app for the inexperienced R user. It provides features to manage complete workflow for biodiversity data cleaning, including data upload; user input - in order to adjust cleaning procedures; data cleaning; and finally, generation of various reports and versions of the data.

We are now working on submitting the bdverse packages to rOpenSci software review, and as soon as the packages meet core requirements, we will officially release the bdverse. The bdverse project won the 2nd prize in the 2018 Ebbe Nielsen Challenge.
Keywords

biodiversity informatic, data quality, R

Presenting author

Tomer Gueta

Presented at

Biodiversity_Next 2019

Funding program

ISF grant No. 127/16

The Technion, Blumenstein family fund

Google Summer of Code program