

bddashboard : An infrastructure for biodiversity dashboards in R

Tomer Gueta[‡], Rahul Chauhan[§], Thiloshon Nagarajah^{||}, Vijay Barve^{¶,¶}, Povilas Gibas[□], Martynas Jočys[□], Rahul Saxena[«], Sunny Dhoke[»], Yohav Carmel[‡]

[‡] Department of Civil and Environmental Engineering, The Technion – Israel Institute of Technology, Haifa, Israel

[§] Galgotias University, Greater Noida, India

^{||} Informatics Institute of Technology, Colombo, Sri Lanka

[¶] Post Doctoral Researcher, Terrestrial Parasite Tracker TCN, West Lafayette, Indiana, United States of America

[¶] Florida Museum of Natural History, Gainesville, United States of America

[□] Vilnius University, Vilnius, Lithuania

[«] International Institute of Information Technology, Bhubaneswar, India

[»] Indian Institute Of Information Technology, Nagpur, India

Corresponding author: Tomer Gueta (tomer.gu@gmail.com)

Abstract

The [bdverse](#) is a collection of packages that form a general framework for facilitating biodiversity science in R (programming language). Exploratory and diagnostic visualization can unveil hidden patterns and anomalies in data and allow quick and efficient exploration of massive datasets. The development of an interactive yet flexible dashboard that can be easily deployed locally or remotely is a highly valuable biodiversity informatics tool. To this end, we have developed '[bddashboard](#)', which serves as an agile framework for biodiversity dashboard development. This project is built in R, using the Shiny package (RStudio, Inc 2021) that helps build interactive web apps in R. The following key components were developed:

Core Interactive Components The basic building blocks of every dashboard are interactive plots, maps, and tables. We have explored all major visualization libraries in R and have concluded that 'plotly' (Sievert 2020) is the most mature and showcases the best value for effort. Additionally, we have concluded that 'leaflet' (Graul 2016) shows the most diverse and high-quality mapping features, and DT (DataTables library) (Xie et al. 2021) is best for rendering tabular data. Each component was modularized to better adjust it for biodiversity data and to enhance its flexibility.

Field Selector The field selector is a unique module that makes each interactive component much more versatile. Users have different data and needs; thus, every combination or selection of fields can tell a different story. The field selector allows users to change the X and Y axis on plots, to choose the columns that are visible on a table, and to easily control map settings. All that in real-time, without reloading the page or disturbing the reactivity. The field selector automatically detects how many columns a plot needs and

what type of columns can be passed to the X-axis or Y-axis. The field selector also displays the completeness of each field.

Plot Navigation We developed the plot navigation module to prevent unwanted extreme cases. Technically, drawing 1,000 bars on a single bar plot is possible, but this visualization is not human-friendly. Navigation allows users to decide how many values they want to see on a single plot. This technique allows for fast drawing of extensive datasets without affecting page reactivity, dramatically improving performance and functioning as a fail-safe mechanism.

Reactivity Reactivity creates the connection between different components. The changes in input values automatically flow to the plots, text, maps, and tables that use the input, and cause them to update. Reactivity facilitates drilling down functionality, which enhances the user's ability to explore and investigate the data. We developed a novel and robust reactivity technique that allows us to add a new component and effectively connect it with all existing components within a dashboard tab, using only one line of code.

Generic Biodiversity Tabs We developed five useful dashboard tabs (Fig. 1): (i) the Data Summary tab to give a quick overview of a dataset; (ii) the Data Completeness tab helps users get valuable information about missing records and missing Darwin Core fields; (iii) the Spatial tab is dedicated to spatial visualizations; (iv) the Taxonomic tab is designed to visualize taxonomy; and (v) the Temporal tab is designed to visualize time-related aspects.

Performance and Agility To make a dashboard work smoothly and react quickly, hundreds of small and large modules, functions, and techniques must work together. Our goal was to minimize dashboard latency and maximize its data capacity. We used asynchronous modules to write non-blocking code, clusters in map components, and preprocessing and filtering data before passing it to plots to reduce the load. The 'bddashboard' package modularized architecture allows us to develop completely different interactive and reactive dashboards within mere minutes.

Keywords

biodiversity informatics, Shiny, bdverse

Presenting author

Tomer Gueta

Presented at

TDWG 2021

Funding program

Israel Science Foundation (grant 127/16)

Technion's Blumenstein family fund

Google Summer of Code program

References

- Graul C (2016) leafletR: Interactive Web-Maps Based on the Leaflet JavaScript Library. R package version 0.4-0. URL: <http://cran.r-project.org/package=leafletR>.
- RStudio, Inc (2021) shiny: Easy web applications in R. 1.6.0. URL: <http://shiny.rstudio.com>
- Sievert C (2020) Interactive web-based data visualization with R, plotly, and shiny. Chapman and Hall/CRC URL: <https://plotly-r.com>
- Xie Y, Cheng J, Tan X (2021) DT: A wrapper of the JavaScript library 'DataTables'. R package version 0.18. URL: <https://CRAN.R-project.org/package=DT>



Figure 1.
bddashboard's five generic tabs: (i) Data Summary; (ii) Data Completeness; (iii) Spatial; (iv) Taxonomic; and (v) Temporal.