Using R for an Antarctic Virtual Research environment

Hsun-Yi Hsieh (1), Ben Raymond (2), Jean_Olivier Irison (3), Sophie Mormede (4), Yi-Ming Gan (1), and Anton Van de Putte (1)

(1) Royal Institute of Natural Sciences, Belgium, Biodiversity & Ecosystems Data & Information centre, Belgium (hhsunyi@naturalsciences.be), (2) Australian Antarctic Division, (3) Maître de Conférences, Univ. Paris VI, Observatoire Océanologique, (4) National Institute of Water and Atmospheric Research

The Antarctic Biodiversity portal (www.biodiversity.aq) is a gateway to a wide variety of Antarctic biodiversity Information and tools. Launched in 2015 as SCAR-MarBIN and the register of Antarctic Marine Species (RAMS) the system has grown in scope from purely marine to also include terrestrial information.

Biodiversity.aq is an official product of the Scientific Committee on Antarctic Research (SCAR). This product is currently supported as one of the Belgian contributions to the European Lifewatch-ERIC (European Research Infrastructure Consortium). The goal of Lifewatch is to provide access to Distributed observatories/sensor networks; Interoperable databases, existing (data-)networks, using accepted standards; High Performance Computing (HPC) and Grid power, including the use of the start-of-art of the so-called Cloud and Big Data paradigms technologies; Software and tools for visualization, analysis, and modeling."

In this presentation, we focus on the analytical tools of biodiversity.aq. As a Virtual Research Environment, biodiversity.aq does not only provide access to Antarctic biodiversity data; it also provides access to analytical tools. The efforts include providing an inventory of existing tools and repositories but also developing new tools such as the R packages atlasR and biotaxa.

biotaxa is a visualization and analytical tool to study the accumulation curves of taxonomic discoveries over time. It aims to identify taxa that are relatively well and less discovered, and used the well-discovered taxa as the templates for predicting the species richness of less-discovered taxa and, together, the entire taxonomic richness of an interested region. The end product is an R product that allows users to use an example and their own datasets for visualization and modelling. The tool will be used in the Register of Antarctic Species (RAS) to learn the development curve of the taxonomic backbone. We collaborate with rOpenSci, an international community of R tools development, on this product.

We will present an overview of the biodiversity.aq virtual research environment and the usage of R packages and the app, and the prospects ahead.