

Rhakhis: A workflow for managing the WFO taxonomic backbone

Roger Hyam[‡], Alan Elliott[‡], William Ulate[§]

[‡] Royal Botanic Garden Edinburgh, Edinburgh, United Kingdom

[§] Missouri Botanical Garden, St. Louis, MO, United States of America

Corresponding author: Roger Hyam (rhyam@rbge.org.uk)

Abstract

In 2021, the [World Flora Online](#) (WFO) Council agreed that the team at the [Royal Botanic Garden Edinburgh](#) would take on the technical role of managing the [WFO Taxonomic Backbone](#) (WFO-TB). This presentation outlines the implementation of a system to manage the associated data and explores possible future developments.

The WFO-TB is a global consensus checklist of plants including bryophytes, pteridophytes, gymnosperms and angiosperms. The checklist data consists of two parts: facts concerning the nomenclatural acts that establish the names under the nomenclatural code, and consensus expert opinion on the placement of those names into a single, authoritative taxonomy. WFO-TB is unique in that it is both global in scope and curated by a large team of experts from multiple institutions. There are currently around 280 specialist contributors organised into 37 Taxonomic Expert Networks (TENs).

The primary function of the WFO-TB is to provide structure to the hundreds of thousands of descriptions, images and other pieces of Content that make up the [WFO's main web portal](#), but periodic snapshots of the taxonomic data are also made available as public downloads and published as the [WFO Plant List](#). The WFO Plant List is unique in that it remembers each version of WFO-TB that was published and links between them thus providing a stable, citable resource. In addition, data can be released in any format that researchers may find useful including through [ChecklistBank](#) and so be potentially incorporated into the [Catalogue of Life](#) and [GBIF](#) (Global Biodiversity Information Facility). All checklist data are released with a Creative Commons (CC) license of [CC0 so that others may freely build upon, enhance and reuse the works for any purposes without restriction under copyright or database law](#).

Prior to 2022, the WFO-TB was managed as part of the main WFO data resources supporting the WFO Portal by the team at Missouri Botanical Garden. Moving responsibility for hosting the backbone to the team at Edinburgh added significant resources to the project as a whole, but created technical challenges. The first requirement of the new system (called Rhakhis, a Greek form of rachis, the 'backbone' of

a leaf or inflorescence) was to demonstrate that it could feed data back to Missouri in a way that could be incorporated into the WFO infrastructure without causing disruption to the Content curation process. This was completed in early 2022, and, by June, the primary copy of the WFO-TB data had been entrusted to Rhakhis. The June 2022 version of the WFO Plant List and the WFO-TB was published using the previous system, switching to publication from Rhakhis in December 2022.

The system architecture of Rhakhis is quite simple. A [MySQL](#) database holds all the data and is exposed via a [GraphQL API](#) to a web-based user interface written in Javascript using [React and Bootstrap](#). User authentication is handled via a link to [ORCID](#) (Open Researcher and Contributor ID). Authorisation for editors is delegated hierarchically down the taxonomic tree. This provides the ability for TENs and the TEN Manager to oversee and manage the live data directly and to delegate authority to colleagues. The TEN Manager has access to a plain HTML bulk loader interface that enables the ingestion of CSV and files in [DwC-A](#) (Darwin Core Archive) format supplied by TENs as well as updates from other data sources, such as [IPNI](#) (International Plant Names Index) and [WCVF](#) (RBG Kew's World Checklist of Vascular Plants). Rhakhis is designed as a standalone data management tool for taxonomists involved in the project rather than a public website, and is hosted on a WFO server on the Google Cloud.

The WFO Plant List is run on a separate system and designed to be a performant public-facing website. Periodic snapshots of the WFO-TB are imported into a [Solr Index](#) that is then exposed via another GraphQL API as well as [Semantic Web](#)-compliant HTTPS URIs. A web-based user interface to WFO Plant List is implemented as part of the [Craft CMS](#) (Content Management System) that also runs the About pages of the WFO Web Portal, but it would be possible to build other interfaces to this data.

The periodic snapshots of the WFO-TB, which are published through the WFO Plant List, are archived in [Zenodo](#) and assigned a [DOI](#) (Digital Object Identifier), as well as contributed to Catalogue of Life ChecklistBank. The archive formats currently supported include: Darwin Core Archive and Catalogue of Life Data Package, but other formats will be considered in the future. We are interested in investigating the creation of data papers, e.g., through the [Biodiversity Data Journal](#) at [Pensoft](#), to provide additional accreditation for contributors and exposure of their data.

Keywords

World Flora Online, taxonomy, botany, taxonomists, Rhakhis

Presenting author

Roger Hyam

Presented at

TDWG 2022

Acknowledgements

The World Flora Online is a collaborative project. Any outputs are only possible because many participating individuals and institutions give freely of their time, skills and resources.

Funding program

Institute funded

Hosting institution

Royal Botanic Garden Edinburgh

Ethics and security

There are no ethical issues.

Author contributions

Roger Hyam wrote text. Alan Elliott and William Ulate contributed design and data throughout project as well as reviewed text.

Conflicts of interest

The authors declare no conflicts of interest.