



Freie Universität Berlin - Botanic Garden and Botanical Museum Berlin
(FUB-BGBM)

Agreed strategy for mapping local taxonomies (MS4)

Walter G. Berendsohn (FUB-BGBM)
Eckhard von Raab-Straube (FUB-BGBM)
Andreas Müller (FUB-BGBM)
Anton Güntsch (FUB-BGBM)

TETTRIs

Transforming European Taxonomy through Training, Research, and Innovations

Date: 21/02/2023 1 / 5Doc version: V03



Document: Milestone MS4
Version: v. 1.0
Date: 16/04/2024

Disclaimer: Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Date: 21/02/2023 2 / 5Doc version: V03

Document history:

Revision	Date	Created by	Short Description of Changes
0.1	31.3.2024	W.G. Berendsohn	Initial draft
1.0	16.4.2024	W.G. Berendsohn	Revised draft
2.0	23.5.2024	W.G. Berendsohn	Revised draft incorporating comments of WP2 members

MS4 Agreed strategy for mapping local taxonomies

This milestone report outlines the background of the mapping exercise, provides the basic strategy behind our activities, and points to the wiki pages where the ongoing work towards the project deliverables are documented (see footnotes).

Virtual aggregators of data on taxa, scientific names, and geographic distribution of organisms operate on global and regional levels, playing a normative role in all research concerning organisms and biodiversity. These aggregators may serve as the glue that binds together local data, ranging from spreadsheets maintained by individual researchers to large databases containing taxonomic checklists for countries, entire regions of the world [1] or taxonomic groups, especially when they work closely together (as currently planned by the TETTRIs 3PP project “TNLS - Taxonomic Name Linking Service” for several such aggregators).

The strategy of TETTRIs targets both the holders of local data and the aggregators themselves. It aims to enable [2] and motivate local users to verify their data with the aggregators and, optionally, to link to the aggregators’ services. Additionally, it seeks to instigate the development of services [3] at the aggregators’ side that facilitate usage and the establishment of such linkages.

Organisms are classified into taxa (e.g. species, genera, families), which are named according to fixed rules to enable the sharing of biodiversity information related to them. Taxa represent scientific concepts and develop as new data emerge, so that scientific names, which are tied to unique, name-bearing specimens, do not provide unambiguous identifiers for taxa [4]. Nevertheless, scientific names are the most commonly used tags for biodiversity information. Therefore, our strategy for mapping taxonomies must start with the names. Since variants of name strings exist, we promote the use of unique identifiers for names on the aggregators’ side to unambiguously relate all relevant names and their variants to information about them [5]. We further promote the resolvability of name IDs so that end-users can consult them to find the associated information on the aggregator’s side.

So, we will:

- Promote the use of name-matching services as a means of data quality control by documenting existing services[6] and their use in concrete use cases [7].
- Stimulate software developers to integrate name matching into their software (or to provide standard output that is easily used in matching services).
- Motivate aggregators to provide functional and useful services.

Our focus is on the following aggregators (primary, secondary, and lookup [8]):

- For the European context, PESI (eu-nomen) with its content-providing databases Index / Species Fungorum, Euro+Med PlantBase, Fauna Europaea, and ERMS. The unified services are provided by the Flanders Institute for de Zee (VLIZ) in Belgium, which also provides the global Ocean

Date: 21/02/2023 4 / 5 Doc version: V03



Biogeographic Information System OBIS and a number of taxonomic services in the context of the European LifeWatch ERIC.

- Globally, Catalogue of Life and the connected COL/GBIF ChecklistBank.
- Furthermore, we target taxonomic services that provide a reconciliation interface usable by tools such as OpenRefine [9], for which we can develop detailed workflows useful for people with and without a technical background. This currently includes Global Names Architecture, World Flora Online Plant List and the International Plant Name Index (see [6] for details).

[1] https://wiki.bgbm.org/tettris/Metadata_for_%22local_lists%22

[2] https://wiki.bgbm.org/tettris/End-user_workflows_for_name_matching

[3] https://wiki.bgbm.org/tettris/Aggregator_services_wish_list

[4] https://wiki.bgbm.org/tettris/Potential_caveats

[5] https://wiki.bgbm.org/tettris/Target_aggregator%27s_resolvable_name_ID

[6] https://wiki.bgbm.org/tettris/Existing_name_checking_mechanisms

[7] https://wiki.bgbm.org/tettris/Use_cases_for_name_matching

[8] https://wiki.bgbm.org/tettris/Global_or_regional_aggregators

[9] <https://openrefine.org/>