

The metadata configuration file

Appendix 1 to *NCBITaxonomy.jl* - rapid biological names finding and reconciliation

Timothée Poisot^{1,2} Rory Gibb^{3,4,5} Sadie J. Ryan^{6,7,8} Colin J. Carlson^{10,9}

¹ Université de Montréal, Département de Sciences Biologiques, Montréal QC, Canada ¹⁰ Center for Global Health Science and Security, Georgetown University Medical Center, Georgetown University, Washington, D.C., United States of America ² Québec Centre for Biodiversity Science, Montréal, QC, Canada ³ Centre on Climate Change and Planetary Health, London School of Hygiene and Tropical Medicine, London, UK ⁴ Centre for Mathematical Modelling of Infectious Diseases, London School of Hygiene and Tropical Medicine, London, UK ⁵ Current address: Centre for Biodiversity and Environment Research, University College London, London, UK ⁶ Emerging Pathogens Institute, University of Florida, Gainesville, FL, United States of America ⁷ School of Life Sciences, University of KwaZulu-Natal, Durban, South Africa ⁸ Department of Geography, University of Florida, Gainesville, FL, United States of America ⁹ Department of Microbiology and Immunology, Georgetown University Medical Center, Georgetown University, Washington, D.C., United States of America

Metadata overview

The `metadata.json` file, at the root of the project, is where all of the information that are not text are stored. This includes the manuscript title, affiliations, the abstract and keywords, and the citation style.

Note that the license is, by default, CC-BY 4.0. This is hardcoded in the template. There are no plans to support other CC licenses for now.

Manuscript metadata

The `title` field (a single string of text) stores the full manuscript title. The `keywords` array is a list of strings, which represents the keywords associated to the manuscript.

Affiliations

The information on authorship and affiliations are stored in an array of author blocks, with a *very specific* format. The authors are filled in the order in which they appear, and the affiliations are filled in the same way.

3.1. Author metadata An author is defined by *at least* two fields: `familyname` and `givennames`. In addition, you can (should) specify an `orcid` and an `email`. For some authors, you might add a `status` array, the values of which can be `corresponding` and `equal`.

3.2. Affiliations The affiliations are part of the author object, and are represented as an array of text. Note that the conversion to numbered affiliations is handled by the script, and that the affiliations that have the same string will be matched.

Abstract

This template supports three types of abstracts, indicated in the metadata file as `abstract`:

A regular abstract is defined as

```
"abstract": "A very long string"
```

An itemized abstract is an array of strings, each representing a bullet point:

```
"abstract": [  
  "Point 1",  
  "Point 2"  
]
```

A structured abstract is an object with key-value pairs :

```
"abstract": {  
  "Location": "Worldwide",  
  "Organisms": "Mammals"  
}
```

5

Citation style

The `citationstyle` key corresponds to the name, with `.csl` omitted, of a CSL stylesheet stored in the [citation style language](#) repository. Note that there is no difference between main and dependent styles, the build engine will take the correct steps to get the correct style. The default is `"citationstyle": "ecology-letters"`. There is a longer section about references management in the main text.

Bezanson, J. et al. 2017. [Julia: A Fresh Approach to Numerical Computing](#). - SIAM Review 59: 65–98.

Chamberlain, S. A. and Szöcs, E. 2013. [Taxize: Taxonomic search and retrieval in R](#). - F1000Research 2: 191.

Federhen, S. 2012. The NCBI taxonomy database. - Nucleic acids research 40: D136–D143.

Norman, K. E. A. et al. 2020. [Taxadb: A high-performance local taxonomic database interface](#). - Methods in Ecology and Evolution 11: 1153–1159.

Schoch, C. L. et al. 2020. NCBI Taxonomy: A comprehensive update on curation, resources and tools. - Database in press.