The `dwctaxon` R package for editing and validating taxonomic data
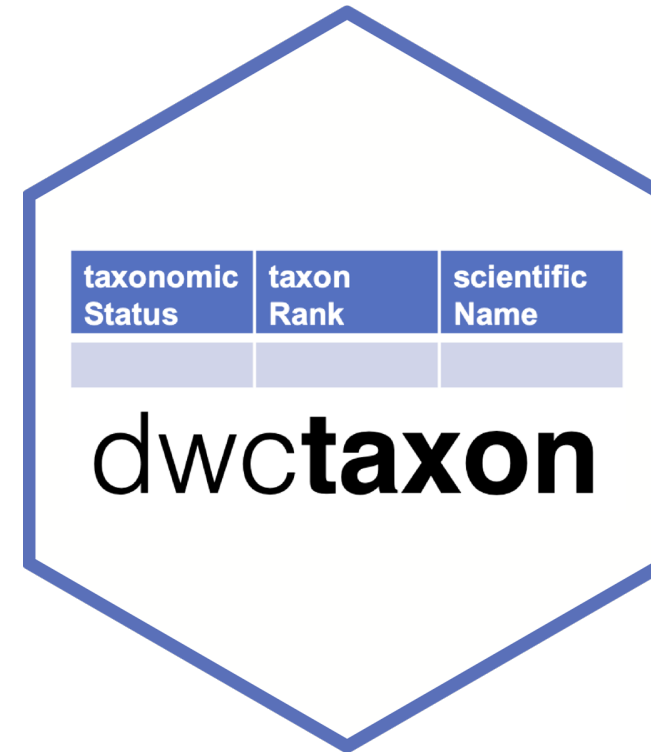
Joel H. Nitta[1]

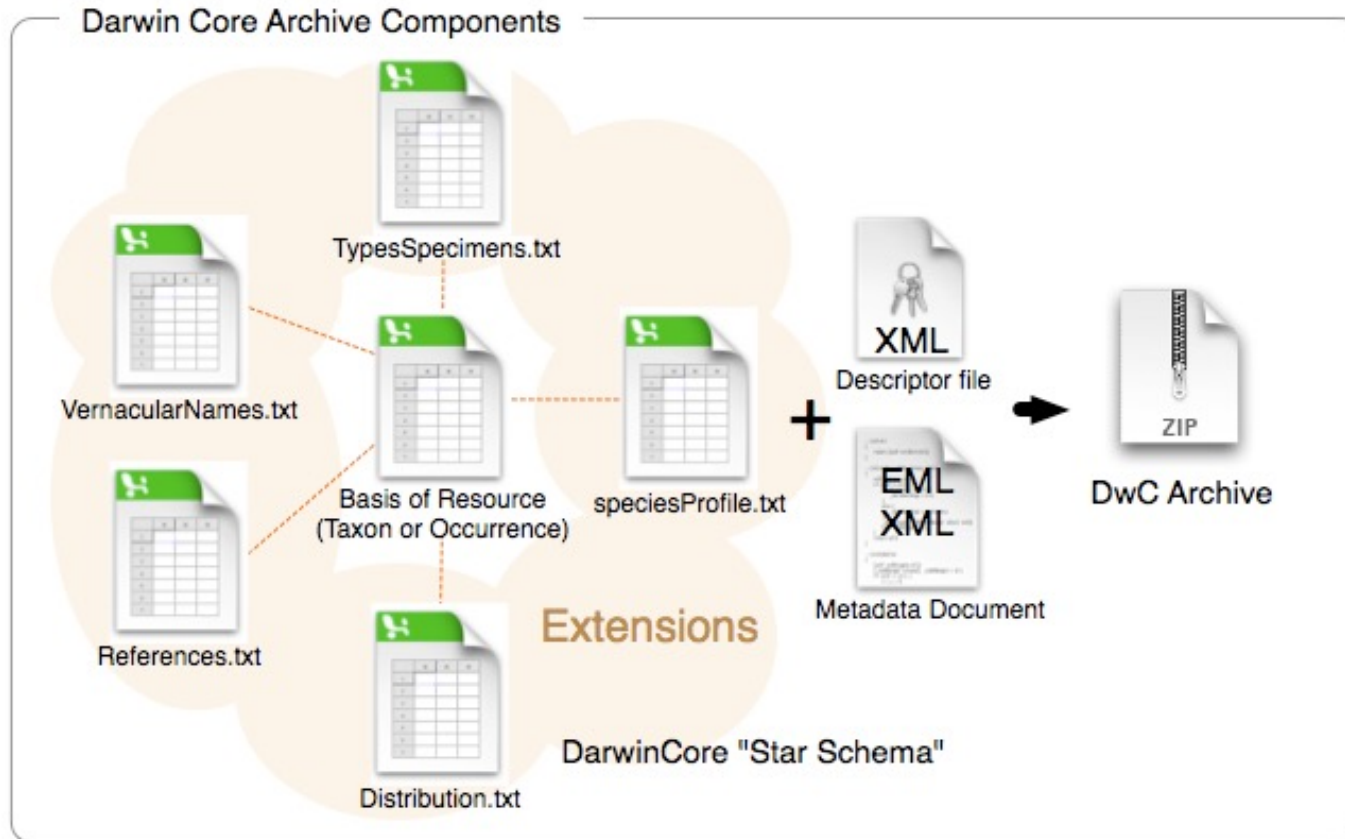Wataru Iwasaki[2]

[1]Chiba University, Japan

[2]The University of Tokyo, Japan

https://www.joelnitta.com

Darwin Core (DwC) is a **standard** for biodiversity data

`dwctaxon` helps <u>edit</u> and <u>validate</u> DwC taxonomic data

- Use cases
  - Taxonomic database management
  - Taxon checklists
  - Name standardization

Database of Vascular Plants of Canada (VASCAN)

Latest version published by Université de Montréal Biodiversity Centre on Apr 18, 2023

The Database of Vascular Plants of Canada or VASCAN (http://data.canadensys.net /vascan) is a comprehensive and curated checklist of all vascular plants reported in Canada, Greenland (Denmark), and Saint Pierre and Miquelon (France). VASCAN was developed at the Université de Montréal Biodiversity Centre and is maintained by a

Example: Vascular Plants of Canada
http://data.canadensys.net/ipt/resource.do?r=vascan

Editing taxonomic data is complicated!

Old version

- Accepted species 1: *Cephalomanes crassum*
  - Synonym: *Trichomanes crassum*
- Accepted species 2: *Cephalomanes densinervium*
  - Synonym: *Trichomanes densinervium*

New version (*C. crassum* → synonym of **C. densinervium**)

- Accepted species: **Cephalomanes densinervium**
  - Synonym 1: *Cephalomanes crassum*
  - Synonym 2: *Trichomanes crassum*
  - Synonym 3: *Trichomanes densinervium*

Need to account for **all** synonyms!

`dct_change_status()` handles synonym mapping **automatically**

```
1  # C. crassum → synonym of C. densinervium
2  dct_change_status(
3    tax_dat = filmies,
4    sci_name = "Cephalomanes crassum (Copel.) M. G. Price",
5    new_status = "synonym",
6    usage_name = "Cephalomanes densinervium (Copel.) Copel."
7  )
```

```
# A tibble: 4 × 4
  taxonID  acceptedNameUsageID taxonomicStatus scientificName
  <chr>    <chr>               <chr>           <chr>
1 54133783 54115098            synonym         Trichomanes crassum Copel.
2 54115097 54115098            synonym         Cephalomanes crassum (Copel.)
M.…
3 54133784 54115098            synonym         Trichomanes densinervium Copel.
4 54115098 <NA>                accepted name   Cephalomanes densinervium
(Copel…
```

Taxonomic data can be full of errors!

- Non-unique names

- Mistaken ranks

- Incorrect synonym mapping

- Etc...

https://media.npr.org/assets/img/2018/10/17/gettyimages-528623386_slide-564362f49c9eec2b225c747098405186a740c536-s600-c85.webp

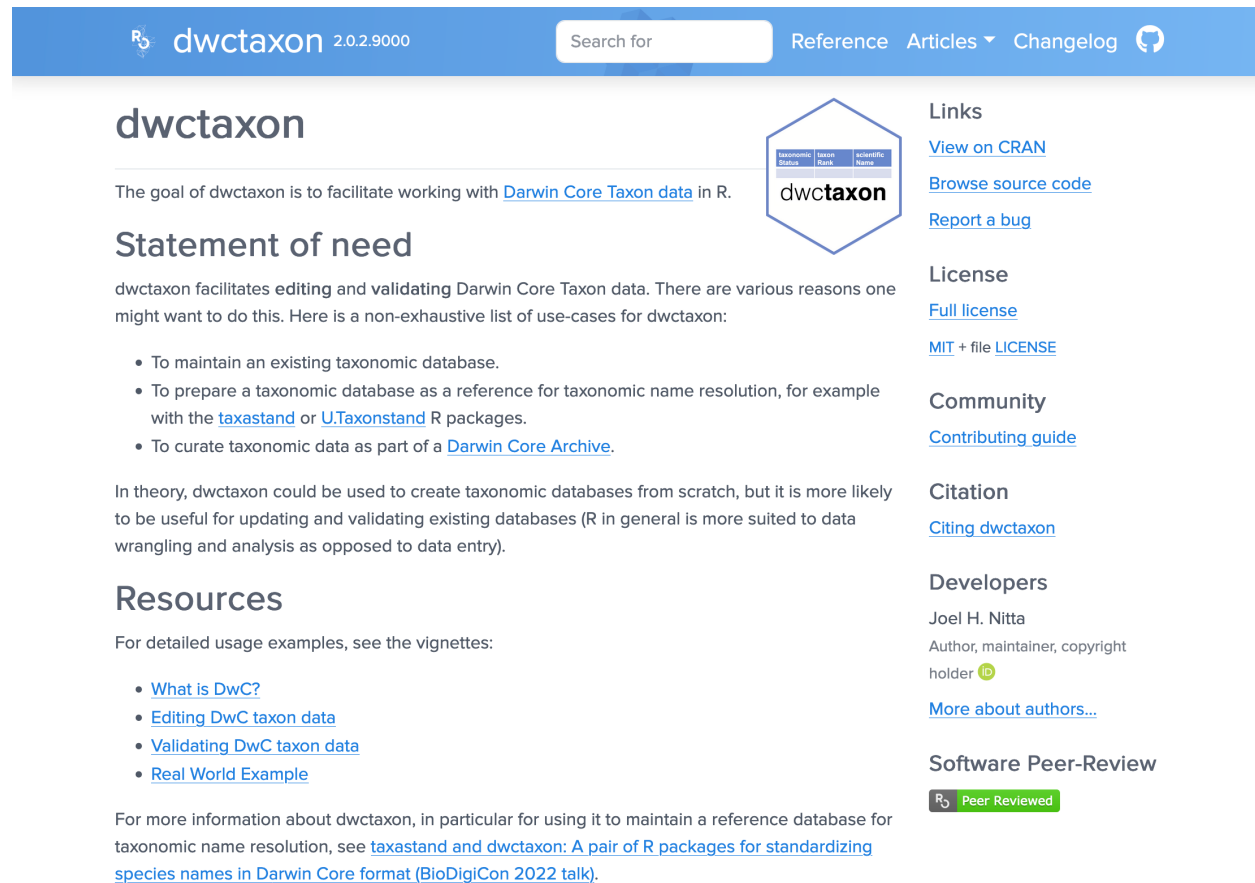# `dct_validate()` finds problems in taxonomic data

```
1  dct_change_status(
2    tax_dat = filmies,
3    sci_name = "Trichomanes crassum Copel.",
4    new_status = "synonym",
5    usage_name = "Trichomanes densinervium Copel."
6  ) |>
7  dct_validate()
```

```
Error: `check_mapping` failed.
`taxonID`(s) detected whose `acceptedNameUsageID` value does not map to
`taxonID` of an existing name.
Bad `taxonID`: 54133783
Bad `scientificName`: Trichomanes crassum Copel.
```

Please give `dwctaxon` a try!

- https://github.com/ropensci/dwctaxon

Thank you!