

# Package: authoritative (via r-universe)

February 4, 2025

**Title** Parse and Deduplicate Author Names

**Version** 0.1.0.9000

**Description** Utilities to parse authors fields from DESCRIPTION files and general purpose functions to deduplicate names in database, beyond the specific case of R package authors.

**License** MIT + file LICENSE

**URL** <https://github.com/epiverse-connect/authoritative>

**BugReports** <https://github.com/epiverse-connect/authoritative/issues>

**Depends** R (>= 4.1.0)

**Imports** stringi, utils

**Suggests** knitr, rmarkdown, spelling, testthat (>= 3.0.0)

**VignetteBuilder** knitr

**Config/Needs/website** epiverse-trace/epiversetheme, tidyverse, igraph, netUtils

**Config/testthat/edition** 3

**Config/testthat/parallel** true

**Encoding** UTF-8

**Language** en-GB

**LazyData** true

**Roxygen** list(markdown = TRUE, roclets = c("`collate", "`rd", "`namespace", "`devtag::dev\_roclet"))

**RoxygenNote** 7.3.2

**Config/Needs/build** moodymudskipper/devtag

**Config/pak/sysreqs** libicu-dev

**Repository** <https://bisaloo.r-universe.dev>

**RemoteUrl** <https://github.com/epiverse-connect/authoritative>

**RemoteRef** HEAD

**RemoteSha** 5b925e03f72ca4f8cddc47874e26f70d8c2df64e

## Contents

cran_epidemiology_packages . . . . .	2
expand_names . . . . .	2
parse_authors . . . . .	3
parse_authors_r . . . . .	4

<b>Index</b>	<b>6</b>
--------------	----------

---

cran\_epidemiology\_packages  
*A data.frame of historical metadata from CRAN packages epidemiology.*

---

### Description

A data.frame of historical metadata from CRAN packages epidemiology.

### Usage

```
cran_epidemiology_packages
```

### Format

A data.frame with 5 variables:

**Package** package name

**Version** package version

**Authors@R** authors as listed in the Authors@R field from the DESCRIPTION file

**Author** authors as listed in the Author field from the DESCRIPTION file

**Maintainer** package maintainer

---

expand\_names                    *Expand names from abbreviated forms or initials*

---

### Description

Expand names from abbreviated forms or initials

### Usage

```
expand_names(short, expanded)
```

### Arguments

short                    A character vector of potentially abbreviated names

expanded                A character vector of potentially expanded names

**Details**

When you have a list of abbreviated and non-abbreviated names and you want to deduplicate them, this function can be used as `expand_names(x, x)`, which will return the most expanded version available in `x` for each name

**Value**

A character vector with the same length as `short`

**Examples**

```
expand_names(
  c("W A Mozart", "Wolfgang Mozart", "Wolfgang A Mozart"),
  "Wolfgang Amadeus Mozart"
)

# Real-case application example
# Deduplicate names in list, as described in "details"
epi_pkg_authors <- cran_epidemiology_packages |>
  subset(!is.na(`Authors@R`), `Authors@R`, drop = TRUE) |>
  parse_authors_r() |>
  # Drop email, role, ORCID and format as string rather than person object
  lapply(function(x) format(x, include = c("given", "family"))) |>
  unlist()

# With all duplicates
length(unique(epi_pkg_authors))

# Deduplicate
epi_pkg_authors_normalized <- expand_names(epi_pkg_authors, epi_pkg_authors)

length(unique(epi_pkg_authors_normalized))
```

---

parse_authors	<i>Parse the Author field from a DESCRIPTION file</i>
---------------	---

---

**Description**

Parse the Author field from a DESCRIPTION file into a person object

**Usage**

```
parse_authors(author_string)
```

**Arguments**

`author_string` A character containing the Author or Maintainer field from a DESCRIPTION file

**Value**

A character vector, or a list of character vectors of length equals to the length of author\_string

**Examples**

```
# Read from a DESCRIPTION file directly
utils_description <- system.file("DESCRIPTION", package = "utils")
utils_authors <- read.dcf(utils_description, "Author")

parse_authors(utils_authors)

# Read from a database of CRAN metadata
cran_epidemiology_packages$Author |>
  parse_authors() |>
  unlist() |>
  unique() |>
  sort()
```

---

parse_authors_r	<i>Parse the Authors@R field from a DESCRIPTION file</i>
-----------------	--

---

**Description**

Parse the Authors@R field from a DESCRIPTION file into a person object

**Usage**

```
parse_authors_r(authors_r_string)
```

**Arguments**

```
authors_r_string
  A character containing the Authors@R field from a DESCRIPTION file
```

**Value**

A person object, or a list of person objects of length equals to the length of authors\_r\_string

**Examples**

```
# Read from a DESCRIPTION file directly
pkg_description <- system.file("DESCRIPTION", package = "authoritative")
authors_r_pkg <- read.dcf(pkg_description, "Authors@R")

parse_authors_r(authors_r_pkg)

# Read from a database of CRAN metadata
cran_epidemiology_packages |>
```

*parse\_authors\_r*

5

```
subset(!is.na(`Authors@R`), `Authors@R`, drop = TRUE) |>  
parse_authors_r() |>  
head()
```

# Index

## \* datasets

cran\_epidemiology\_packages, [2](#)

cran\_epidemiology\_packages, [2](#)

expand\_names, [2](#)

parse\_authors, [3](#)

parse\_authors\_r, [4](#)