

Package: rperseus (via r-universe)

August 30, 2026

Title Get Texts from the Perseus Digital Library

Version 0.1.2

Description The Perseus Digital Library is a collection of classical texts. This package helps you get them. The available works can also be viewed here: <http://cts.perseids.org/>.

Depends R (>= 3.3.0)

License MIT + file LICENSE

Encoding UTF-8

LazyData true

Imports stringr, httr, xml2, dplyr, purrr, rlang, rvest, ggplot2, stringi, tibble

RoxygenNote 6.0.1

Suggests knitr, rmarkdown, testthat, covr, roxygen2, tidytext

VignetteBuilder knitr

URL <https://docs.ropensci.org/rperseus>,
<https://github.com/ropensci/rperseus>

BugReports <https://github.com/ropensci/rperseus/issues>

Config/pak/sysreqs libicu-dev libxml2-dev libssl-dev

Repository <https://ropensci.r-universe.dev>

Date/Publication 2023-06-02 00:21:47 UTC

RemoteUrl <https://github.com/ropensci/rperseus>

RemoteRef master

RemoteSha 934698cd9dbf2f6923a78d5e581bfa4feba99bbf

Contents

get_perseus_text	2
greek_stop_words	3
parse_excerpt	3

perseus_catalog	4
perseus_parallel	4
rperseus	5

Index	6
--------------	----------

get_perseus_text	<i>Get a primary text by URN.</i>
------------------	-----------------------------------

Description

Get a primary text by URN.

Usage

```
get_perseus_text(urn, excerpt = NULL)
```

Arguments

- urn Valid uniform resource number (URN) obtained from [perseus_catalog](#).
- excerpt An index to excerpt the text. For example, the first four "verses" of a text might be 1.1-1.4. If NULL, the entire work is returned.

Value

A seven column tbl_df with one row for each "section" (splits vary from text—could be line, chapter, etc.). Columns:

- text** character vector of text
- urn** Uniform Resource Number
- group_name** Could refer to author (e.g. "Aristotle") or corpus (e.g. "New Testament")
- label** Text label, e.g. "Phaedrus"
- description** Text description
- language** Text language, e.g. "grc" = Greek, "lat" = Latin, "eng" = English
- section** Text section or excerpt if specified

Examples

```
get_perseus_text("urn:cts:greekLit:tlg0013.tlg028.perseus-grc2")
get_perseus_text("urn:cts:latinLit:stoa0215b.stoa003.opp-lat1")
get_perseus_text(urn = "urn:cts:greekLit:tlg0031.tlg009.perseus-grc2", excerpt = "5.1-5.5")
```

greek_stop_words	<i>A dictionary of Greek stop words</i>
------------------	---

Description

A dictionary of Greek stop words

Usage

```
greek_stop_words
```

Format

A data frame with 223 rows and 1 variable:

word Greek stop word

Source

Compiled manually by filtering prepositions, pronouns, conjunctions, particles, and articles.

parse_excerpt	<i>Parse a Greek excerpt</i>
---------------	------------------------------

Description

This function parses a Greek excerpt from the Perseus Digital Library. Parsing includes part of speech, gender, case, mood, voice, tense, person, number, and degree.

Usage

```
parse_excerpt(urn, excerpt)
```

Arguments

urn	a valid urn from the perseus_catalog object.
excerpt	a valid excerpt, e.g. 5.1-5.5

Value

a data frame

Examples

```
parse_excerpt("urn:cts:greekLit:tlg0031.tlg002.perseus-grc2", "5.1-5.4")
```

perseus_catalog	<i>Metadata for texts available via the Perseus Digital Library.</i>
-----------------	--

Description

A dataset containing the texts available from the Perseus Digital Library.

Usage

```
perseus_catalog
```

Format

A data frame with 2291 rows and 5 variables:

urn Uniform Resource Number

group_name Could refer to author (e.g. "Aristotle") or corpus (e.g. "New Testament")

label Text label, e.g. "Phaedrus"

description Text description

language Text language, e.g. "grc" = Greek, "lat" = Latin, "eng" = English, "hpt" = Hebrew pointed text, "hct" = Hebrew consonantal text, "ger" = German, "oth" = other

Source

<http://cts.perseids.org/api/cts/?request=GetCapabilities>

perseus_parallel	<i>Render a text parallel with ggplot2</i>
------------------	--

Description

Render a text parallel with ggplot2

Usage

```
perseus_parallel(perseus_df, words_per_row = 6)
```

Arguments

perseus_df a data frame obtained from `get_perseus_text`. Can contain multiple texts.

words_per_row adjusts the words displayed per "row".

Value

a ggplot object

Examples

```
## Not run:
tibble::tibble(label = c("Colossians", rep("1 Thessalonians", 2), "Romans"),
               excerpt = c("1.4", "1.3", "5.8", "8.35-8.39")) %>%
  dplyr::left_join(perseus_catalog) %>%
  dplyr::filter(language == "grc") %>%
  dplyr::select(urn, excerpt) %>%
  as.list() %>%
  purrr::pmap_df(get_perseus_text) %>%
  perseus_parallel()

## End(Not run)
```

rperseus

rperseus: A package to obtain texts from the Perseus Digital Library

Description

The package contains a catalog and a function. The catalog is [perseus_catalog](#) and can be perused to locate a particular text and its corresponding URN. Then pass the URN to [get_perseus_text](#) to obtain the text.

Author(s)

David Ranzolin <daranzolin@gmail.com>

Index

* **datasets**

greek_stop_words, [3](#)

perseus_catalog, [4](#)

get_perseus_text, [2](#), [5](#)

greek_stop_words, [3](#)

parse_excerpt, [3](#)

perseus_catalog, [2](#), [4](#), [5](#)

perseus_parallel, [4](#)

rperseus, [5](#)

rperseus-package (rperseus), [5](#)