Package: scimetr (via r-universe)

May 16, 2026

Type Package

Title Analysis of Scientific Publication Data with R

Version 1.2.0

Date 2026-03-01

Description Tools for quantitative research in scientometrics and bibliometrics. This package provides routines for importing bibliographic data from Clarivate Web of Science (<https://www.webofscience.com/wos/>) and performing bibliometric analysis.

Depends R (>= 4.1.0), graphics

Imports dplyr, tidyr, stringr, stringi, scales, ggplot2, rlang, openxlsx

Suggests DT, rmarkdown, knitr

License GPL (>= 2)

URL https://rubenfcasal.github.io/scimetr/, https://github.com/rubenfcasal/scimetr/

BugReports https://github.com/rubenfcasal/scimetr/issues/

Encoding UTF-8

VignetteBuilder knitr

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.2

Config/pak/sysreqs libicu-dev

Repository https://rubenfcasal.r-universe.dev

Date/Publication 2026-03-16 20:21:04 UTC

RemoteUrl https://github.com/rubenfcasal/scimetr

RemoteRef HEAD

RemoteSha 17c3e3c5b219662888b9eb80dae297a59cfd41d4

Contents

`scimetr-package` *scimetr: Analysis of Scientific Publication Data with R*

Description

Tools for quantitative research in scientometrics and bibliometrics. This package provides routines for importing bibliographic data from Clarivate Web of Science (https://www.webofscience.com/wos/) and performing bibliometric analysis. For more information visit https://rubenfcasal.github.io/scimetr/articles/scimetr.html.

Author(s)

Maintainer: Ruben Fernandez-Casal <rubenfcasal@gmail.com> (ORCID)

Authors:

- Borja Lafuente-Rego <borjalrego@gmail.com> (ORCID)
- Maria Jose Lombardia <maria.jose.lombardia@udc.es> (ORCID)
- Javier Tarrio-Saavedra <javier.tarrio@udc.es> (ORCID)
- Julian Costa-Bouzas <julian.costa.bouzas@udc.es> (ORCID)

Other contributors:

- Yesica Fernandez-Ramos <yesica.ramos@udc.es> [contributor]
- Guillermo Lopez-Taboada <guillermo.lopez.taboada@udc.es> [contributor]

See Also

Useful links:

- https://rubenfcasal.github.io/scimetr/
- https://github.com/rubenfcasal/scimetr/
- Report bugs at https://github.com/rubenfcasal/scimetr/issues/

add_jcr *Add JCR data to a bibliographic database*

Description

Extends the bibliographic database by adding JCR metrics to sources, per year and WoS category.

Usage

```
add_jcr(db, jcrdb)

## S3 method for class 'wos.jcr'
plot(
  x,
  filter,
  plot = TRUE,
  all = FALSE,
  ask = plot && all && interactive(),
  ...
)
```

Arguments

db	a bibliographic database (a wos.db object; typically the output of the function db_bib()).
jcrdb	a JCR database (a jcr.db object; typically the output of the function db_jcr()).
x	a wos.jcr object.
filter	vector of document identifiers (usually a result of get_id_docs()).
plot	logical; if TRUE (default), the plots are drawn, otherwise only the ggplot2 object is (invisibly) returned.
all	logical; if TRUE, function plot.wos.db() is called (additional parameters ... are passed to this function), otherwise only a JCR metrics plot is generated.
ask	logical; if TRUE, the user is asked before each plot (see par(ask=.)).
...	further arguments passed to or from other methods.

Value

An S3 object of class wos.jcr. A wos.db object with additional components JCRSour and JCRCatSour.

See Also

db_bib, db_jcr.

db_bib *Bibliographic data base creation*

Description

db_bib.wos.data() converts a data.frame with WoS variables (as returned by import_wos; see wosdf) into a relational database (a list of data.frames).

Usage

```
db_bib(data, ...)

## S3 method for class 'wos.data'
db_bib(data, label = "", progress = interactive(), verbose = FALSE, ...)

## S3 method for class 'wos.db'
print(x, ...)

## S3 method for class 'wos.db'
plot(
  x,
  filter,
  which = 1:3,
  plot = TRUE,
  warning = FALSE,
  ask = plot && (length(which) > 1) && interactive(),
  ...
)
```

Arguments

data	a data.frame with bibliographic data.
...	further arguments passed to or from other methods.
label	character string describing the data.
progress	logical; if TRUE the progress is printed.
verbose	logical; if TRUE additional information is printed.
x	an object used to select a method.
filter	vector of document identifiers (usually a result of get_id_docs).
which	if a subset of the plots is required, specify a subset of the numbers 1:3.
plot	logical; if TRUE (default), the plots are drawn, otherwise only the list of ggplot2 objects is (invisibly) returned.
warning	logical; if FALSE (default), warnings are ignored.
ask	logical; if TRUE, the user is asked before each plot (see par(ask=.)).

Value

An S3 object of class `wos.db`. A `list` with the following components:

- `Docs`: document-level records (year, type, citations, identifiers).
- `Authors` and `AutDoc`: author dictionary and document-author links.
- `OI` and `OIDoc`: ORCID identifiers and their linkage to documents.
- `RI` and `RIDoc`: ResearcherID identifiers and their linkage to documents.
- `Affiliations` and `AffDoc`: affiliation dictionary and document links.
- `Addresses` and `AddAutDoc`: address and document-author-address links.
- `Sources`: document sources (journals, books, proceedings, ...).
- `Categories`, `Areas`, `CatSour` and `AreaSour`: thematic classifications and their linkage to sources.
- `WSIndex` and `SourWSI`: Web of Science indexes and their linkage to sources.
- `label` and `date`: metadata that helps tracking the dataset identity over time.

This object has specialized print and plot methods.

See Also

`wosdf`, `import_wos`.

`summary.wos.db`.

`plot.summary.wos.db`, `plot.summary.year.wos`

Examples

```
db <- db_bib(wosdf)
print(db)
summary(db)
```

---

db_jcr *Import JCR data from Web of Science (WoS)*

Description

Reads JCR data from excel files downloaded from WoS and generates a relational database (a list of data.frames). It is assumed that the file name format is *JCR_[WE]_[PY]*, where *[WE]* is the WoS index (SCIE, SSCI, ...) and *[PY]* the JCR year.

Usage

```
db_jcr(path = ".", files = dir(path, pattern = "*.xlsx"), verbose = TRUE)
```

Arguments

path	character; path to the directory containing the files. Defaults to the working directory.
files	character vector with the file names. Defaults to filenames in directory path with extension *.xlsx"*.
verbose	logical; indicating whether the name of the file being processed is printed. Defaults to TRUE.

Details

A subscription to Web of Science is required to download JCR data. For successful execution, Excel files must retain the original header fields from Clarivate exports: Title20, ISO_ABBREV, TITLE, ISSN, EISSN, ISSUES/YEAR, COUNTRY, LANGUAGE, 1ST_YR_PUB, categories, TOT_CITES, CITES_JCR_YR, CITES_JCR_YR1, CITES_JCR_YR2, IF_NUMERATOR, CITES_JCR_YR3, CITES_JCR_YR4, CITES_JCR_YR5, 5YR_IF_NUMERATOR, ITEMS_JCR_YR, ITEMS_JCR_YR1, ITEMS_JCR_YR2, IF_DENOMINATOR, ITEMS_JCR_YR3, ITEMS_JCR_YR4, ITEMS_JCR_YR5, 5YR_IF_DENOMINATOR, IMPACT_FACTOR, IMMEDIACY_INDEX, CITED_HALF_LIFE, 5YR_IMPACT_FACTOR, EIGENFACTOR, NORM_EIGENFACTOR, ARTL_INFLUENCE, FREQUENCY, PUBCODE, PUBLISHER_NAME, CATEGORY_CODE, CATEGORY_DESCRIPTION, CATEGORY_RANKING, QUARTILE_RANK, JIF_PERCENTILE.

Value

An S3 object of class jcr.db. A list with components: Sources, Categories, JCRSour and JCRCatSour.

See Also

db_bib, add_jcr.

---

dbjcr *Mathematics UDC 2018-2023 bibliographic database*

Description

Bibliographic database with JCR metrics (a wos.jcr S3 object) corresponding to a WoS search by the Affiliation field of *Universidade da Coruña (UDC)* in the research area "Mathematics" during the years 2018–2023 (generated fom wosdf data set, using the functions db_bib, db_jcr and add_jcr).

Usage

```
dbjcr
```

Format

An object of class wos.jcr (inherits from wos.db) of length 22.

`get_id_docs` *Get table identifiers*

Description

Retrieve internal identifiers (entity keys) for the corresponding relational table (authors, sources, categories, areas, ...). Any field in the target table can be used as a condition, and multiple conditions are combined with & (only IDs for which all conditions evaluate to TRUE are returned). *Tidy evaluation* can be used to construct the logical expressions, see `dplyr::filter()`.

Usage

```
get_id_docs(
  db,
  ...,
  id_sources,
  id_authors,
  id_addresses,
  id_oi,
  id_ri,
  id_affiliations
)

get_id_authors(db, ...)

get_id_addresses(db, ...)

get_id_areas(db, ...)

get_id_categories(db, ...)

get_id_wsi(db, ...)

get_id_sources(db, ..., id_categories, id_areas, id_wsi)

get_id_oi(db, ...)

get_id_ri(db, ...)

get_id_affiliations(db, ...)
```

Arguments

db	Object of `class` `wos.db`, as returned by db_bib.
...	Logical predicates. Multiple conditions are combined with & (see `filter`).
id_sources	optional; sources identifiers.

id_authors	optional; author identifiers (values of db$Authors$ids).
id_addresses	optional; addresses identifiers.
id_oi	optional; ORCID identifiers or codes (values of dbOIOI).
id_ri	optional; RI identifiers or codes (values of dbRIRI).
id_affiliations	optional; affiliations identifiers or names (values of db$Affiliations$C3).
id_categories	optional; categories identifiers or names (values of db$Categories$WC)
id_areas	optional; research area identifiers or names (values of db$Areas$SC)
id_wsi	optional; WoS Index identifiers or names (values of dbWSIWE)

Value

An integer vector of identifiers. For instance, get_id_docs returns the document identifiers (values of the db$Docs$idd variable) corresponding to identifiers of authors, categories, areas, addresses and/or sources. Logical expressions defined in terms of the variables in db$Docs can also be used as arguments.

See Also

db_bib, filter.

---

get_jcr *Get JCR metrics*

Description

get_jcr() combines document indexes with their source JCR metrics per year.

get_jcr_cat() combines document indexes with their source JCR metrics per year and WoS category.

Usage

```
get_jcr(db, filter)

get_jcr_cat(db, filter, best = TRUE)
```

Arguments

db	a bibliographic database with JCR information (a wos.jcr object; typically the output of the function add_jcr()).
filter	vector of document identifiers (usually a result of get_id_docs()).
best	logical; if TRUE (default), only the results for the WoS category with the best ranking for each document are returned.

Value

A `data.frame` with document indexes and their source JCR metrics.

See Also

`add_jcr()`, `db_jcr()`, `db_bib()`.

Examples

```
head(get_jcr(dbjcr))
head(get_jcr_cat(dbjcr, best = TRUE))
```

`import_wos` *Import bibliographic data downloaded from Web of Science (WoS).*

Description

Reads bibliography entries from UTF-8 encoded Tab-delimited files containing "Full Record and Cited References" (see wosdf and vignette Downloading data from the Web of Science).

Usage

```
import_wos(path = ".", pattern = "*.txt", all = TRUE, progress = NULL)
```

Arguments

`path`	character; path to the directory containing the files.
`pattern`	regular expression; only matching files will be loaded. Defaults to `"*.txt"`.
`all`	logical; indicating whether sources without ISSN are included (`TRUE`) or not (`FALSE`) in the result. Defaults to `TRUE`.
`progress`	logical; indicating whether a progress bar should be displayed. Defaults to `TRUE` when R is used interactively and there is more than one file, and `FALSE` otherwise.

Details

A subscription to Web of Science is required to download bibliometric data.

Value

A `data.frame` with rows corresponding to sources and columns to WoS variables.

See Also

wosdf, `db_bib()`.

summary.wos.db *Summaries of a bibliographic database*

Description

summary() methods for bibliometric databases.

Usage

```
## S3 method for class 'wos.db'
summary(object, filter, index = c("H", "G"), nmax = 10, ...)

## S3 method for class 'summary.wos.db'
print(x, ...)

author_metrics(db, id_authors)

## S3 method for class 'summary.wos.db'
plot(
  x,
  which = 1:5,
  pie = FALSE,
  others = !pie,
  accuracy = 0.1,
  plot = TRUE,
  ask = length(which) > 1 && interactive(),
  ...
)

## S3 method for class 'wos.jcr'
summary(object, filter, all = FALSE, ...)

## S3 method for class 'summary.wos.jcr'
print(x, digits = 2, all = TRUE, ...)

## S3 method for class 'summary.wos.jcr'
plot(x, plot = TRUE, all = FALSE, ask = plot && all && interactive(), ...)
```

Arguments

object	an object for which a summary is desired.
filter	vector of document identifiers (optional, usually a result of get_id_docs()).
index	citation indexes.
nmax	number of top levels.
...	further arguments passed to or from other methods.

x	an object used to select a method.
db	Object of class wos.db, as returned by db_bib.
id_authors	optional; author identifiers
which	if a subset of the plots is required, specify a subset of the numbers 1:5.
pie	logical; if TRUE, pie charts are drawn.
others	logical; if FALSE, only nmax top levels categories are shown.
accuracy	numerical scalar; see label_number(accuracy=.).
plot	logical; if TRUE (default), the plots are drawn, otherwise only the list of ggplot2 objects is (invisibly) returned.
ask	logical; if TRUE, the user is asked before each plot, see par(ask=.).
all	logical; if TRUE, the corresponding wos.db method is called (additional parameters ... are passed to it), otherwise only results of JCR metrics are returned.
digits	minimal number of significant digits, see print.default.

Value

A list of summary statistics (an object of class summary.wos.db and/or summary.wos.jcr), which has specialized print() and plot() methods.

See Also

db_bib(), add_jcr().

Examples

```
# Bibliographic database
db <- db_bib(wosdf)
summary(db)
# Bibliographic database with JCR metrics
summary(dbjcr)
```

---

summary_year *Summaries per year*

Description

The summary_year methods return summaries by year of a bibliometric database.

Usage

```
summary_year(object, ...)

## S3 method for class 'wos.db'
summary_year(object, filter, ...)

## S3 method for class 'summary.year.wos'
print(x, ...)

## S3 method for class 'summary.year.wos'
plot(
  x,
  which = 1:3,
  boxplot = FALSE,
  plot = TRUE,
  ask = plot && length(which) > 1 && interactive(),
  ...
)

## S3 method for class 'wos.jcr'
summary_year(object, filter, all = FALSE, ...)

## S3 method for class 'summary.year.jcr'
print(x, digits = 3, all = TRUE, ...)

## S3 method for class 'summary.year.jcr'
plot(x, plot = TRUE, all = FALSE, ask = plot && interactive(), ...)
```

Arguments

object	an object for which a summary is desired.
...	further arguments passed to or from other methods.
filter	vector of document identifiers (optional, usually a result of get_id_docs).
x	an object used to select a method.
which	if a subset of the plots is required, specify a subset of the numbers 1:3.
boxplot	logical; if TRUE, boxplots are drawn (for plots from 2 to 3).
plot	logical; if TRUE (default), the plots are drawn, otherwise only the list of ggplot2 objects is (invisibly) returned.
ask	logical; if TRUE, the user is asked before each plot, see par(ask=.).
all	logical; if TRUE, the corresponding wos.db method is called (additional parameters ... are passed to it), otherwise only results of JCR metrics are returned.
digits	minimal number of *significant* digits, see print.default.

Value

A list of summary statistics (an object of class summary.year.wos and/or summary.year.jcr), which has specialized print() and plot() methods.

See Also

db_bib(), add_jcr().

Examples

```
# Bibliographic database
db <- db_bib(wosdf)
summary_year(db)
# Bibliographic database with JCR metrics
summary_year(dbjcr)
```

---

wosdf *Mathematics UDC WoS Core Collection 2018-2023 data*

Description

The data set consists of 293 publications corresponding to a WoS search by the Affiliation field of *Universidade da Coruña (UDC)* (Affiliation: OG = Universidade da Coruna) in the research area "Mathematics" (SC = Mathematics), indexed in the Web of Science Core Collection:

- Science Citation Index Expanded (SCI-EXPANDED).
- Social Sciences Citation Index (SSCI).
- Arts & Humanities Citation Index (A&HCI).

in the years 2018-2023 (generated using the import_wos function).

Usage

```
wosdf
```

Format

A data frame with 293 rows and 48 columns:

PT Publication Type

AU Author

AF Author Full Name

TI Article Title

SO Source Title

SE Book Series Title

BS Book Series Subtitle

LA Language

DT Document Type

C1 Address

C3 Affiliation

RI Researcher Ids
OI ORCID
NR Cited Reference Count
TC Times Cited, WoS Core
Z9 Times Cited, All Databases
U1 180 Day Usage Count
U2 Since 2013 Usage Count
PU Publisher
PI Publisher City
PA Publisher Address
SN ISSN
EI eISSN
BN ISBN
J9 Journal Abbreviation
JI Journal ISO Abbreviation
PD Publication Date
PY Publication Year
VL Volume
IS Issue
PN Part Number
SU Supplement
SI Special Issue
MA Meeting Abstract
BP Start Page
EP End Page
AR Article Number
DI DOI
D2 Book DOI
EA Early Access Date
PG Number of Pages
WC WoS Category
WE Web of Science Index
SC Research Area
HC Highly Cited Status
HP Hot Paper Status
DA Date of Export
UT Unique WOS ID

Source

Clarivate Web of Science: https://www.webofscience.com/wos/.

See Also

import_wos, db_bib

Index