

# Package ‘GetLattesData’

March 7, 2020

**Title** Reading Bibliometric Data from Lattes Platform

**Version** 1.4

**Date** 2020-03-07

**Description**

A simple API for downloading and reading xml data directly from Lattes <<http://lattes.cnpq.br/>>.

**Depends** R (>= 3.3.0)

**Imports** stringr, XML, dplyr, readr, stringdist, curl, tools

**License** GPL-2

**BugReports** <https://github.com/msperlin/GetLattesData/issues>

**URL** <https://github.com/msperlin/GetLattesData/>

**LazyData** true

**RoxygenNote** 7.0.2

**Suggests** knitr, rmarkdown, testthat, ggplot2

**VignetteBuilder** knitr

**NeedsCompilation** no

**Author** Marcelo Perlin [aut, cre]

**Maintainer** Marcelo Perlin <[marceloperlin@gmail.com](mailto:marceloperlin@gmail.com)>

**Repository** CRAN

**Date/Publication** 2020-03-07 15:30:02 UTC

## R topics documented:

<code>gld.get.papers.info</code>	2
<code>gld_download_lattes_files</code>	2
<code>gld_get_lattes_data</code>	3
<code>gld_get_lattes_data_from_zip</code>	4
<code>gld_get_qualis</code>	5
<code>gld_get_SJR</code>	6
<code>gld_read_zip</code>	6

<b>Index</b>	<b>8</b>
--------------	----------

`gld.get.papers.info` *Extracts information about papers from xml data*

---

**Description**

Extracts information about papers from xml data

**Usage**

```
gld.get.papers.info(l.papers, name.author, id.author)
```

**Arguments**

<code>l.papers</code>	A list with papers information
<code>name.author</code>	Name of author
<code>id.author</code>	ID of author

**Value**

A dataframe with several information about papers (published or accepted)

**Examples**

```
## Not run: # INTERNAL USE
tpublish <- gld.get.papers.info(i.papers, name.author, id.author)

## End(Not run)
```

---

`gld_download_lattes_files`  
*Downloads data from Lattes*

---

**Description**

Downloads data from Lattes

**Usage**

```
gld_download_lattes_files(id, folder.dl = tempdir())
```

**Arguments**

<code>id</code>	Id symbol from Lattes (see <code>gld_get_lattes_data</code> )
<code>folder.dl</code>	Name of folder where to store xml files (default = <code>tempdir()</code> )

**Value**

The name of downloaded zip file

**Examples**

```
## Not run:  
file.out <- gld_download_lattes_files(id = 'K4723925J2')  
  
## End(Not run)
```

---

`gld_get_lattes_data` *Downloads and reads Lattes data based on a vector of Lattes ids*

---

**Description**

DEPRECATED. Please use function `gld_get_lattes_data_from_zip` to read the files

**Usage**

```
gld_get_lattes_data(id.vec, field.qualis = NULL, folder.dl = tempdir())
```

**Arguments**

<code>id.vec</code>	A vector of Lattes ids (e.g. <code>id.vec &lt;- c('K4723925J2', 'K4713546D3')</code> )
<code>field.qualis</code>	Area of Qualis to get Qualis journal rankings (default equals NULL). Eg. <code>area.qualis &lt;- 'ECONOMIA'</code>
<code>folder.dl</code>	Name of folder where to store xml files (default = <code>tempdir()</code> )

**Details**

This function downloads xml data directly from Lattes and reads the resulting file, giving as output three dataframes, a table with description of researchers, list of all publications and a table with completed supervisions

**Value**

Returns a list with two components:

**tpesq** A dataframe with information about researchers

**tpublic** A dataframe with information about publications

**tsupervisions** A dataframe with information about all supervisions

**Examples**

```
## Not run:
l.out <- gld_get_lattes_data(id.vec = 'K4713546D3',
                           field.qualis = 'ECONOMIA')
print(l.out$tpesq)

## End(Not run)
```

---

```
gld_get_lattes_data_from_zip
  Reads zip files from Lattes
```

---

**Description**

This function reads zipped files from Lattes, giving as output a list with several dataframes

**Usage**

```
gld_get_lattes_data_from_zip(zip.files, field.qualis = NULL)
```

**Arguments**

<code>zip.files</code>	A vector with location of zip files downloaded from Lattes website
<code>field.qualis</code>	Area of Qualis to get Qualis journal rankings (default equals NULL). Eg. <code>area.qualis &lt;- 'ECONOMIA'</code>

**Value**

Returns a list with two components:

**tpesq** A dataframe with information about researchers

**public** A dataframe with information about publications

**tsupervisions** A dataframe with information about all supervisions

**Examples**

```
## Not run:
# get files from pkg (you can download from other researchers in lattex website)
f.in <- system.file('extdata/3262699324398819.zip', package = 'GetLattesData')

# set qualis
field.qualis = 'ECONOMIA'

# get data
l.out <- gld_get_lattes_data_from_zip(f.in, field.qualis = field.qualis )
```

```
# print it
print(l.out$tpesq)
print(l.out$public.published)

## End(Not run)
```

---

gld_get_qualis	<i>Reads Qualis table</i>
----------------	---------------------------

---

## Description

Read Qualis data available within the package. The original data is downloaded from <https://sucupira.capes.gov.br/sucupira/public/index.xhtml>

## Usage

```
gld_get_qualis(field.qualis = "ALL")
```

## Arguments

`field.qualis` Area of Qualis to get Qualis journal rankings (default equals NULL). Eg. `area.qualis <- 'ECONOMIA'`

## Details

Latest Qualis: 2013-2016 | Latest Qualis update: 2017-09-04

## Value

A dataframe with Qualis table, including columns for ISSN, field and ranking

## Examples

```
df.qualis <- gld_get_qualis()
df.qualis
```

---

gld_get_SJR	<i>Reads SJR table</i>
-------------	------------------------

---

### Description

Reads locally available SJR table. Original data provided as an excel file in <http://scimagojr.com/journalrank.php>.

### Usage

```
gld_get_SJR()
```

### Details

Latest SJR: 2016 | Latest update: 2017-09-04

### Value

A dataframe with SJR table

### Examples

```
df.SJR <- gld_get_SJR()
df.SJR
```

---

gld_read_zip	<i>Reads xml data from lattes zip file</i>
--------------	--

---

### Description

Reads xml data from lattes zip file

### Usage

```
gld_read_zip(zip.in)
```

### Arguments

zip.in            A (single) name of zip file containing a xml file

### Value

A list with the following items:

**tpesq** A dataframe with information about researchers

**tpublic** A dataframe with information about publications

**Examples**

```
f.in <- system.file('extdata/3262699324398819.zip', package = 'GetLattesData')
my.l <- gld_read_zip(f.in)
my.l
```

# Index

`gld.get.papers.info`, 2  
`gld_download_lattes_files`, 2  
`gld_get_lattes_data`, 3  
`gld_get_lattes_data_from_zip`, 4  
`gld_get_qualis`, 5  
`gld_get_SJR`, 6  
`gld_read_zip`, 6