

# Package ‘rawr’

June 8, 2020

**Type** Package

**Title** Retrieve Raw R Code from Popular Tutorials and Websites

**Version** 0.1.1

**Author** Steve Condylis [aut, cre] (<<https://orcid.org/0000-0003-0599-844X>>)

**Maintainer** Steve Condylis <[steve.condylis@gmail.com](mailto:steve.condylis@gmail.com)>

**BugReports** <https://github.com/stevecondylis/rawr/issues>

**License** MIT + file LICENSE

**URL** <https://github.com/stevecondylis/rawr>

**Description** Retrieves pure R code from popular R websites, including github <<https://github.com>>, kaggle <<https://www.kaggle.com>>, datacamp <<https://www.datacamp.com>>, and R blogs made using R blogdown <<https://github.com/rstudio/blogdown>>.

**Encoding** UTF-8

**LazyData** true

**Imports** rvest, dplyr, jsonlite, xml2, stringr, clipr

**RoxygenNote** 7.1.0

**Suggests** testthat

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2020-06-08 09:20:02 UTC

## R topics documented:

|                 |   |
|-----------------|---|
| blogdown        | 2 |
| datacamp        | 3 |
| gist_url_to_raw | 4 |
| github          | 4 |
| identify_domain | 5 |
| kaggle          | 6 |
| medium          | 7 |
| output_to_file  | 8 |

|                         |    |
|-------------------------|----|
| rawr . . . . .          | 8  |
| rawr_to_clip . . . . .  | 10 |
| r_bloggers . . . . .    | 11 |
| stackoverflow . . . . . | 12 |
| tidytext . . . . .      | 12 |

|              |           |
|--------------|-----------|
| <b>Index</b> | <b>14</b> |
|--------------|-----------|

---

|          |   |
|----------|---|
| blogdown | <i>Retrieve raw R code from a blogdown web page</i> |
|----------|---|

---

### Description

Retrieve raw R code from a blogdown web page

### Usage

```
blogdown(url, method, padding)
```

### Arguments

|         |  |
|---------|--|
| url     | Link to a blogdown web page  |
| method  | Not all websites are formatted consistently. To overcome this, try a different method by setting the method parameter to integers 2 and greater to try other available methods                         |
| padding | Specify what goes between the last character of one code block and the first character of the next code block. Default is a two new lines, which appears visually as one new line between code blocks. |

### Value

A character vector of length 1 containing the R code from the target url. Code from other languages (e.g. javascript, bash, or python) will also be returned if present.

### Examples

```
library(dplyr)
blogdown("https://www.jtimm.net/2019/04/14/lexical-change-procrustes/")

# Same as above but provided to cat for easy viewing
blogdown("https://www.jtimm.net/2019/04/14/lexical-change-procrustes/") %>%
  cat
```

---

`datacamp`*Retrieve raw R code from a datacamp tutorial*

---

**Description**

Retrieve raw R code from a datacamp tutorial

**Usage**

```
datacamp(url, method, padding)
```

**Arguments**

|                      |   |
|----------------------|---|
| <code>url</code>     | Link to datacamp tutorial   |
| <code>method</code>  | Not all websites are formatted consistently. To overcome this, try a different method by setting the method parameter to integers 2 and greater to try other available methods                    |
| <code>padding</code> | Specify what goes between the last character of one code block and the first character of the next code block. Default is a two new lines, which appears visually as one new line between blocks. |

**Value**

A character vector of length 1 containing the R code from the target url.

**Examples**

```
library(dplyr)
datacamp("https://www.datacamp.com/community/tutorials/sentiment-analysis-R")
datacamp("https://www.datacamp.com/community/tutorials/R-nlp-machine-learning")

# Same as above but provided to cat for easy viewing
datacamp("https://www.datacamp.com/community/tutorials/sentiment-analysis-R") %>%
  cat
datacamp("https://www.datacamp.com/community/tutorials/R-nlp-machine-learning") %>%
  cat

# If datacamp() doesn't return a sensible result, try another method
"https://www.datacamp.com/community/tutorials/keras-r-deep-learning" %>% datacamp(.) %>% cat
# no result

# Let's try another method
"https://www.datacamp.com/community/tutorials/keras-r-deep-learning" %>%
  datacamp(method = 2) %>%
  cat # from inspection of output we see this method returns code correctly

# Note that rawr::rawr() has the same capability
```

```
"https://www.datacamp.com/community/tutorials/keras-r-deep-learning" %>%  
  rawr(method = 2) %>% cat
```

---

|                 |   |
|-----------------|---|
| gist_url_to_raw | <i>Convert github gist url into equivalent url for raw code</i> |
|-----------------|---|

---

### Description

Convert github gist url into equivalent url for raw code

### Usage

```
gist_url_to_raw(url)
```

### Arguments

url                    Link to a github gist

### Value

A character vector of length 1 containing url for the raw version of the github gist

### Examples

```
gist_url_to_raw("https://gist.github.com/korkridake/ea96f6569f5a4f9a1a552145e6151235")
```

---

|        |   |
|--------|---|
| github | <i>Retrieve raw R code from an .R file hosted on github website</i> |
|--------|---|

---

### Description

Retrieve raw R code from an .R file hosted on github website

### Usage

```
github(url, method, padding)
```

**Arguments**

|         |  |
|---------|--|
| url     | Link to an R file on github website  |
| method  | Not all websites are formatted consistently. To overcome this, try a different method by setting the method parameter to integers 2 and greater to try other available methods                         |
| padding | Specify what goes between the last character of one code block and the first character of the next code block. Default is a two new lines, which appears visually as one new line between code blocks. |

**Value**

A character vector of length 1 containing the R code from the target url. All code at the target url (including javascript, ruby, python) will be returned.

**Examples**

```
library(dplyr)
github("https://github.com/hadley/vis-eda/blob/master/travel.R")

# Same as above but provided to cat for easy viewing
github("https://github.com/hadley/vis-eda/blob/master/travel.R") %>%
  cat
```

---

|                 |                                       |
|-----------------|---------------------------------------|
| identify_domain | <i>Identify the domain of the url</i> |
|-----------------|---------------------------------------|

---

**Description**

Identify the domain of the url

**Usage**

```
identify_domain(urls)
```

**Arguments**

|      |                |
|------|----------------|
| urls | Vector of urls |
|------|----------------|

**Value**

A character vector of length 1 containing the website domain.

## Examples

```
test_domains <- c("https://github.com/hadley/vis-eda/blob/master/travel.R",
  "https://www.datacamp.com/community/tutorials/sentiment-analysis-R",
  "https://www.tidytextmining.com/sentiment.html",
  "https://www.kaggle.com/vrtjso/mercari-eda-more-info-than-you-can-imagine")

identify_domain(test_domains)
```

---

kaggle

*Retrieve raw R code from a kaggle notebook*


---

## Description

Retrieve raw R code from a kaggle notebook

## Usage

```
kaggle(url, method, padding)
```

## Arguments

|         |  |
|---------|--|
| url     | Link to a kaggle notebook  |
| method  | Not all websites are formatted consistently. To overcome this, try a different method by setting the method parameter to integers 2 and greater to try other available methods                         |
| padding | Specify what goes between the last character of one code block and the first character of the next code block. Default is a two new lines, which appears visually as one new line between code blocks. |

## Value

A character vector of length 1 containing the R code from the target url. Returned value will contain rmarkdown or python code if that was the language used in the kaggle notebook.

## Examples

```
library(dplyr)
kaggle("https://www.kaggle.com/vrtjso/mercari-eda-more-info-than-you-can-imagine")
kaggle("https://www.kaggle.com/captcalculator/a-very-extensive-mercari-exploratory-analysis")
kaggle("https://www.kaggle.com/adityaecdrd/mnist-with-keras-for-beginners-99457")

# Same as above but provided to cat for easy viewing
```

```
# R
kaggle("https://www.kaggle.com/vrtjso/mercari-eda-more-info-than-you-can-imagine") %>%
  cat

# rmarkdown
kaggle("https://www.kaggle.com/captcalculator/a-very-extensive-mercari-exploratory-analysis") %>%
  cat

# python
kaggle("https://www.kaggle.com/adityaecdrid/mnist-with-keras-for-beginners-99457") %>%
  cat
```

---

|        |  |
|--------|--|
| medium | <i>Retrieve r code from Medium article or Towards Data Science article</i> |
|--------|--|

---

## Description

Retrieve r code from Medium article or Towards Data Science article

## Usage

```
medium(url, method, padding)
```

## Arguments

|         |  |
|---------|--|
| url     | Link to a medium or towards data science article   |
| method  | Not all websites are formatted consistently. To overcome this, try a different method by setting the method parameter to integers 2 and greater to try other available methods                         |
| padding | Specify what goes between the last character of one code block and the first character of the next code block. Default is a two new lines, which appears visually as one new line between code blocks. |

## Value

A character vector of length 1 containing the R code from the target url. All code at the target url (including javascript, ruby, python) will be returned.

## Examples

```
library(dplyr)
url <- "https://medium.com/@data_datum/ggplot2-or-how-to-make-fancy-graphics-with-r-fabae7fc16d4"
medium(url)
```

---

|                |  |
|----------------|--|
| output_to_file | <i>Open contents in a new tempfile or file</i> |
|----------------|--|

---

### Description

Open contents in a new tempfile or file

### Usage

```
output_to_file(content, file_path)
```

### Arguments

|           |   |
|-----------|---|
| content   | Some (text) content to go into the new file |
| file_path | Specifies a file path and file name.        |

### Examples

```
## Not run:  
code_sample <- "#Sample code\nx <- 6\n2 * 2"  
output_to_file(code_sample)  
  
## End(Not run)
```

---

|      |  |
|------|--|
| rawr | <i>Retrieve raw R code from any github, kaggle, stack overflow, blog-down, or data camp web page</i> |
|------|--|

---

### Description

Retrieve raw R code from any github, kaggle, stack overflow, blogdown, or data camp web page

### Usage

```
rawr(url, to_file, file_path, method, padding)
```



**Arguments**

|           |   |
|-----------|---|
| url       | Link to an R file on supported website (github, kaggle, datacamp, tidytext)   |
| to_file   | Boolean to state whether to return output into a new file. Default is false.  |
| file_path | Specifies a file path and file name. The default is to create a randomly named temp file (see ?tempfile for further information)  |
| method    | Not all websites are formatted consistently. To overcome this, rawr may have more than one method for any site. If what rawr returns is not sensible, try setting method parameter to integers 2 and greater to try other available methods |
| padding   | Specify what goes between the last character of one code block and the first character of the next code block. Default is a two new lines, which appears visually as one new line between code blocks.                                      |

**Value**

A character vector of length 1. rawr attempts to retrieve and return the raw R code it finds at the target url. In the case of blogdown pages, all code will be returned (not just) R code, and in the case of kaggle, all of R (R markdown) and Python code will be returned.

**Examples**

```
## Not run:
library(dplyr)
rawr("https://github.com/hadley/vis-eda/blob/master/travel.R")

# Same as above but provided to cat for easy viewing
rawr("https://github.com/hadley/vis-eda/blob/master/travel.R") %>%
  cat

# Use on multiple urls

domains <- c("https://github.com/hadley/vis-eda/blob/master/travel.R",
            "https://www.datacamp.com/community/tutorials/sentiment-analysis-R",
            "https://www.tidytextmining.com/sentiment.html",
            "https://www.kaggle.com/vrtjso/mercari-eda-more-info-than-you-can-imagine",
            "https://www.jtimm.net/2019/04/14/lexical-change-procrustes/")

domains %>% sapply(rawr)

# When default method doesn't return a sensible result, try a different method

"https://www.datacamp.com/community/tutorials/keras-r-deep-learning" %>% rawr %>% cat
# no result

# Let's try another method
"https://www.datacamp.com/community/tutorials/keras-r-deep-learning" %>%
  rawr(method = 2) %>%
  cat # from inspection of output we see this method returns code correctly

# Write output to a temp file
```

```

"https://www.datacamp.com/community/tutorials/R-nlp-machine-learning" %>%
  rawr(to_file = TRUE)

# Or to a new file (provide a file name)
"https://www.datacamp.com/community/tutorials/R-nlp-machine-learning" %>%
  rawr(to_file = TRUE, file_path = "yourfile.R")

## End(Not run)

```

---

|              |  |
|--------------|--|
| rawr_to_clip | <i>Retrieve raw R code from any github, kaggle, stack overflow, blog-down, or data camp web page</i> |
|--------------|--|

---

### Description

Retrieve raw R code from any github, kaggle, stack overflow, blogdown, or data camp web page

### Usage

```
rawr_to_clip(url, method, padding)
```

### Arguments

|         |   |
|---------|---|
| url     | Link to an R file on supported website (github, kaggle, datacamp, tidytext)   |
| method  | Not all websites are formatted consistently. To overcome this, rawr may have more than one method for any site. If what rawr returns is not sensible, try setting method parameter to integers 2 and greater to try other available methods |
| padding | Specify what goes between the last character of one code block and the first character of the next code block. Default is a two new lines, which appears visually as one new line between code blocks.                                      |

### Value

The raw R code from the URL provided written to the system clipboard (and so can be pasted in the usual way with ctrl + v / cmd + v). No object is returned to the R interpreter.

### Examples

```

library(dplyr)
rawr_to_clip("https://cran.r-project.org/web/packages/tidytext/vignettes/topic_modeling.html")

# Also works with the magrittr pipe
"https://github.com/hadley/vis-eda/blob/master/travel.R" %>% rawr_to_clip

```

---

|            |   |
|------------|---|
| r_bloggers | <i>Retrieve raw R code from a kaggle notebook</i> |
|------------|---|

---

### Description

Retrieve raw R code from a kaggle notebook

### Usage

```
r_bloggers(url, method, padding)
```

### Arguments

|         |  |
|---------|--|
| url     | Link to an R-bloggers article  |
| method  | Not all websites are formatted consistently. To overcome this, try a different method by setting the method parameter to integers 2 and greater to try other available methods                         |
| padding | Specify what goes between the last character of one code block and the first character of the next code block. Default is a two new lines, which appears visually as one new line between code blocks. |

### Value

A character vector of length 1 containing the R code from the target url.

### Examples

```
library(dplyr)
paste0("https://www.r-bloggers.com/how-to-create-a-bar-",
"chart-race-in-r-mapping-united-states-city-population-1790-2010/") %>%
r_bloggers

# Same as above but provided to cat for easy viewing
paste0("https://www.r-bloggers.com/how-to-create-a-bar-",
"chart-race-in-r-mapping-united-states-city-population-1790-2010/") %>%
r_bloggers %>% cat
```

---

|               |  |
|---------------|--|
| stackoverflow | <i>Retrieve raw R code from Stack Overflow website</i> |
|---------------|--|

---

**Description**

Retrieve raw R code from Stack Overflow website

**Usage**

```
stackoverflow(url, method, padding)
```

**Arguments**

|         |  |
|---------|--|
| url     | Link to a page on Stack Overflow website (or any Stack Exchange)   |
| method  | Not all websites are formatted consistently. To overcome this, try a different method by setting the method parameter to integers 2 and greater to try other available methods                         |
| padding | Specify what goes between the last character of one code block and the first character of the next code block. Default is a two new lines, which appears visually as one new line between code blocks. |

**Value**

A character vector of length 1 containing the R code from the target url. All code at the target url (including javascript, ruby, python) will be returned.

**Examples**

```
library(dplyr)
stackoverflow("https://stackoverflow.com/questions/58248102/date-input-dt-r-shiny")

# Same as above but provided to cat for easy viewing
stackoverflow("https://stackoverflow.com/questions/58248102/date-input-dt-r-shiny") %>%
  cat
```

---

|          |  |
|----------|--|
| tidytext | <i>Retrieve raw R code tidytext tutorial</i> |
|----------|--|

---

**Description**

Retrieve raw R code tidytext tutorial

**Usage**

```
tidytext(url, method, padding)
```

```
tidytextmining(url, method, padding = "\n\n")
```

**Arguments**

|                      |   |
|----------------------|---|
| <code>url</code>     | Link to tidytext tutorial   |
| <code>method</code>  | Not all websites are formatted consistently. To overcome this, try a different method by setting the method parameter to integers 2 and greater to try other available methods                        |
| <code>padding</code> | Specify what goes between the last character of one code block and the first character of the next code block. Default is a two new lines, which appears visually as one new line between codeblocks. |

**Value**

A character vector of length 1 containing the R code from the target url.

**Examples**

```
library(dplyr)
tidytext("https://www.tidytextmining.com/sentiment.html")

# Same as above but provided to cat for easy viewing
tidytext("https://www.tidytextmining.com/sentiment.html") %>%
  cat
```

# Index

`blogdown`, [2](#)

`datacamp`, [3](#)

`gist_url_to_raw`, [4](#)  
`github`, [4](#)

`identify_domain`, [5](#)

`kaggle`, [6](#)

`medium`, [7](#)

`output_to_file`, [8](#)

`r_bloggers`, [11](#)

`rawr`, [8](#)

`rawr_to_clip`, [10](#)

`stackoverflow`, [12](#)

`tidytext`, [12](#)

`tidytextmining (tidytext)`, [12](#)