

Package: paperwizard (via r-universe)

October 29, 2024

Title Scrape News Sites using 'readability.js'

Version 0.2.0.9000

Description uses Mozillas readability.js to scrape text from websites.

License MIT + file LICENSE

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.2

Imports adaR, anytime, cli, jsonlite, processx, tibble

URL <https://github.com/schochastics/paperwizard>

BugReports <https://github.com/schochastics/paperwizard/issues>

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

RemoteUrl <https://github.com/schochastics/paperwizard>

RemoteRef HEAD

RemoteSha 98667867b84e94511540666b6510512e2d733fe7

Contents

| | |
|--------------------------|----------|
| pw_deliver | 2 |
| pw_npm_install | 2 |
| pw_report | 3 |
| Index | 4 |

| | |
|------------|------------------------------------|
| pw_deliver | <i>Scrape using Readability.js</i> |
|------------|------------------------------------|

Description

Scrape using Readability.js

Usage

```
pw_deliver(x, type = c("static", "dynamic"))
```

Arguments

| | |
|------|---|
| x | Either a vector of urls or a data.frame returned by <code>paperboy::pb_collect()</code> . |
| type | either "static" or "dynamic" if articles are scraped |

Value

A tibble similar to the output of `paperboy::pb_deliver()`.

| | |
|----------------|--|
| pw_npm_install | <i>Run NPM install Run NPM install to install dependencies</i> |
|----------------|--|

Description

Run NPM install Run NPM install to install dependencies

Usage

```
pw_npm_install()
```

Value

An installed lib

| | |
|------------------------|--------------------------------------|
| <code>pw_report</code> | <i>Summary of delivered articles</i> |
|------------------------|--------------------------------------|

Description

Summary of delivered articles

Usage

```
pw_report(x, n = 100)
```

Arguments

| | |
|----------------|---|
| <code>x</code> | result from pw_deliver() |
| <code>n</code> | integer cutoff when articles are considered too short (default 100) |

Value

nothing. called for side effects

Index

paperboy::pb_collect(), 2
paperboy::pb_deliver(), 2
pw_deliver, 2
pw_deliver(), 3
pw_npm_install, 2
pw_report, 3