

Test

Table of contents

```
library(httr)
library(jsonlite)
library(dplyr)
```

Attache Paket: 'dplyr'

Die folgenden Objekte sind maskiert von 'package:stats':

filter, lag

Die folgenden Objekte sind maskiert von 'package:base':

intersect, setdiff, setequal, union

```
library(purrr)
```

Attache Paket: 'purrr'

Das folgende Objekt ist maskiert 'package:jsonlite':

flatten

```
library(openxlsx)
```

```
extract_textquote_prefix <- function(annotation_data) {  
  # Initialize a vector to store the extracted prefixes  
  prefixes <- vector("list", length(annotation_data$target))  
  
  # Loop through each row of the annotation data  
  for (i in seq_along(annotation_data$target)) {  
    # Extract the selectors for the current target  
    selectors <- annotation_data$target[[i]]$selector  
  
    # Filter for "TextQuoteSelector" and extract the prefix  
    if (!is.null(selectors)) {  
      prefixes[[i]] <- selectors[[1]] |>  
        dplyr::filter(type == "TextQuoteSelector") |>  
        dplyr::select(prefix) |>  
        dplyr::pull()  
    } else {  
      prefixes[[i]] <- NA # Handle cases where target/selector is missing  
    }  
  }  
  
  # Return a flattened vector or a list depending on requirements  
  annotation_data |>  
    dplyr::bind_cols(as.data.frame(purrr::flatten_chr(prefixes)) |>  
      rename(quote = "purrr::flatten_chr(prefixes)"))  
}
```

```
# List of URLs to loop over
```

```
urls <- c(  
  "https://research-it-swiss-tph.github.io/quarto_training/index.html",  
  "https://research-it-swiss-tph.github.io/quarto_training/notes/quarto_intro.html",  
  "https://research-it-swiss-tph.github.io/quarto_training/notes/notebook_structure.html",  
  "https://research-it-swiss-tph.github.io/quarto_training/notes/python_r_short_demo.html",  
  "https://research-it-swiss-tph.github.io/quarto_training/practicals/exercise1_gettingsta",  
  "https://research-it-swiss-tph.github.io/quarto_training/practicals/exercise2_gettingsta",  
  "https://research-it-swiss-tph.github.io/quarto_training/practicals/exercise3.html",  
  "https://research-it-swiss-tph.github.io/quarto_training/practicals/exercise3_instructio",  
  "https://research-it-swiss-tph.github.io/quarto_training/practicals/exercise4_instructio",  
  "https://research-it-swiss-tph.github.io/quarto_training/practicals/exercise5_instructio"
```

```

"https://research-it-swiss-tph.github.io/quarto_training/correction/exercice3_test.html"
"https://research-it-swiss-tph.github.io/quarto_training/correction/exercise3_correction
"https://research-it-swiss-tph.github.io/quarto_training/correction/exercise3_step_by_st
"https://research-it-swiss-tph.github.io/quarto_training/correction/exercise4_step_by_st
)

```

```

all_annotations <- tibble()
for (url in urls) {
  api_url <- paste0("https://api.hypothes.is/api/search?uri=", url)
  response <- GET(api_url)

  if (status_code(response) == 200) {
    annotations <- fromJSON(content(response, "text"))
    if (!is.null(annotations$rows)) {
      annotation_data <- extract_textquote_prefix(annotations$rows)
      all_annotations <- bind_rows(all_annotations, annotation_data)
    }
  } else {
    message(paste("Failed to fetch data for URL:", url))
  }
}

```

No encoding supplied: defaulting to UTF-8.
 No encoding supplied: defaulting to UTF-8.
 No encoding supplied: defaulting to UTF-8.
 No encoding supplied: defaulting to UTF-8.
 No encoding supplied: defaulting to UTF-8.
 No encoding supplied: defaulting to UTF-8.
 No encoding supplied: defaulting to UTF-8.
 No encoding supplied: defaulting to UTF-8.
 No encoding supplied: defaulting to UTF-8.
 No encoding supplied: defaulting to UTF-8.
 No encoding supplied: defaulting to UTF-8.
 No encoding supplied: defaulting to UTF-8.
 No encoding supplied: defaulting to UTF-8.

```

all_annotations <- all_annotations |>
  as.data.frame() |>
  dplyr::select(created,

```

```

        text,
        quote,
        user,
        uri) |>
dplyr::rename(date = created,
              comment = text) |>
dplyr::mutate(date = as.Date(date),
              user = sub(".*:(.*)@.*", "\\1", user))
all_annotations

```

```

      date
1 2025-01-21
2 2025-01-23
3 2025-01-20
4 2025-01-17
5 2025-01-17
6 2025-01-17
7 2025-01-17
8 2025-01-17
9 2025-01-16
10 2025-01-28
11 2025-01-21
12 2025-01-20
13 2025-01-20
14 2025-01-20
15 2025-01-17
16 2025-01-16
17 2025-01-17
18 2025-01-17
19 2025-01-17
20 2025-01-16
21 2025-01-23
22 2025-01-16

```

```

1
2
3
4
5
6
7
8

```

Some personal comment:\nFor me this can be a very important reason that I hesitate to use

9
10
11
12
13
14
15
16
17
18
19
20
21
22

```
                                quote    user
1      markdown text and code.\n\nHeader hlanget
2                                <NA> hlanget
3      sing R\n\n\n\n\n\n\n\n\nIn this example,    zhuzh
4      diate files.\n\n6 Best practices\n\n        zhuzh
5      hon is now accessible in R. The              zhuzh
6      jects in R\n 6 Best practices\n            zhuzh
7      riables by defining them in the             zhuzh
8      Data/anaconda3/python.exe" with            zhuzh
9      section of an Quarto document.             zhuzh
10     h them;\nBeautify the plot using hlanget
11     tives\n\nOpen the Quarto document hlanget
12     us exercise where you learnt to             zhuzh
13     r FormatsPDFOther Linksdf1.RData           zhuzh
14     f1.RDataaswisstph_template.docx            zhuzh
15     s\n\n\n\n\nAdd the bibliography file         zhuzh
16     te.docx Step-by-step correction             zhuzh
17     g command which needs to be run            zhuzh
18     using different parameters\n\n\n\n\n        zhuzh
19     PDFOther Linksdf1.RDataadf2.RData          zhuzh
20 e your report\n\n\n\n\n\n\n\nTip\n\n\n\n\n\nSee    zhuzh
21                                <NA> zhuzh
22     and results\n  \nOther FormatsPDF          zhuzh
```

```
1      https://research-it-swiss-tph.github.io/quarto_training/notes/notebook
2      https://research-it-swiss-tph.github.io/quarto_training/notes/python_r
3      https://research-it-swiss-tph.github.io/quarto_training/notes/python_r
4      https://research-it-swiss-tph.github.io/quarto_training/notes/python_r
5      https://research-it-swiss-tph.github.io/quarto_training/notes/python_r
```

```
6             https://research-it-swiss-tph.github.io/quarto_training/notes/python_r
7             https://research-it-swiss-tph.github.io/quarto_training/notes/python_r
8             https://research-it-swiss-tph.github.io/quarto_training/notes/python_r
9     https://research-it-swiss-tph.github.io/quarto_training/practicals/exercise1_gett
10    https://research-it-swiss-tph.github.io/quarto_training/practicals/exercise3_in
11    https://research-it-swiss-tph.github.io/quarto_training/practicals/exercise4_in
12    https://research-it-swiss-tph.github.io/quarto_training/practicals/exercise4_in
13    https://research-it-swiss-tph.github.io/quarto_training/practicals/exercise4_in
14    https://research-it-swiss-tph.github.io/quarto_training/practicals/exercise4_in
15    https://research-it-swiss-tph.github.io/quarto_training/practicals/exercise4_in
16    https://research-it-swiss-tph.github.io/quarto_training/practicals/exercise4_in
17    https://research-it-swiss-tph.github.io/quarto_training/practicals/exercise5_in
18    https://research-it-swiss-tph.github.io/quarto_training/practicals/exercise5_in
19    https://research-it-swiss-tph.github.io/quarto_training/practicals/exercise5_in
20    https://research-it-swiss-tph.github.io/quarto_training/practicals/exercise5_in
21 https://research-it-swiss-tph.github.io/quarto_training/correction/exercise3_step_by_step
22 https://research-it-swiss-tph.github.io/quarto_training/correction/exercise3_step_by_step
```

```
# Write the annotation data to a CSV file
openxlsx::write.xlsx(all_annotations, 'annotations.xlsx')

# Print a message to indicate completion
cat("Annotations have been saved to 'annotations.csv'.")
```

Annotations have been saved to 'annotations.csv'.