

StringSearchInFiles.R

S. Schiwek

2020-05-01

Contents

Libraries	1
Functions	1
Results	3

Libraries

```
library(readr)
library(stringr)
library(knitr)
```

Functions

Function that searches string in file and returns how often it appeared

```
findFiles <- function(file, string = string, myignorecase = T){
  Code <- read_file(paste0(file))
  #result.grepl<-grepl(pattern = string,x = Code,ignore.case = myignorecase)
  result.count <- str_count(Code, regex(string, ignore_case = myignorecase))
  return(data.frame(resultcount = result.count))
}
```

Display the area surroundings of string in file

```
getSurroundings <- function(string = 'iasviewability', file, area = 200, myignorecase = T){
  Code <- read_file(file)
  result.count <- str_count(Code, regex(string, ignore_case = myignorecase))
  result <- vector(mode = "character", length = result.count)
  stringchars <- nchar(string)
  position <- 0
  #return(result)
  for(i in 1:result.count){
    #find position of ith appearance
    loopcode <-
      substring(Code,
        position + ((i - 1) < 0)*stringchars,
        nchar(Code))
  }
}
```

```

    )
    positionupdate <- regexpr(pattern = string,
                             text = loopcode,
                             ignore.case = myignorecase
    )
    position <- position + positionupdate
    result[i] <- substring(text = loopcode,
                          first = max(positionupdate - area, 1),
                          last = min(positionupdate + area + stringchars,
                                    nchar(loopcode)
                          )
    )
}
return(result)
}

```

Function stringSearch

Arguments	
folder	folder you want to search
string	what you are searching for inside the code in Regex
myignorecase	should your search be case insensivite?
datefilter	add a 'date modified' filter, if you want to. Only result younger than this will be shown
pattern	(Default: R, python and SQL -files)
surroundingarea	characters of surrounding area

```

stringSearch <-
function(
  folder = 'C:/nanoathenareports',
  string = 'iasviewability',
  myignorecase = T,
  datefilter = NULL,
  pattern = '\\.R$|\\.py$|\\.r$|\\.sql$|\\.SQL$',
  surroundingarea = 20
){
  # list of ALL files in folder
  # recursive: Search also sub folders
  filelist <- list.files(folder,
                        pattern = pattern,
                        recursive = T,
                        full.names = T)
  # filter out files, that match criteria
  # datefilter:
  if(!is.null(datefilter)){
    filelist <- filelist[difftime(datefilter, file.info(filelist)$mtime) < 0]}
}

```

```

# sort by 'last modified' date
filelist <- filelist[order(file.info(filelist)$mtime, decreasing = T)]

# Does file-content match string?
fileSearch <- lapply(X = filelist, FUN = findFiles, string = string)
fileSearch <- do.call(what = rbind, args = fileSearch)

#
fulllist <- data.frame(fileslist = filelist,
                      appearancecount = fileSearch$resultcount,
                      stringsAsFactors = F
)

# Filter only files, that contain string
results.tmp <- fulllist[fulllist$appearancecount > 0, ]

# show surroundings of that string
a <- lapply(X = results.tmp$fileslist,
           FUN = getSurroundings,
           string = string,
           area = surroundingarea)
names(a) <- results.tmp$fileslist
returnlist <- list()
returnlist[[1]] <- data.frame(results.tmp,
                             lastchange = format(file.info(results.tmp$fileslist)$mtime),
                             stringsAsFactors = F
)
returnlist[[2]] <- a
names(returnlist) <- c('files', 'surroundings')
return(returnlist)
}

```

enter folder you want to search

```
# source(file = paste0(getwd(), '/StringSearchInFiles/functions.R'))
```

Results

Files with that pattern

Contains:

- Actual File name
- Count of appearances of this pattern inside that file.
- Date when that file was last modified

```

result <-
stringSearch(
  folder = '/Users/ullijeratsch/rprog/ausstellung_andreas',

```

```

    string = 'iasviewability'
  )

knitr::kable(result$files,
              col.names = c("File name",
                           "Count of appearances",
                           "Date last modified")
              )

```

	File name	Count of appearances	Date last modified
1	/Users/ullijeratsch/rprog/ausstellung_andreas/StringSearchInFiles/stringSearchInFiles.R	1	2023-03-10 14:10:10
3	/Users/ullijeratsch/rprog/ausstellung_andreas/StringSearchInFiles/functions.R	2	2023-03-10 14:10:10

Show surrounding areas of appearances

```

for (myfile in 1:length(result$surroundings)) {
  cat(paste0('\r\n\r\n', 'File: ', names(result$surroundings[myfile]), '\r\n\r\n'))
  for(i in 1:length(result$surroundings[[myfile]])){
    cat(paste0('\r\n\r\n # of appearance: ', i, '\r\n\r\n', result$surroundings[[myfile]][i]))
  }
}

##
##
## File: /Users/ullijeratsch/rprog/ausstellung_andreas/StringSearchInFiles/stringSearchInFiles.R
##
##
##
## # of appearance: 1
##
## n Regex
## string = 'iasviewability'
## )
##
## # Files with tha
##
## File: /Users/ullijeratsch/rprog/ausstellung_andreas/StringSearchInFiles/functions.R
##
##
##
## # of appearance: 1
##
## s<-function(string='iasviewability',file,area=200,myign
##
## # of appearance: 2
##
## iasviewability',file,area=200,myign

```