

新华网 – XinHua.net

Samuel Joslin

12/1/2018

新华网 or XinHua.net is an popular news outlet in China. My goal is to create a number of functions to allow users to interact with headlines on XinHua.net

In order to user all the functions, users will need to aquire a unique XinHua.net API Key and a GoogleCloud API key. Below I include mine for demonstration purposes.

```
library(httr, quietly = T)
library(dplyr, quietly = T)
library(lubridate, quietly = T)
library(base64, quietly = T)

samAPIkey <- "8665b12062004f058c454ea7c4b2bd9a"
googleAPI_key <- "AIzaSyC0rcUDCTzVLfCgTekyFYhT_S4hJthGyak"
```

The user can define their API as a global vairable to make using the functions easier. This way, the user's will simply need to input the variable name into the functions, rather than enter their entire key everytime.

getHeadlines__TimeSpan

The getHeadlines__TimeSpan function takes a beginning date and end date. It returns all the headlines posted by XinHua in between those dates. Its important to note date1 must be earlier than date2.

```
getHeadlines_TimeSpan <- function(date1,date2, APIkey){
  base <- "https://newsapi.org/v2/everything?"
  sources <- "xinhua-net"
  url1 <- paste(base,"sources=",sources,"&","from=",date1,"&","to=",date2,"&sortBy=popularity&","a
    APIkey, sep = "")

  x <- httr::GET(url=url1)
  x <- jsonlite::fromJSON(rawToChar(x$content))
  return(x$articles$title)
}
```

getHeadlines__TimeSpan– Example

```
head(getHeadlines_TimeSpan(2018-11-19,2018-11-20, samAPIkey))
```

```
## [1] "Cuba rejects expansion of US economic sanctions - Xinhua"
## [2] "Spotlight: Escalating fighting in Yemen's Hodeidah sparks fears over hospitals' fate - Xinhua"
## [3] "Chinese police detains car-ramming attacker - Xinhua"
## [4] "S.Korea to prepare well for possible visit by DPRK leader to Seoul: ministry - Xinhua"
## [5] "S. African court dismisses application to set aside report on land reform - Xinhua"
## [6] "Myanmar state counselor meets Nepali prime minister to boost bilateral ties - Xinhua"
```

getHeadlines__keyword

The getHeadlines__keyword function accepts a keyword as a character vector to search by and returns the headlines regarding the input word.

```
getHeadlines_keyword <- function(keyword, APIkey){
  base <- "https://newsapi.org/v2/everything?"
  sources <- "xinhua-net"
  url1 <- paste(base,"sources=",sources,"&","q=",keyword,
                "&sortBy=popularity&","apiKey=",
                APIkey, sep = "")
  x <- httr::GET(url=url1)
  x <- jsonlite::fromJSON(rawToChar(x$content))
  return(x$articles$title)
}
```

getHeadlines__keyword- Example

```
head(getHeadlines_keyword(" 美国",samAPIkey))
```

```
## [1] "新闻分析：中美超算“你追我赶” 中国优势可圈可点 - 新华网"
## [2] "李克强会见美国国会议员访华团 - 新华网"
## [3] "美国9月份贸易逆差升至7个月新高 - 新华网"
## [4] "王毅会见美国前国务卿基辛格 - 新华网"
## [5] "专访：与中国改革开放大潮共激荡——访著名美籍华人靳羽西 - 新华网"
## [6] "杨洁篪将赴美国主持第二轮中美外交安全对话 - 新华网"
```

Chinese to English Headline Translation

This function takes a vector of chinese characters and returns a data.frame where the first column is the input and the second column is the translation.

```
chinese_to_en_translator <- function(headlines,gAPI_key){
  nh <- length(headlines)
  en_trans <- rep(NA,nh)
  for (i in 1:nh){
```

```

ch <- headlines[i]
source <- "zh"
target <- "en"
body <- paste("{",
              "'q':", ch, "'",
              "'source':", source, "'",
              "'target':", target, "'",
              "'format':'text'",
              "}", sep="")

base <- "https://translation.googleapis.com/language/translate/v2?key="
url <- paste(base, gAPI_key, sep="")

x <- httr::POST(url, body = body)
x <- jsonlite::fromJSON(rawToChar(x$content))
en_trans[i] <- x$data$translations$translatedText

}
return(data.frame(Original = headlines, Translation = en_trans))
}

```

Chinese to English Headline Translation– Example

```

chinese_headlines <- head(getHeadlines_keyword(" 美国",samAPIkey))
print(chinese_to_en_translator(chinese_headlines,googleAPI_key), right =F)

```

Original

1 新闻分析：中美超算“你追我赶” 中国优势可圈可点 - 新华网

2 李克强会见美国国会议员访华团 - 新华网

3 美国9月份贸易逆差升至7个月新高 - 新华网

4 王毅会见美国前国务卿基辛格 - 新华网

5 专访：与中国改革开放大潮共激荡——访著名美籍华人靳羽西 - 新华网

6 杨洁篪将赴美国主持第二轮中美外交安全对话 - 新华网

Translation

1 News analysis: China and the United States super-calculation "you chase me" China's advantage

2 Li Keqiang Meets with US Congress Members' Visit to China - Xinhuanet

3 US trade deficit rose to a seven-month high in September - Xinhuanet

4 Wang Yi Meets with Former US Secretary of State Kissinger - Xinhuanet

5 Interview: Coexist with China's tide of reform and opening up - Interview with famous American

6 Yang Jiechi will go to the United States to host the second round of Sino-US diplomatic security

Text to Speech

The `text_to_speech` function takes a character string and dictates it.

```
text_to_speech <-function(text, gAPIkey){

  text <-paste(text,collapse=" ")
  text <-gsub("'", "", text)

  input <-paste("'input':{'text':'", text,"'}",sep="")
  voice <-"'voice':{'languageCode':'en-GB','name':'en-GB-Standard-C','ssmlGender':'FEMALE'}"
  output <-"'audioConfig':{'audioEncoding':'MP3'}"

  s_body <-paste("{", input,",", voice,",", output,"}",sep="")
  base_url <- "https://texttospeech.googleapis.com/v1/text:synthesize?key="
  base_url <-url <-paste(base_url,gAPIkey,sep="")

  x <-httr::POST(url,body =s_body)
  x <-jsonlite::fromJSON(rawToChar(x$content))

  Sys.setenv(output_string=x$audioContent)

  system("echo $output_string > sound_out.txt")
  system("base64 --decode sound_out.txt > sound_out.mp3")
  system("afplay sound_out.mp3")
}
```

`text_to_speech`— example

```
chinese_headlines <- head(getHeadlines_keyword(" 美国",samAPIkey))
text_to_speech(chinese_headlines[1],googleAPI_key)
```