

Facebook Page Data Extraction using R: Prepared for Internet Policy Africa Meet March 1 2018

Richard Ngamita

March 01 2018

```
# Note: At the moment, the only information that can be scraped from Facebook
# is the content of public pages.
# To Scrape data from Facebooks API,
# we'll go ahead and use the R packages for this.
# To get access to the Facebook API, you need an OAuth code.
# You can get yours going to the following URL: https://developers.facebook.com/tools/explorer

# Once you're there:
# 1. Click on "Get Access Token"
# 2. Copy the long code ("Access Token") and paste it here below, substituting the fake one I wrote:

# List of packages required for this Twitter Search API data analysis
pkg <- c("Rfacebook")

# Check if packages are not installed and assign the
# names of the packages not installed to the variable new.pkg
new.pkg <- pkg[!(pkg %in% installed.packages())]

# If there are any packages in the list that aren't installed,
# install them
if (length(new.pkg)) {
  install.packages(new.pkg, repos = "http://cran.rstudio.com")
}

# Load libraries for accessing the Twitter search API.
library(Rfacebook)

## Loading required package: httr
## Loading required package: rjson
## Loading required package: httpuv
##
## Attaching package: 'Rfacebook'
## The following object is masked from 'package:methods':
##
##   getGroup

# extract Facebook data. Add the long Access Token you just got from developers.facebook.com page.
fb_oauth = 'EAACEdEose0cBAHWQXktKdNUZAmiFbmlnbGKQwmmlrE9tiMw1iZBxoIwHcwbZCTr38FQoYsy0pbcsxGGqhoiIJl1fhp'

# Now get the user details you're logged in as.
getUsers("me", token=fb_oauth, private_info=TRUE)

##           id           name username first_name middle_name
## 1 104178460415833 Kampalr Usergroup      NA           NA      NA
##   last_name gender locale likes picture birthday location hometown
## 1      NA    NA     NA    NA      NA        NA      NA      NA
```

```
## relationship_status
## 1 NA

# Does it return your Facebook public information? Yes? Then we're ready to go. See also ?fbOAuth for
# information on how to get a long-lived OAuth token.

# The following line downloads the ~200 most recent posts on the facebook page of BBC Africa
page <- getPage("bbcafrica", token=fb_oauth, n=20, reactions=TRUE, api="v2.9")

## 20 posts

# What information is available for each of these posts?
page[1,]

##               id likes_count      from_id from_name
## 1 285361880228_10156359784675229      3834 285361880228 BBC Africa
##
##               message
## 1 Ghana man teaching computing without computers becomes viral hit http://bbc.in/2F0qKtr
##               created_time type      link story comments_count
## 1 2018-02-28T21:35:02+0000 link http://bbc.in/2F0qKtr      NA      169
##   shares_count love_count haha_count wow_count sad_count angry_count
## 1      1542      560      39      205      24      1

# page[which.max(page$likes_count),]
page[which.max(page$likes_count),]

##               id likes_count      from_id from_name
## 1 285361880228_10156359784675229      3834 285361880228 BBC Africa
##
##               message
## 1 Ghana man teaching computing without computers becomes viral hit http://bbc.in/2F0qKtr
##               created_time type      link story comments_count
## 1 2018-02-28T21:35:02+0000 link http://bbc.in/2F0qKtr      NA      169
##   shares_count love_count haha_count wow_count sad_count angry_count
## 1      1542      560      39      205      24      1

page[which.max(page$comments_count),]

##               id likes_count      from_id from_name
## 1 285361880228_10156359784675229      3834 285361880228 BBC Africa
##
##               message
## 1 Ghana man teaching computing without computers becomes viral hit http://bbc.in/2F0qKtr
##               created_time type      link story comments_count
## 1 2018-02-28T21:35:02+0000 link http://bbc.in/2F0qKtr      NA      169
##   shares_count love_count haha_count wow_count sad_count angry_count
## 1      1542      560      39      205      24      1

page[which.max(page$shares_count),]

##               id likes_count      from_id from_name
## 1 285361880228_10156359784675229      3834 285361880228 BBC Africa
##
##               message
## 1 Ghana man teaching computing without computers becomes viral hit http://bbc.in/2F0qKtr
##               created_time type      link story comments_count
## 1 2018-02-28T21:35:02+0000 link http://bbc.in/2F0qKtr      NA      169
##   shares_count love_count haha_count wow_count sad_count angry_count
## 1      1542      560      39      205      24      1

# save extracted data to csv file
# Windows compute d:/internet_policy_fb.csv or right path to writable drive.
```

```
write.csv(page[1:20,], "~/internet_policy_fb.csv")
page[1:20,]
```

	id	likes_count	from_id	from_name
## 1	285361880228_10156359784675229	3834	285361880228	BBC Africa
## 2	285361880228_10156360473220229	370	285361880228	BBC Africa
## 3	285361880228_10156360542185229	78	285361880228	BBC Africa
## 4	285361880228_10156360780085229	578	285361880228	BBC Africa
## 5	285361880228_10156360831200229	26	285361880228	BBC Africa
## 6	285361880228_10156360863845229	134	285361880228	BBC Africa
## 7	285361880228_10156360938225229	671	285361880228	BBC Africa
## 8	285361880228_10156361081260229	71	285361880228	BBC Africa
## 9	285361880228_10156361163305229	59	285361880228	BBC Africa
## 10	285361880228_10156361258290229	109	285361880228	BBC Africa
## 11	285361880228_10156361328210229	81	285361880228	BBC Africa
## 12	285361880228_10156361408115229	56	285361880228	BBC Africa
## 13	285361880228_10156361451665229	89	285361880228	BBC Africa
## 14	285361880228_10156361462485229	373	285361880228	BBC Africa
## 15	285361880228_10156361504505229	83	285361880228	BBC Africa
## 16	285361880228_10156361616070229	92	285361880228	BBC Africa
## 17	285361880228_10156361752325229	151	285361880228	BBC Africa
## 18	285361880228_10156361809615229	51	285361880228	BBC Africa
## 19	285361880228_10156361819540229	37	285361880228	BBC Africa
## 20	285361880228_10156362121265229	47	285361880228	BBC Africa

##

1

2

3

4

5

Today's African #proverb: Water does not

6

7

8

9

10

11

12

Fans react as Sk

13

14

15

16

17

18

19 More than 200 Cameroonian teachers demanding back pay and allowances have been arrested and detain

20

	created_time	type
## 1	2018-02-28T21:35:02+0000	link
## 2	2018-03-01T02:33:58+0000	link
## 3	2018-03-01T03:05:00+0000	link
## 4	2018-03-01T05:29:39+0000	link
## 5	2018-03-01T06:00:20+0000	link
## 6	2018-03-01T06:15:26+0000	link
## 7	2018-03-01T06:45:48+0000	link
## 8	2018-03-01T08:18:18+0000	link

```

## 9 2018-03-01T09:06:39+0000 link
## 10 2018-03-01T10:18:20+0000 link
## 11 2018-03-01T10:54:39+0000 link
## 12 2018-03-01T11:45:02+0000 photo
## 13 2018-03-01T12:09:01+0000 link
## 14 2018-03-01T12:15:21+0000 photo
## 15 2018-03-01T12:38:21+0000 link
## 16 2018-03-01T13:22:09+0000 link
## 17 2018-03-01T14:22:42+0000 video
## 18 2018-03-01T14:42:02+0000 link
## 19 2018-03-01T14:44:54+0000 link
## 20 2018-03-01T16:39:38+0000 link
##
## 1 http://bbc.in/2F
## 2 http://bbc.in/2F
## 3 http://bbc.in/2F
## 4 http://bbc.in/2F
## 5 http://bbc.in/2C
## 6 http://bbc.in/2F
## 7 http://bbc.in/2F
## 8 http://bbc.in/2H
## 9 http://bbc.in/2F
## 10 http://bbc.in/2F
## 11 http://bbc.in/2F
## 12 https://www.facebook.com/bbcafrica/photos/a.292202500228.180573.285361880228/10156361408115229/?t
## 13 http://bbc.in/2F
## 14 https://www.facebook.com/bbcafrica/photos/a.292202500228.180573.285361880228/10156361462485229/?t
## 15 http://bbc.in/2F
## 16 http://bbc.in/2F
## 17 https://www.facebook.com/bbcafrica/videos/1015636175232
## 18 http://bbc.in/2F
## 19 http://bbc.in/2H
## 20 http://bbc.in/2F
## story comments_count shares_count love_count haha_count wow_count
## 1 NA 169 1542 560 39 205
## 2 NA 12 60 24 0 5
## 3 NA 3 15 4 0 0
## 4 NA 23 141 NA NA NA
## 5 NA 1 7 1 0 0
## 6 NA 25 45 0 6 2
## 7 NA 40 29 21 0 2
## 8 NA 3 5 2 0 1
## 9 NA 10 4 1 0 0
## 10 NA 33 17 3 0 1
## 11 NA 3 0 1 0 0
## 12 NA 4 4 4 2 0
## 13 NA 17 21 2 2 1
## 14 NA 101 14 14 7 0
## 15 NA 6 9 0 5 2
## 16 NA 26 14 1 0 0
## 17 NA 28 68 17 1 11
## 18 NA 3 5 1 0 0
## 19 NA 25 13 1 0 2
## 20 NA 7 5 3 0 0

```

##	sad_count	angry_count
## 1	24	1
## 2	0	0
## 3	0	0
## 4	NA	NA
## 5	0	0
## 6	9	2
## 7	1	0
## 8	0	0
## 9	0	2
## 10	14	8
## 11	0	0
## 12	0	0
## 13	0	1
## 14	0	0
## 15	0	0
## 16	20	4
## 17	1	1
## 18	0	0
## 19	4	1
## 20	0	0