

```
library(tidyverse)  
rladies_global %>%  
  filter(city == 'Charlottesville')
```



- Wifi: **VividCortex_Guest**
- Follow us on Twitter:
@RLadiesCville

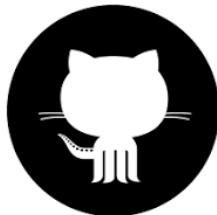
Social Media Analysis with R

November 27, 2018

Sponsors + Friends (for now)



VividCortex



Look at all of this
empty space!



Upcoming Events

- January – Rstudio::conf(2019)
- February 20 – C'ville UseR Meetup: Six decades of personal computing and some consequences for R
- March 2 – Global Diversity CFP Day: more info to come



meetup



Wednesday, February 20, 2019

February meetup - Six decades of personal computing and some consequences for R

global diversity CFP day

Have you always wanted to become a tech conference speaker?

Let 2019 be the year that you make that dream a reality!

Save the date - Saturday 2nd March 2019

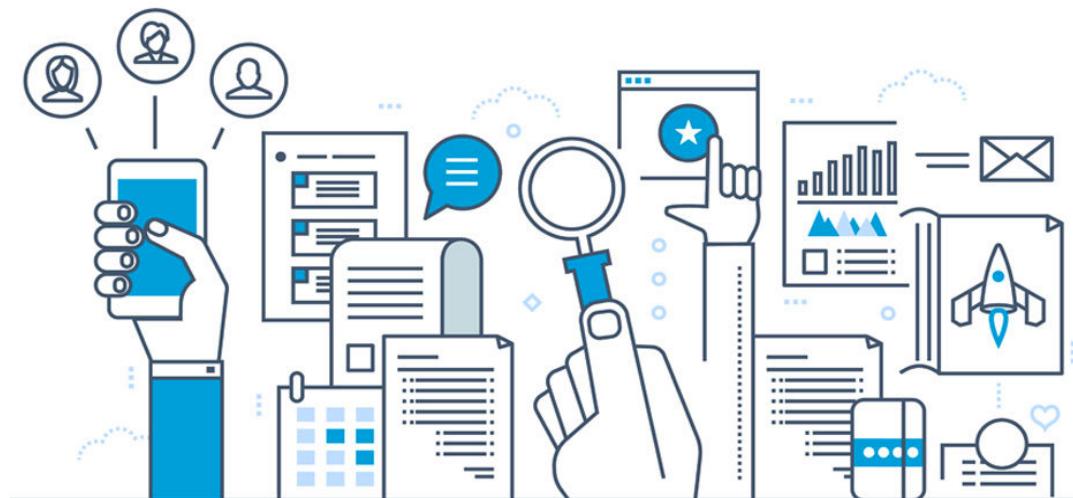


Agenda for today:

- Intro to RESTful APIs
- Twitter Data Mining Workflow
- Sentiment Analysis Example

Ways to analyze social media data

- Frequency
- Sentiment
- Link
- Action
- Patterns
- Streaming



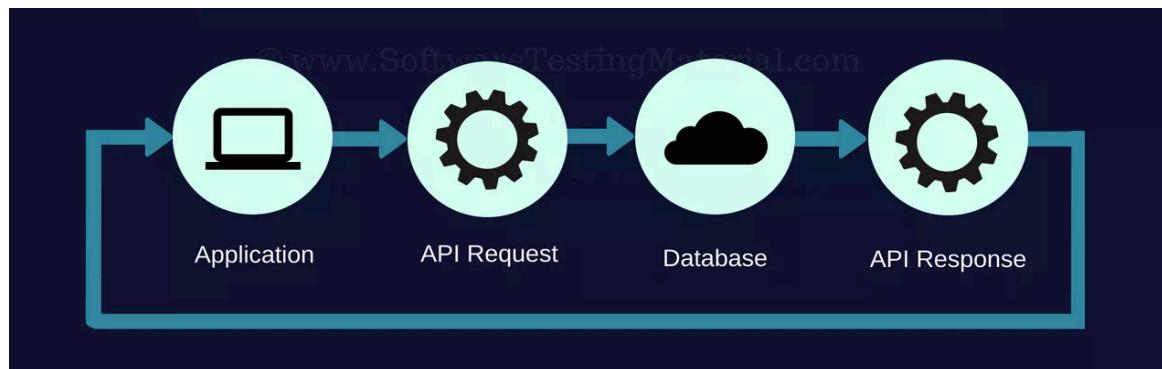


1. What is a RESTful API?



API = Application Programming Interface

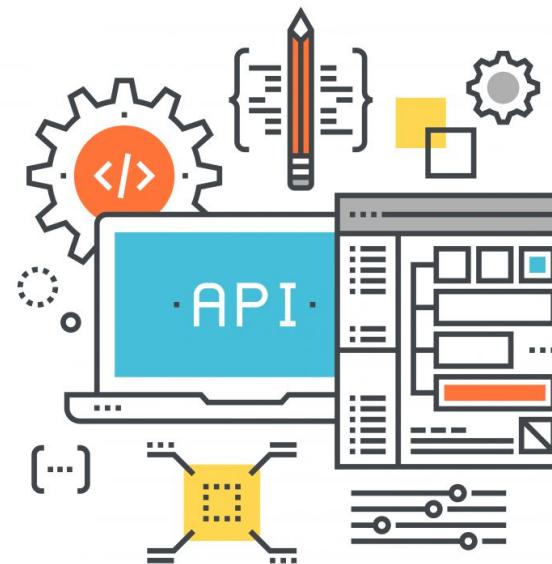
- It is the middle man that allows to applications to talk to each other



What is a RESTful API?

- REpresentational State Transfer technology is an architectural style and approach to communications used in web services development.

1. Client-Server
2. Stateless
3. Cacheable
4. Uniform Interface
5. Layered Systems
6. Code on Demand





Why?

- Almost all social media websites provide RESTful APIs
- Content is usually JSON (easy to read/parse)
- Authentication and authorization is through **Oauth**

Open Authorization Protocol

- Four roles involved:

1. **Resource Owner:** the User (you)



2. **Authorization Server:** the API (Twitter Dev console)



3. **Resource Server:** the stored data (Twitter Data)



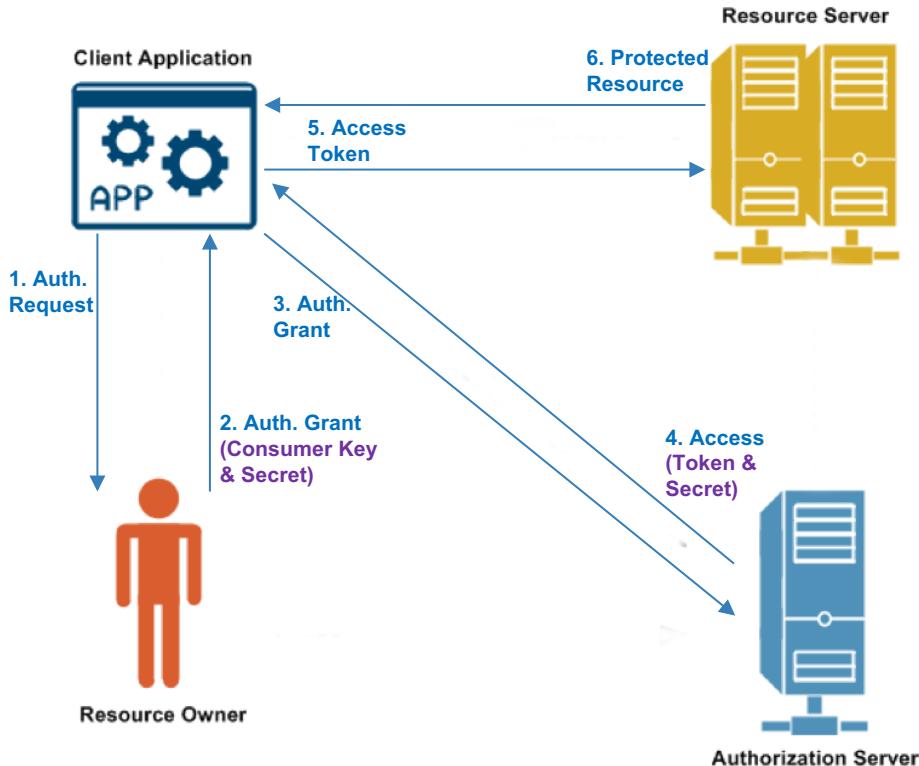
4. **Client** the Application (your Twitter app.)



OAuth Workflow: Authorization & Access



1. **Authorization Request:** the *app* requests authorization to access the service resources from the *user*
2. **Authorization Grant:** the *user* allows the request and the *app* receives an authorization grant
3. **Authorization Grant:** the *app* requests an access token from the *authorization server (API)* by showing 1st authorization grant
4. **Access Token:** if the authorization grant is valid, the *authorization server* issues an access token to the *app*
5. **Access Token:** the *app* requests the data from the *resource server (API)* and presents access token for authentication
6. **Protected Resource:** if the access token is valid, the *resource server* serves the protected resource (data) to the application





Ok so how do I do that with data?

- First you need to access the Twitter API:

The screenshot shows the top navigation bar of the Twitter Developer website. It includes links for 'Developer', 'Use cases', 'Products', 'Docs', and 'More'. There is also an 'Apply' button, a search icon, and a user profile icon.

Get started with Twitter APIs and tools.

Apply for access.

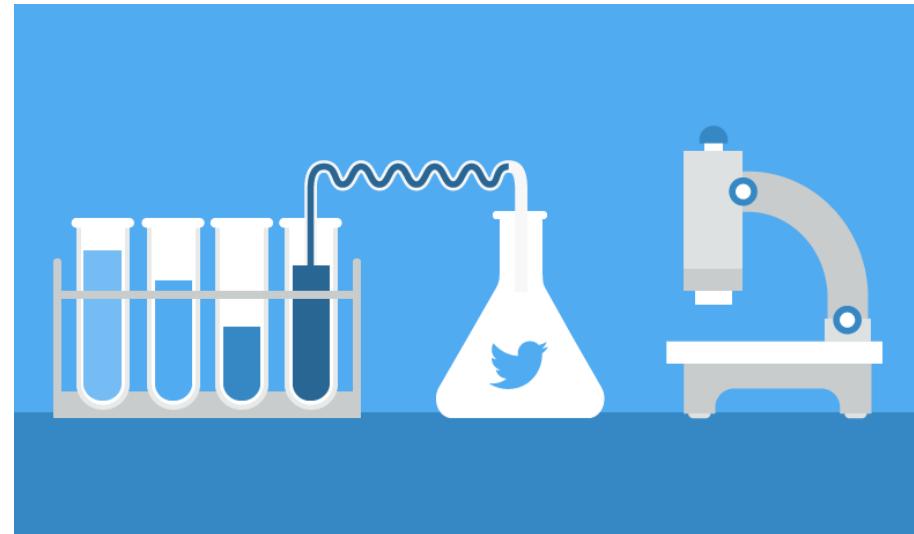
All new developers must apply for a developer account to access Twitter APIs. Once approved, you can begin to use our standard APIs and our new premium APIs.

[Apply for a developer account >](#)

- Register your Twitter account at developer.twitter.com (security questionnaire and approval process ~ 3 days)
- Once your Developer account has been approved, create an app at app.twitter.com

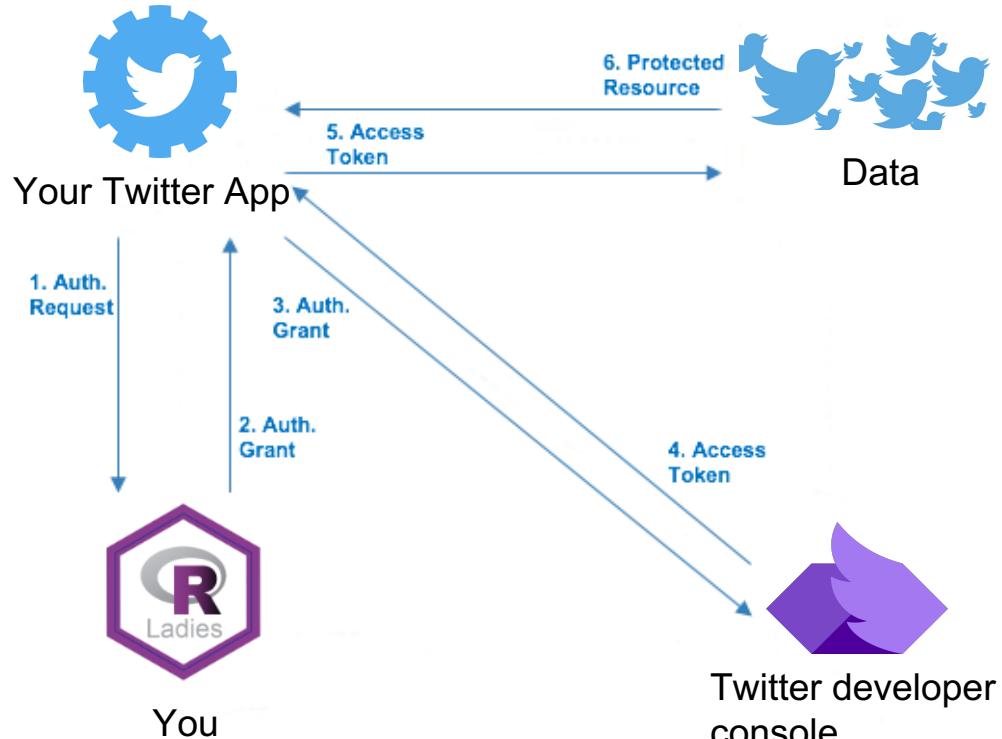


2. Mining Twitter Data



The Big Picture

- Register your app on Twitter and get & store the security keys
- Download the required R packages
- Connect Twitter to R
- Mine the data
- Clean the data
- Analyze & present findings
- Be awesome



Challenges

- Unstructured data
- Incomplete and untidy
- API rate limits



10:33 PM - 1 Jan 13 · Embed this Tweet

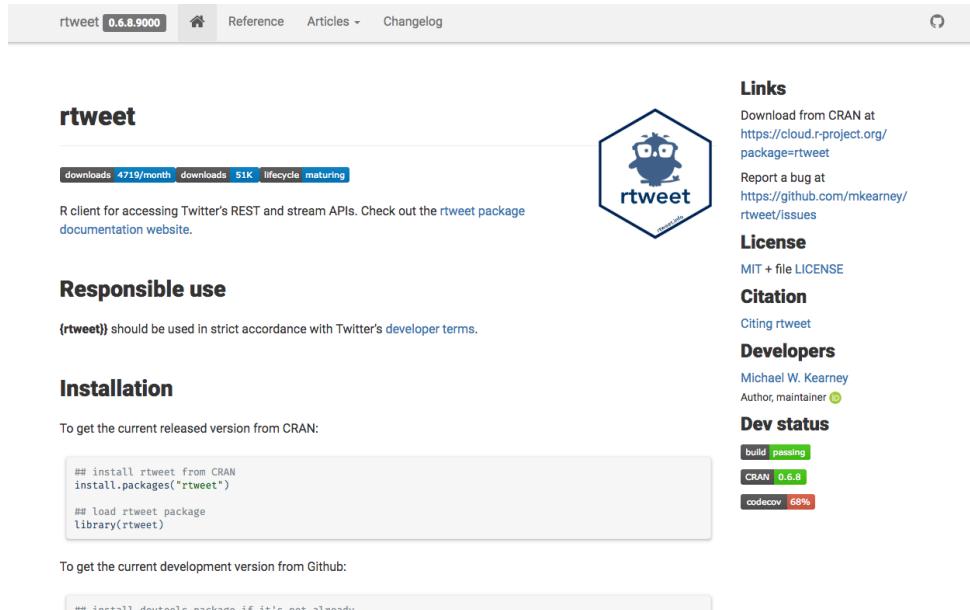
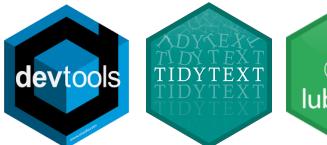


Set up for Success: Packages

```
library(rtweet)
```

The rtweet package, by Michael Kearney has the best documentation and instructions to mining Twitter data.

Also:



The screenshot shows the rtweet package page on CRAN. At the top, there's a navigation bar with links for Home, Reference, Articles, and Changelog. Below the navigation, the package name "rtweet" is displayed along with its version "0.6.8.9000". A large owl logo is prominently featured. The main content area includes sections for "Responsible use" (warning about developer terms), "Installation" (instructions to get the current released version from CRAN), and "Dev status" (showing build, CRAN, and codecov status). The "Links" section provides download and bug reporting information.

rtweet 0.6.8.9000

Reference Articles Changelog

rtweet

downloads 4739/month downloads 51K lifecycle maturing

R client for accessing Twitter's REST and stream APIs. Check out the [rtweet](#) package documentation website.

Responsible use

`{rtweet}` should be used in strict accordance with Twitter's [developer terms](#).

Installation

To get the current released version from CRAN:

```
## install rtweet from CRAN
install.packages("rtweet")  
## load rtweet package
library(rtweet)
```

To get the current development version from Github:

```
## install devtools package if it's not already
```

Links

Download from CRAN at <https://cloud.r-project.org/package=rtweet>
Report a bug at <https://github.com/mkearney/rtweet/issues>

License

MIT + file LICENSE

Citation

Citing rtweet

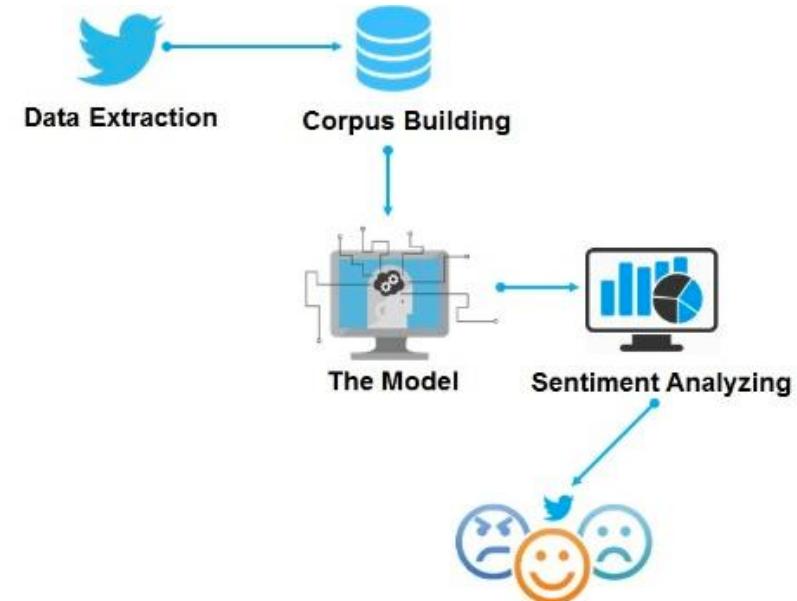
Developers

Michael W. Kearney
Author, maintainer 

Dev status

build passing
CRAN 0.6.8
codecov 68%

3. Twitter Sentiment Analysis Example



Question: is Donald Trump mostly a jerk on Twitter?

- Extract & tidy data from his handle
- Apply sentiment analysis
- Analyze sentiment breakdown
- Make predictions



[Launch Markdown file](#)





Thanks!

- <https://www.tidytextmining.com/>
- <https://rtweet.info/>
- <http://rpubs.com/wctucker/302110>

 RLadiesCville

 Samantha_Toet

 SamanthaToet