



Tweet Collections

Multimedia, Hypertext, and Information Access

Dr. Fox

Virginia Tech, Blacksburg, VA 24061

4/26/18

By: Brandon Chang, Kirk Chenault, Chris Keener, Joseph Widrig



Outline

- Overview
- Testing
 - Assessment
- Deliverables
- Accomplishments
 - Lessons Learned
- Ideas for the Future
- Acknowledgements/References

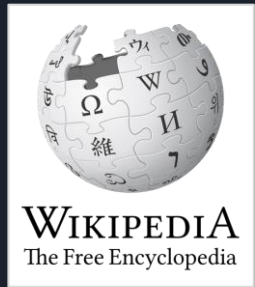
Overview



1



2



3

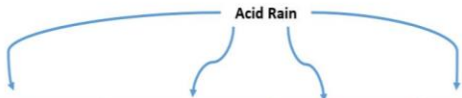


4



Overview Further Details

- Tweet collections may be vaguely named
- CSV contains these names
- We must assign meaning/more data

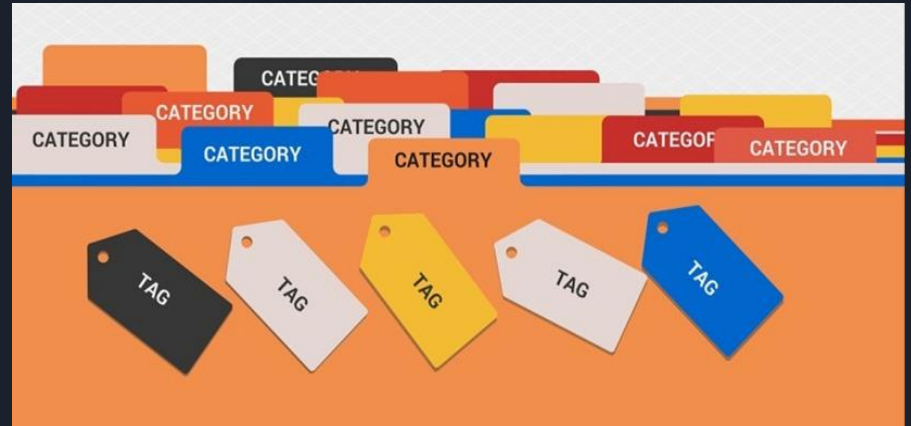


| Database | ID | Source | Wikipedia | Description | Category | Event Name | Count | Latest Date |
|----------|-------|--------|--------------------------------------|---|--|------------------------|-------|-------------|
| GIVEN | GIVEN | GIVEN | Link to an article about 'acid rain' | The most comprehensive and relevant sentence found in the Wikipedia article | Search pre-existing categories for the most relevant for 'acid rain' | Title of the Wiki page | GIVEN | GIVEN |

```
[cbrandon@boxelder TweetCollections]$ ./csv_parser.py
The keywords are:
The keywords are: LAX shooting
The keywords are: #Isaac
The keywords are: hurricane sandy
The keywords are: hurricane
The keywords are: hurricane isaac
The keywords are: @NOAA
The keywords are: California storms
The keywords are: #egypt
The keywords are: #jan25
The keywords are: Nigerian school attack
The keywords are: Nigeria school attack
The keywords are: Connecticut shooting
The keywords are: connecticut school shooting
The keywords are: kentucky shooting
The keywords are: christiansburg mall shooting
The keywords are: santa monica shooting
The keywords are: atlanta school shooting
```

Testing

- Manual review of output from program
- Noticed a need for more categories
- Developing methods for category suggestion
- Could always make more categories





Assessment

- Make current design developer friendly
- Redesign code into a modular format
- Flesh out documentation

>> Header comments
>> Meaningful variable names

>> Encapsulation
>> Easily editable

>> For future developers ease of use

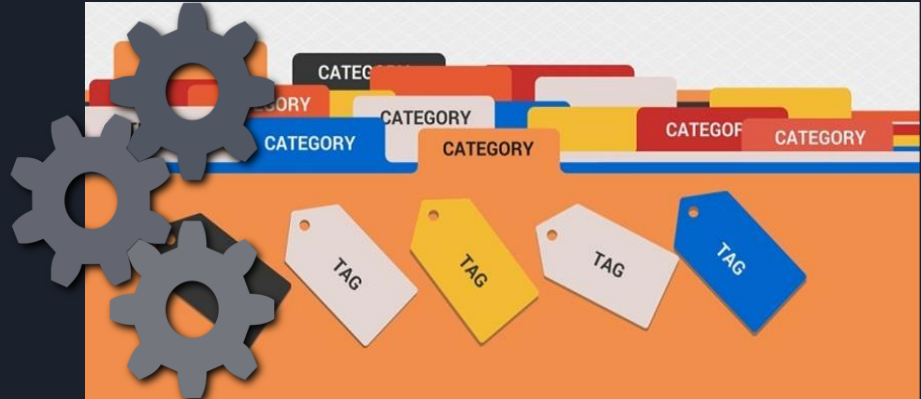
Deliverables

- Developer friendly script
- Documentation
- Excel Spreadsheet



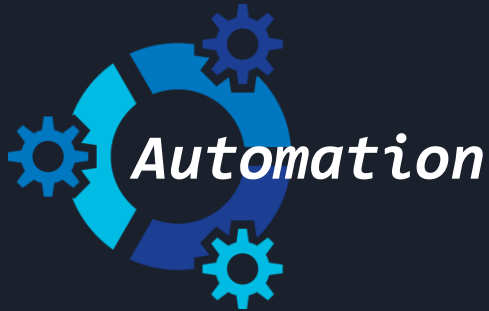
Ideas for the Future

- Modify search terms to improve results
- Automatically create categories when criteria are met
- More advanced metric to measure automated success



Lessons Learned

- Automate testing whenever possible
- Better to stick to one heuristic and improve it
- Frequently interact with client for an agile development process



Acknowledgements & References

Client: **Liuqing Li:** liuqing@vt.edu
Ziqian Song: ziqian@vt.edu



Images: <https://upload.wikimedia.org/wikipedia/commons/thumb/b/b3/Wikipedia-logo-v2-en.svg/1200px-Wikipedia-logo-v2-en.svg.png>

<https://itunes.apple.com/us/app/microsoft-excel/id586683407?mt=8>

GETAR: Thanks go to NSF for support by grant IIS-1619028

GETAR: <http://www.ctrnet.net/>