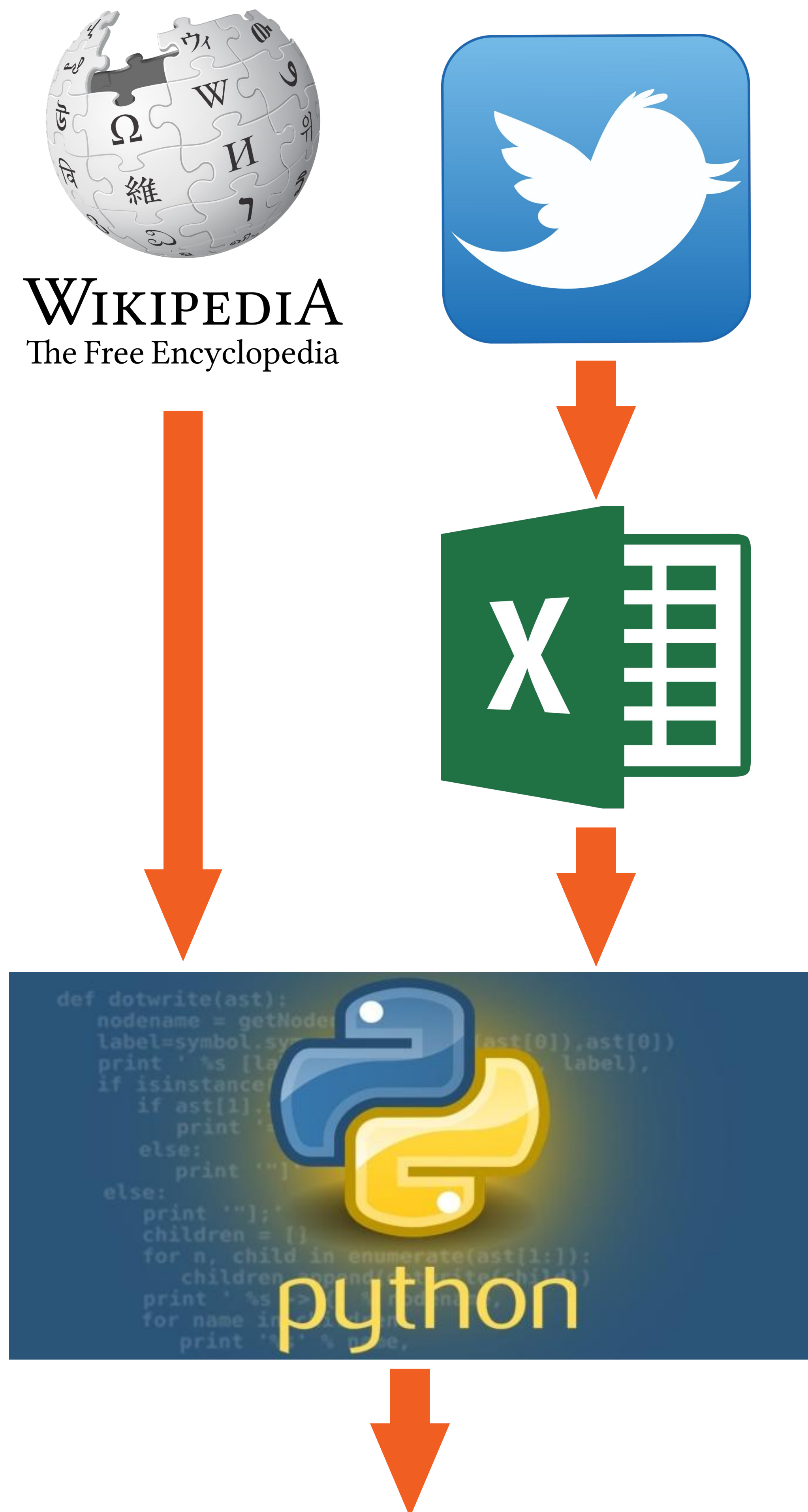


Tweet Collections

Kirk Chenault, Chris Keener, Brandon Chang, Joe Widrig

OVERVIEW:



Collection Terms	Wikipedia	Description	Tags
LAX shooting	https://en.wikipedia.org/wiki/Paul_Ciancia_fatally_shot_a_TSA_officer_and_airport_shooting	Paul Ciancia fatally shot a TSA officer and	airport shooting
#Isaac	https://en.wikipedia.org/wiki/Hurricane_Isaac_formed_on_August_21,_2011	Hurricane Isaac formed on August 21, 2011	hurricane
hurricane sandy	https://en.wikipedia.org/wiki/Hurricane_Sandy,_also_referred_to_as_Superstorm_Sandy	Hurricane Sandy, also referred to as Superstorm Sandy	hurricane
hurricane	https://en.wikipedia.org/wiki/Hurricane	Intended for any hurricane. Originally for	hurricane
hurricane isaac	https://en.wikipedia.org/wiki/Hurricane_Isaac_formed_on_August_21,_2011	Hurricane Isaac formed on August 21, 2011	hurricane
@NOAA	http://www.noaa.gov/	National Oceanic and Atmospheric Administration	hurricane
California storms	https://www.noaa.gov/	Rain storm that started one 3/1/2014 that	rain storm
#egypt	https://en.wikipedia.org/wiki/2011_Egyptian_revolution	Originally for Egyptian revolution, but not	revolution

Figure 1: Snippet of results

PROBLEM:

- Excel sheet contains no metadata

CONSTRAINTS:

- Use Twitter collection input for data
- Use ontology input for categories

SOLUTION:

- Use Python to pull Wikipedia articles
- Filter irrelevant results
- Fill out empty cells in excel
- Return solution as new file

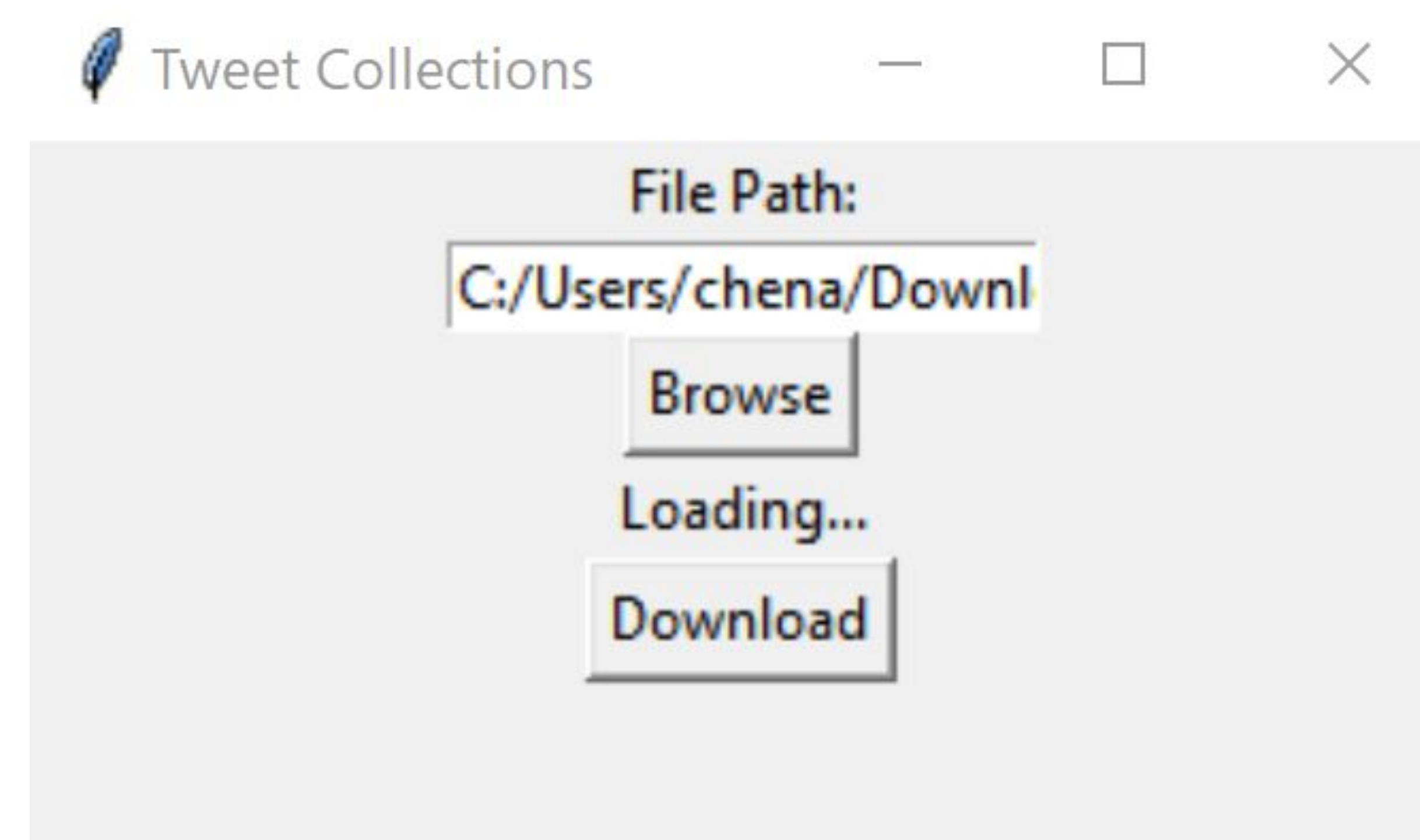


Figure 2: Front end for our tool

- Simple interface
- Input excel sheet
- Output generated excel sheet

#CarryTheLoad
#DefendTheGIBill
#GoSilent
#GratefulNation
#HappyMemorialDay
#Heroes
#HonorThem
#MemorialDayWeekend

Figure 3: Example Hashtags

HEURISTICS:

- Pre-existing Categories
- Date Relevance
- Term Frequency
- Parsing Hashtags
- Improving Source Detection

Category_1
/Disaster_Types/Manmade/Conflict_Based/Terrorism/Shooting
/Disaster_Types/Natural/Geophysical/Storm/Hurricane
/Disaster_Types/Natural/Geophysical/Storm/Hurricane
/Disaster_Types/Natural/Geophysical/Storm/Hurricane
/Disaster_Types/Natural/Geophysical/Storm/Hurricane
/Disaster_Types/Natural/Geophysical/Storm/Hurricane
/Disaster_Types/Natural/Geophysical/Storm/Rain
/Disaster_Types/Manmade/Conflict_Based/Revolution
/Disaster_Types/Manmade/Conflict_Based/Revolution
/Disaster_Types/Manmade/Conflict_Based/School_Shooting
/Disaster_Types/Manmade/Conflict_Based/School_Shooting
/Disaster_Types/Manmade/Conflict_Based/School_Shooting
/Disaster_Types/Manmade/Conflict_Based/School_Shooting
/Disaster_Types/Manmade/Conflict_Based/School_Shooting
/Disaster_Types/Manmade/Conflict_Based/School_Shooting

Figure 4: Category Structure

Tweet Collections

Kirk Chenault, Chris Keener, Brandon Chang, Joe Widrig

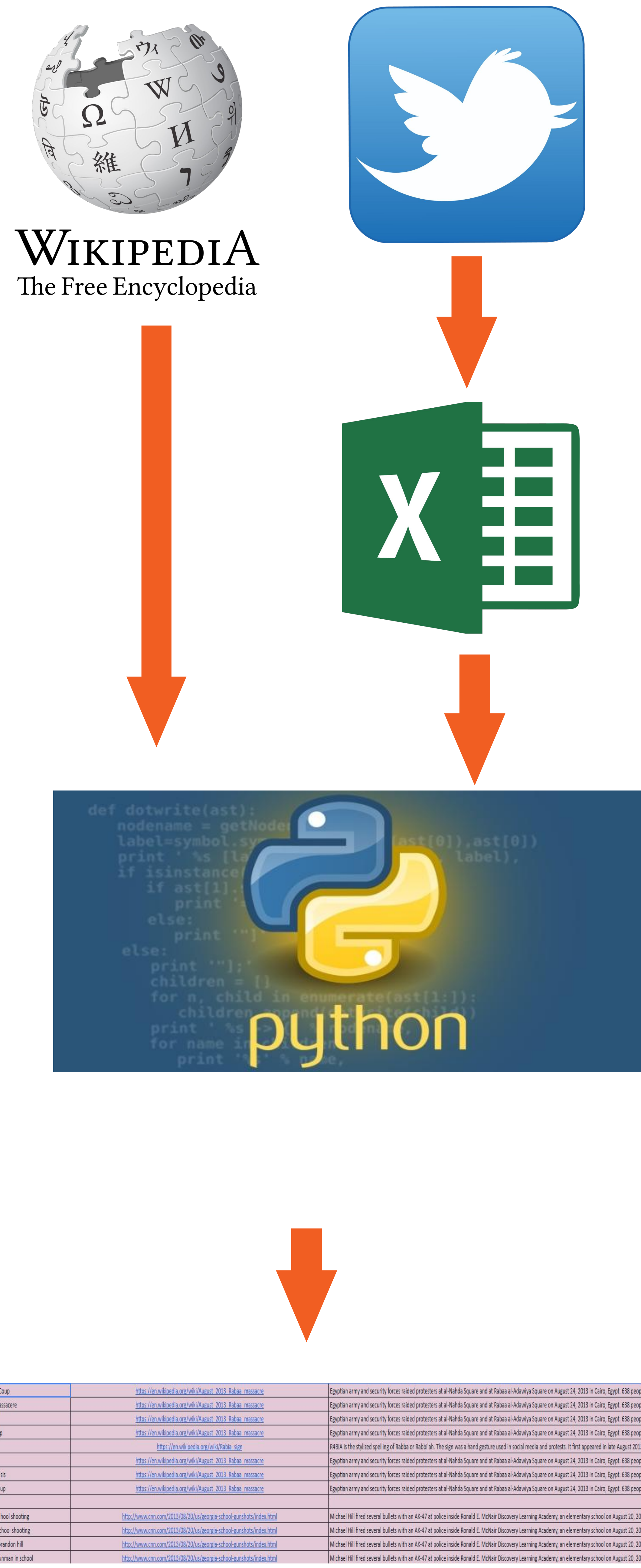


Figure 1

PROBLEM:

- Excel sheet contains no metadata

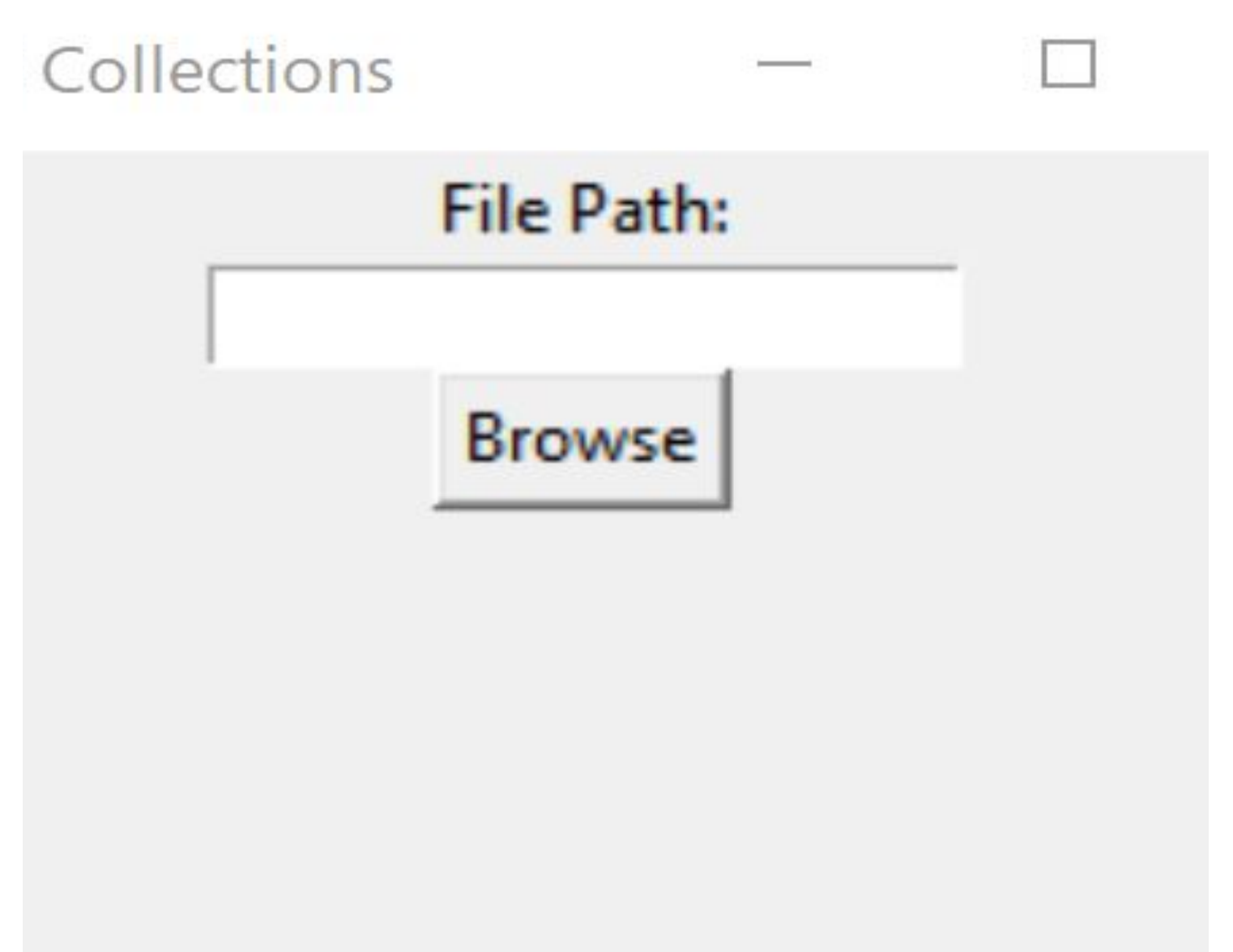
CONSTRAINTS:

- Use Twitter collection input for data
- Use ontology input for categories

SOLUTION:

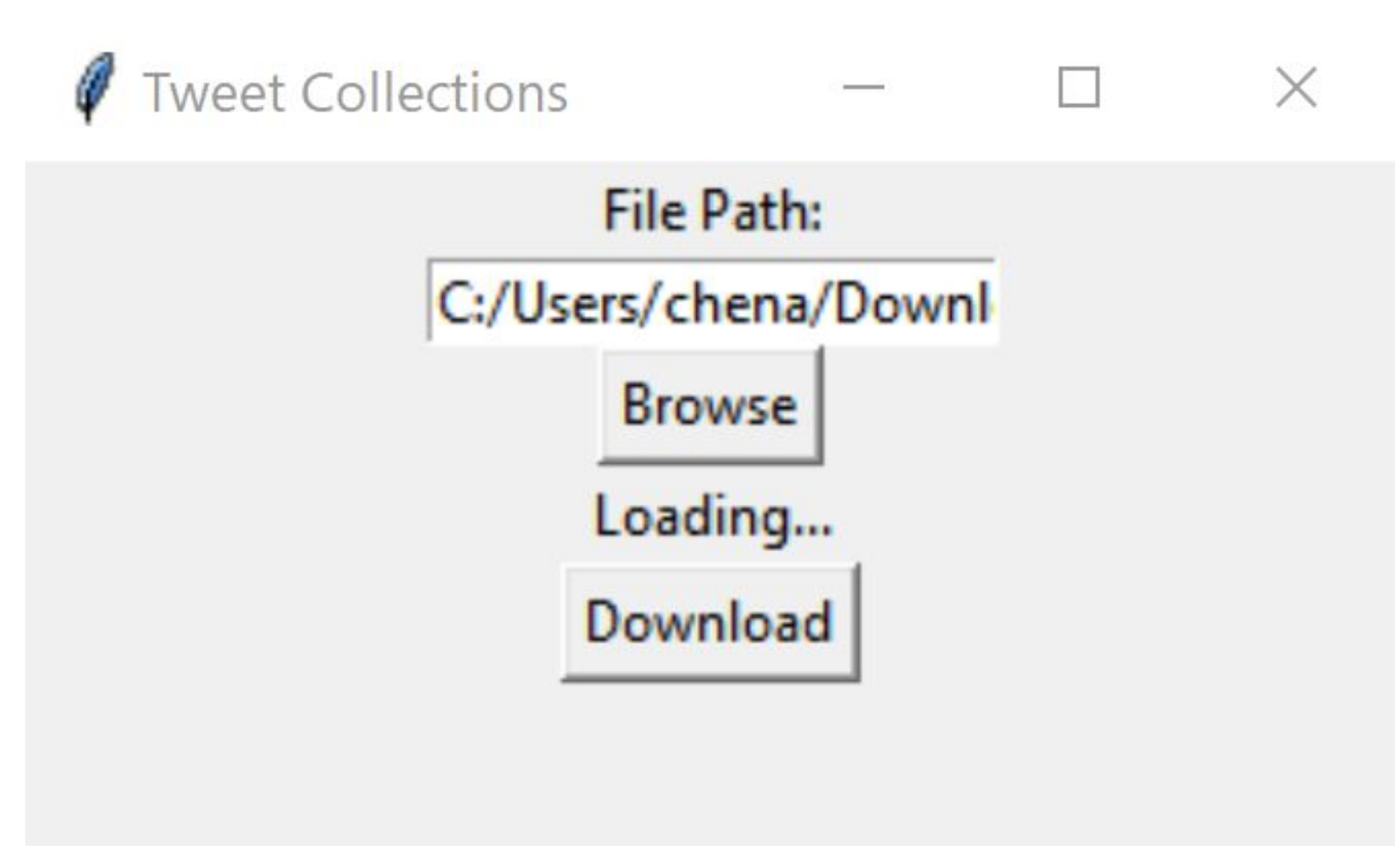
- Use Python to pull Wikipedia articles
- Filter irrelevant results
- Fill out empty cells in excel
- Return solution as new file

Figure 2



- Simple interface
- Input excel sheet

Figure 3



- Output generated excel sheet

#CarryTheLoad
#DefendTheGIBill
#GoSilent
#GratefulNation
#HappyMemorialDay
#Heroes
#HonorThem
#MemorialDayWeekend

Problems and future work
Asdfasdfasdfasdfasdfasdfsdf
Asdfasdfasdfasdfasdfasdfasfas
Asdfasdfasdffasfdasdf

```
In [7]: # Import urllib2
import urllib2
# temp = urllib2.urlopen("https://en.wikipedia.org/w/api.php?action=query&title=Main%20Page&prop=revisions&rvprop=content")
# temp = wikipedia.search("Pizza")
# temp = wikipedia.page("Pizza").content
temp = wikipedia.search("Pizza")

out[1]: ['Pizza',
        'Hawaiian pizza',
        'Chicago-style pizza',
        'Blackjack Pizza',
        'History of pizza',
        'Newman pizza',
        'Stanzio Pizza']
```

Figure 7

Problems and future work
Asdfasdfasdfasdfasdfasdfsdf
Asdfasdfasdfasdfasdfasdfasfas
Asdfasdfasdffasfdasdf
sdfasdfasdfasdfasdfasfas
Asdfasdfasdffasfdasdf

sdfasdfasdfasdfasdfasfas
Asdfasdfasdffasfdasdf
sdfasdfasdfasdfasdfasfas
Asdfasdfasdffasfdasdfsdfasdfasdfa
sfasdfasfas
Asdfasdfasdffasfdasdf

Tweet Collections

Kirk Chenault, Chris Keener, Brandon Chang, Joe Widrig

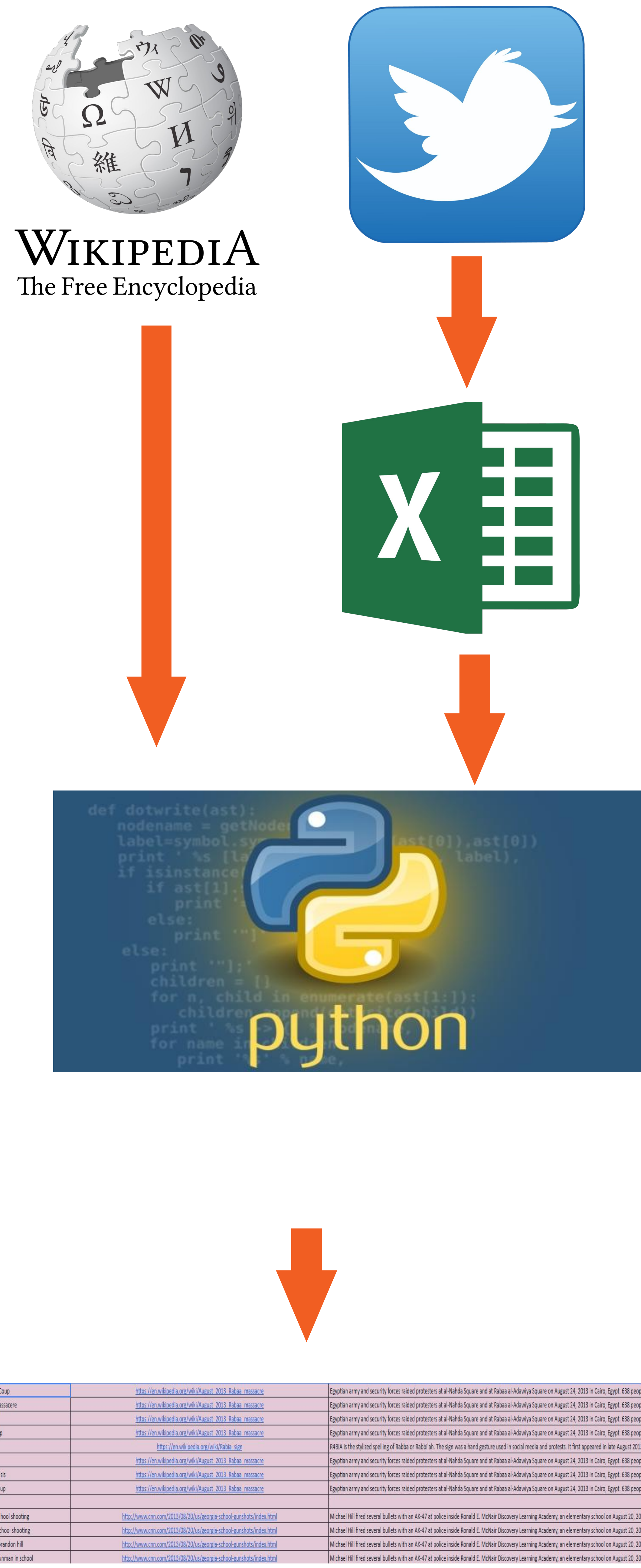


Figure 1

PROBLEM:

- Excel sheet contains no metadata

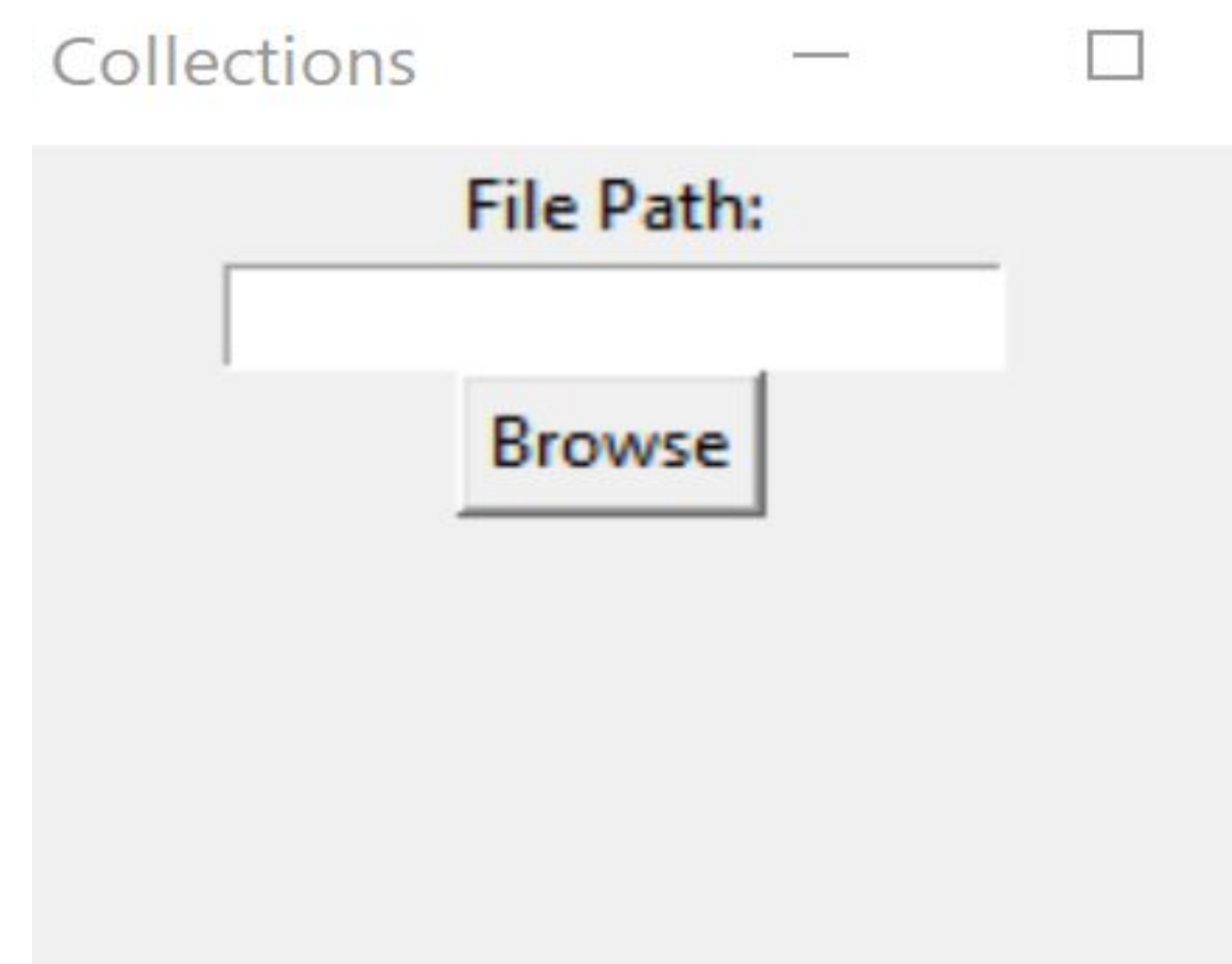
CONSTRAINTS:

- Use Twitter collection input for data
- Use ontology input for categories

SOLUTION:

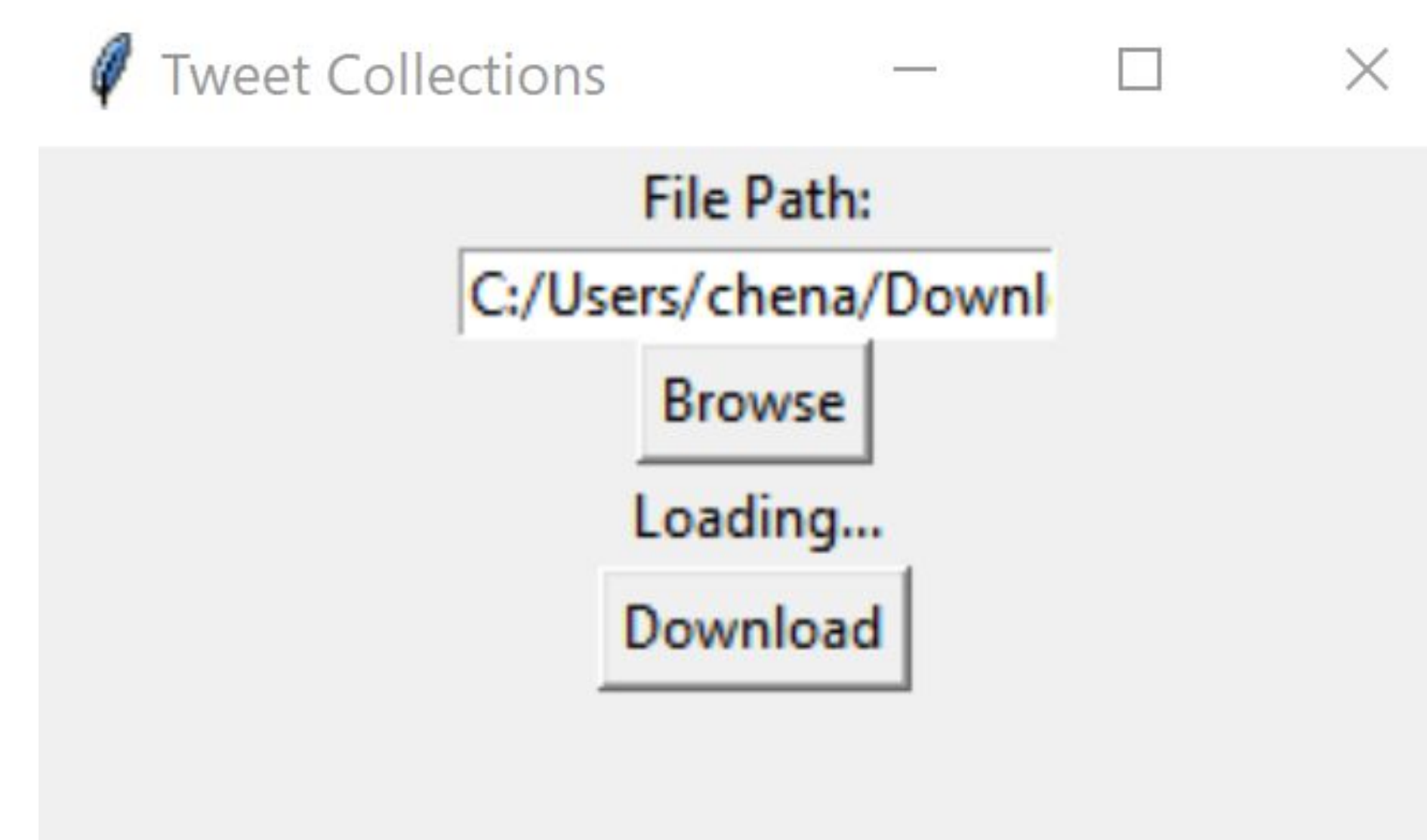
- Use Python to pull Wikipedia articles
- Filter irrelevant results
- Fill out empty cells in excel
- Return solution as new file

Figure 2



- Simple interface
- Input excel sheet

Figure 3



- Output generated excel sheet

#CarryTheLoad
#DefendTheGIBill
#GoSilent
#GratefulNation
#HappyMemorialDay
#Heroes
#HonorThem
#MemorialDayWeekend

Problems and future work
Asdfasdfasdfasdfasdfasdfsdf
Asdfasdfasdfasdfasdfasdfasfas
Asdfasdfasdffasfdasdf

```
In [7]: # Import urllib2
import urllib2
# temp = urllib2.urlopen("https://en.wikipedia.org/w/api.php?action=query&title=Main%20Page&prop=revisions&rvprop=content")
# temp = wikipedia.search("Pizza")
# temp = wikipedia.page("Pizza").content
temp = wikipedia.search("Pizza")

out[1]: ['Pizza',
        'Hempelian pizza',
        'Domino's Pizza',
        'Pizza Hut',
        'Chicago-style pizza',
        'Blackjack Pizza',
        'History of pizza',
        'Neapolitan pizza',
        'White Pizza']
```

Figure 7

Problems and future work
Asdfasdfasdfasdfasdfasdfsdf
Asdfasdfasdfasdfasdfasdfasfas
Asdfasdfasdffasfdasdf
sdfasdfasdfasdfasdfasfas
Asdfasdfasdffasfdasdf

sdfasdfasdfasdfasdfasfas
Asdfasdfasdffasfdasdf
sdfasdfasdfasdfasdfasfas
Asdfasdfasdffasfdasdfsdfasdfasdfa
sfasdfasfas
Asdfasdfasdffasfdasdf