



# TRUSTED NEWS CODE USER GUIDE

Week 7 Assignment

## ABSTRACT

This guide was created for UMUC Course, CMSC 495, Group 3 and describes the requirements and usage of the Trusted News Code (TNC) Software.

### Group 3 Members

Name: Christiano, Andrew

Name: Fernandez, Yrume

Name: Orwick, Brian

Name: Sell, Julia

Class: CMSC 495 - Current Trends and Projects in  
Computer Science

Professor: Dr. Hung Dao

Due: 07 October 2018

### Version Control

Revision #	Date	Name	Descriptions	Contact Info
TNC_UG_0001	10/7/2018	Yrume Fernandez	Created	yrume.fernandez@gmail.com
TNC_UG_0002	10/7/2018	Brian Orwick	Modified	orwick12@outlook.com
TNC_UG_0003	10/7/2018	Andrew Christiano	Modified	ajchristiano91@gmail.com
TNC_UG_0004	10/7/2018	Julia Sell	Modified	selljm14@gmail.com

## Table of Contents

Introduction .....	4
a. Description .....	4
Requirements .....	4
a. System Requirements .....	4
Installation .....	5
Step 1: (Download Code) .....	5
Step 2: (Install Modules) .....	5
Operation .....	6
Step 1: (Run the Program) .....	6
Step 2: (Using Website) .....	7

## Introduction

### Trusted News Code (TNC)

#### a. Description

The Trusted News Code (TNC) software is a Python Web Application which searches the articles on the websites of provided news organizations to retrieve news articles, count the number of times an article is found, and provide the user a list of articles prioritized by their trustworthiness with a link to the main article. This guide is meant for the administrator of the TNC program, as regular users have no need to install the program as it is regularly accessed via the internet.

## Requirements

#### a. System Requirements

The following are the minimum required system and software to support a Python based news indexing system, using Flask, Newspaper and SQLite modules:

- Processors:
  - a. Intel® Core™ i5 processor 4300M at 2.60 GHz
- Memory:
  - a. 8 GB of RAM
- Disk space:
  - a. 3 GB
- Operating systems:
  - a. Windows® 10, macOS (Sierra), or Linux (Ubuntu 16.04)
- Browsers: Microsoft Edge, Google Chrome, or Mozilla Firefox
- Applications:
  - a. Python version 3.7.0
    - i. Additional Python Modules (Install with pip):
      1. Flask
      2. Newspaper3k

The Trusted News Code is maintained on the Social Software Development Site GitHub and can be retrieved at <https://github.com/Mercury-6-Solutions/CMS495/tree/master/src>. Users are encouraged to periodically check the site for new updates, codes, or information pertaining to future enhancements. To install the TNC for initial use, please follow the below Installation guidelines.

(Note) This guide assumes that the user has already installed Python 3.7.0 (or greater).

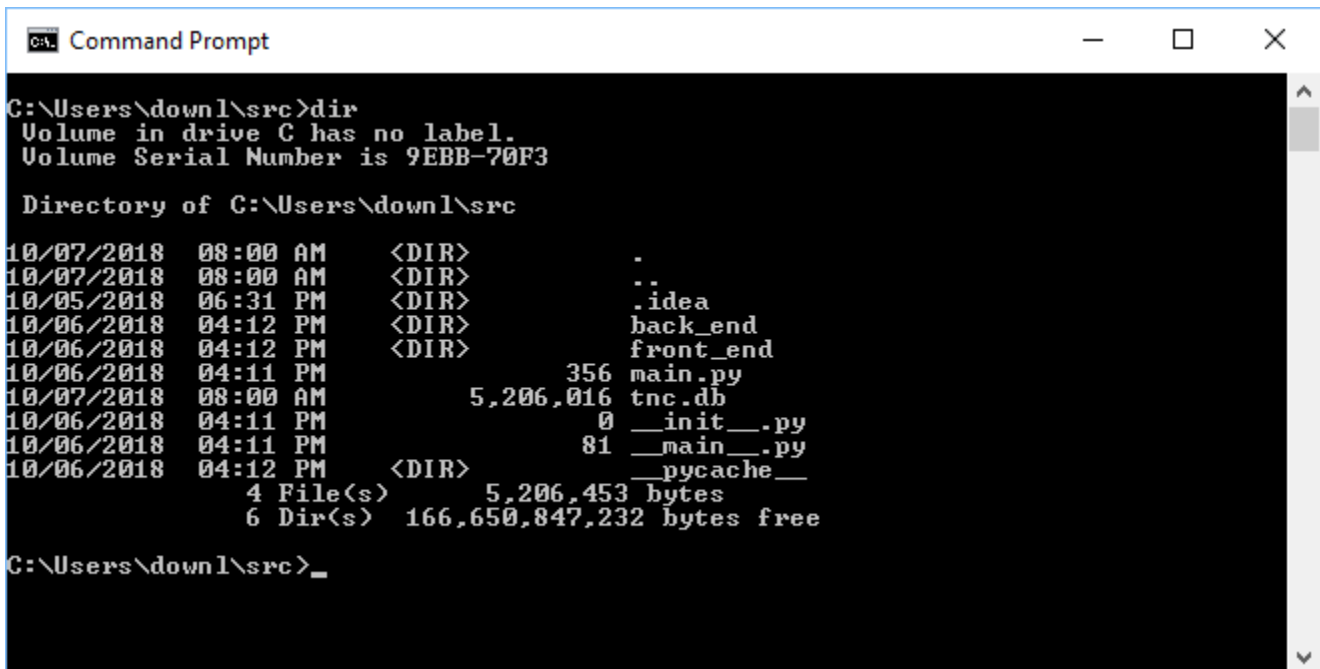
## Installation

### Step 1: (Download Code)

Download the Source Code (src folder) from <https://github.com/Mercury-6-Solutions/CMS495/tree/master/src>.

The contents should be the following:

(Note) Your directory path (C:\xxxx\xxxx\src) may look different depending on where you install the src code.



```
Command Prompt
C:\Users\downl\src>dir
Volume in drive C has no label.
Volume Serial Number is 9EBB-70F3

Directory of C:\Users\downl\src

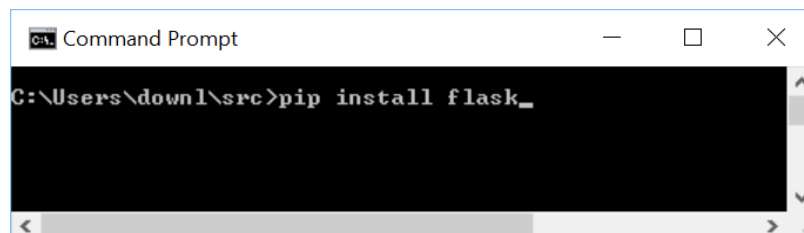
10/07/2018  08:00 AM    <DIR>          .
10/07/2018  08:00 AM    <DIR>          ..
10/05/2018  06:31 PM    <DIR>          .idea
10/06/2018  04:12 PM    <DIR>          back_end
10/06/2018  04:12 PM    <DIR>          front_end
10/06/2018  04:11 PM             356 main.py
10/07/2018  08:00 AM             5,206,016 tnc.db
10/06/2018  04:11 PM              0 __init__.py
10/06/2018  04:11 PM             81 __main__.py
10/06/2018  04:12 PM    <DIR>          __pycache__
               4 File(s)              5,206,453 bytes
               6 Dir(s)  166,650,847,232 bytes free

C:\Users\downl\src>
```

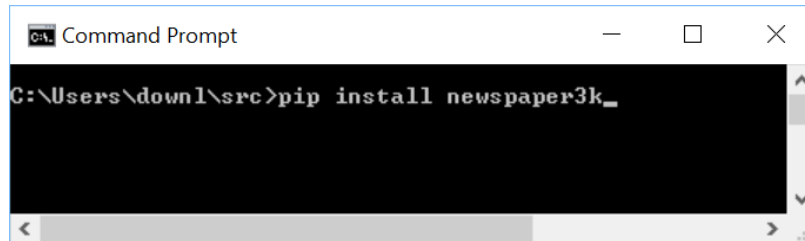
### Step 2: (Install Modules)

Install required python modules (flask, newspaper3k) by running the following commands from the command prompt (command line interface):

- pip install flask
- pip install newspaper3k



```
Command Prompt
C:\Users\downl\src>pip install flask_
```



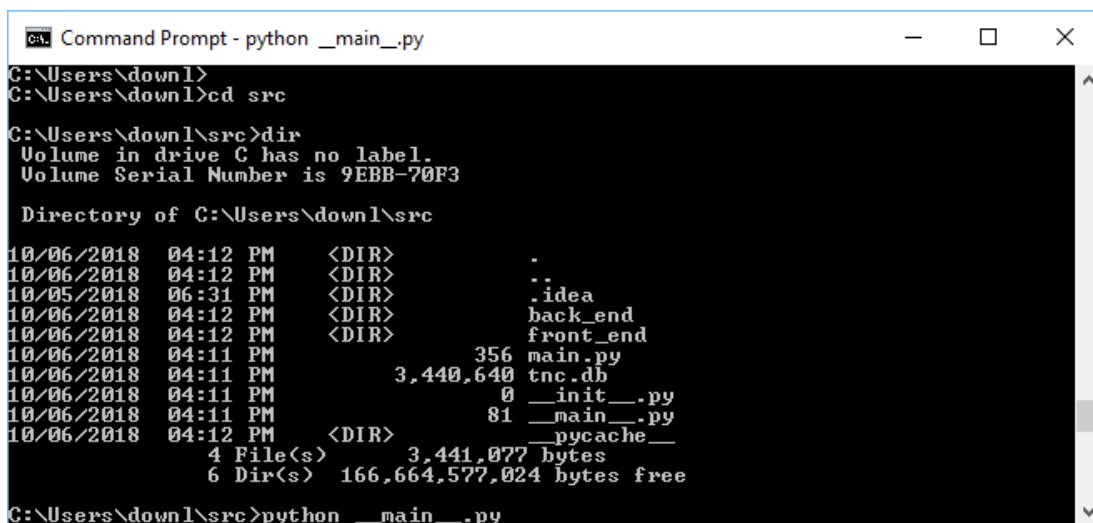
```
C:\Users\downl\src>pip install newspaper3k_
```

## Operation

### Step 1: (Run the Program)

To run the program user opens up a command prompt from within their operating system and performs the following steps:

- a. Go to the src directory
- b. From the src directory type: `python __main__.py`



```
C:\Users\downl>cd src
C:\Users\downl\src>dir
Volume in drive C has no label.
Volume Serial Number is 9EBB-70F3

Directory of C:\Users\downl\src

10/06/2018  04:12 PM    <DIR>          .
10/06/2018  04:12 PM    <DIR>          ..
10/05/2018  06:31 PM    <DIR>          .idea
10/06/2018  04:12 PM    <DIR>          back_end
10/06/2018  04:12 PM    <DIR>          front_end
10/06/2018  04:11 PM                356 main.py
10/06/2018  04:11 PM          3,440,640 tnc.db
10/06/2018  04:11 PM                0 __init__.py
10/06/2018  04:11 PM                81 __main__.py
10/06/2018  04:12 PM    <DIR>          __pycache__
               4 File(s)          3,441,077 bytes
               6 Dir(s)  166,664,577,024 bytes free

C:\Users\downl\src>python __main__.py
```

- c. Program will start retrieving articles, and displaying capture information on command prompt standard output (as shown below)

(Note) Please be patient as this will take some time as program retrieves articles from multiple sources and populates database. The errors you see below are not presented to normal users accessing the TNC program through the internet. These are only visible to the administrator of the TNC program. The errors are showing when articles cannot be downloaded, and therefore cannot be parsed.

```
Command Prompt - python _main_.py
C:\Users\downl\src>python _main_.py
Article 'download()' failed with 404 Client Error: Not Found for url: https://money.cnn.com/2018/09/30/success/pressed-juicery-hayden-slater-fresh-money/index.html on URL https://money.cnn.com/2018/09/30/success/pressed-juicery-hayden-slater-fresh-money/index.html
Article 'download()' failed with 404 Client Error: Not Found for url: https://money.cnn.com/2018/09/30/success/allbirds-fresh-money-joezy-zwillinger/index.html on URL https://money.cnn.com/2018/09/30/success/allbirds-fresh-money-joezy-zwillinger/index.html
Article 'download()' failed with 404 Client Error: Not Found for url: https://money.cnn.com/2018/10/03/tech/amazon-go/index.html on URL https://money.cnn.com/2018/10/03/tech/amazon-go/index.html
Article 'download()' failed with 404 Client Error: Not Found for url: https://money.cnn.com/2018/09/27/tech/grab-drivers-singapore-cargo-uber/index.html on URL https://money.cnn.com/2018/09/27/tech/grab-drivers-singapore-cargo-uber/index.html
Article 'download()' failed with 403 Client Error: Forbidden for url: https://preview.foxnews.com/tech/catch-a-cheater-bank-safely-and-more-tech-qa on URL https://preview.foxnews.com/tech/catch-a-cheater-bank-safely-and-more-tech-qa
Article 'download()' failed with 404 Client Error: Not Found for url: https://ara.reuters.com/video/2018/10/06/-1649?videoId=470858514&newsChannel=internetNews on URL http://ara.reuters.com/video/2018/10/06/-1649?videoId=470858514&newsChannel=internetNews
```

d. Website is running, and program is operating once the following page is displayed:

```
Command Prompt - python _main_.py
Prefix dict has been built successfully.
* Serving Flask app "front_end.Web" (lazy loading)
* Environment: production
  WARNING: Do not use the development server in a production environment.
  Use a production WSGI server instead.
* Debug mode: on
* Restarting with stat
* Debugger is active!
* Debugger PIN: 862-891-584
* Running on http://0.0.0.0:5000/ (Press CTRL+C to quit)
127.0.0.1 - - [07/Oct/2018 08:30:28] "GET / HTTP/1.1" 200 -
127.0.0.1 - - [07/Oct/2018 09:06:30] "GET / HTTP/1.1" 200 -
```

## Step 2: (Using Website)

To test that the website is available to any users who want to access it to retrieve trusted news articles simply enter <http://localhost:5000> on any web browser, and the website will provide a list of articles, sorted from most referenced on the internet to lowest referenced by trusted news sources.

```
id's 1 and 2 have 16.312056737588655 percent word match
url for a is https://cnn.com/2018/10/05/politics/lisa-murkowski-senate-vote-present-kavanaugh/index.html
url for b is https://www.foxnews.com/politics/capitol-police-arrest-101-people-in-washington-dc-amid-kavanaugh-protests
id's 1 and 3 have 31.914893617021278 percent word match
url for a is https://cnn.com/2018/10/05/politics/lisa-murkowski-senate-vote-present-kavanaugh/index.html
url for b is https://www.foxnews.com/opinion/tammy-bruce-susan-collins-speech-tells-us-a-lot-about-the-gop-and-america
id's 1 and 4 have 14.184397163120568 percent word match
url for a is https://cnn.com/2018/10/05/politics/lisa-murkowski-senate-vote-present-kavanaugh/index.html
url for b is https://www.foxnews.com/us/judge-cites-evidence-that-trump-harbors-an-animus-against-non-white-non-european-aliens-in-ruling
id's 1 and 5 have 9.219858156028367 percent word match
url for a is https://cnn.com/2018/10/05/politics/lisa-murkowski-senate-vote-present-kavanaugh/index.html
url for b is https://www.foxnews.com/us/border-patrol-arrests-daca-recipient-accused-of-smuggling-over-34g-worth-of-meth
id's 1 and 6 have 14.893617021276595 percent word match
url for a is https://cnn.com/2018/10/05/politics/lisa-murkowski-senate-vote-present-kavanaugh/index.html
url for b is https://www.foxnews.com/us/mayor-de-blasio-dismisses-homeless-woman-in-nyc-gym-im-doing-my-workout
id's 1 and 7 have 12.056737588652481 percent word match
url for a is https://cnn.com/2018/10/05/politics/lisa-murkowski-senate-vote-present-kavanaugh/index.html
url for b is https://www.foxnews.com/politics/new-york-ag-fires-another-salvo-at-trump-foundation
id's 1 and 8 do not match...
id's 1 and 9 have 19.148936170212767 percent word match
url for a is https://cnn.com/2018/10/05/politics/lisa-murkowski-senate-vote-present-kavanaugh/index.html
url for b is https://www.foxnews.com/politics/susan-collins-is-a-rape-apologist-womens-march-claims-following-kavanaugh-support-announcement
id's 1 and 10 have 7.801418439716312 percent word match
url for a is https://cnn.com/2018/10/05/politics/lisa-murkowski-senate-vote-present-kavanaugh/index.html
url for b is https://www.foxnews.com/entertainment/jersey-shore-star-mike-the-situation-sorrentino-speaks-out-following-8-month-prison-sentence-for-tax-evasion
id's 1 and 11 do not match...
id's 1 and 12 do not match...
id's 1 and 13 do not match...
id's 1 and 14 do not match...
id's 1 and 15 have 7.801418439716312 percent word match
url for a is https://cnn.com/2018/10/05/politics/lisa-murkowski-senate-vote-present-kavanaugh/index.html
url for b is http://www.foxnews.com/entertainment/jersey-shore-star-mike-the-situation-sorrentino-speaks-out-following-8-month-prison-sentence-for-tax-evasion
id's 1 and 16 have 14.893617021276595 percent word match
url for a is https://cnn.com/2018/10/05/politics/lisa-murkowski-senate-vote-present-kavanaugh/index.html
```