Old Dominion University

ODU Digital Commons

Did They Really Tweet That?

Caleb Bradford
Old Dominion University, cbrad022@odu.edu

Michael L. Nelson (Mentor)
Old Dominion University, mnelson@odu.edu

Recommended Citation

Did They Really Tweet That?

● ● ●

Caleb Bradford advised by Dr. Michael Nelson
2022-08-04

An interesting and totally real tweet made by a colleague

Haley Bragg
@haleybragg17

Disinformation on Instagram is literally not a problem.

11:25 AM · Jul 31, 2022

310 Retweets   **214** Quote Tweets   **4.6K** Likes

Fabricated with easy to use web app called Tweetgen

https://www.tweetgen.com/
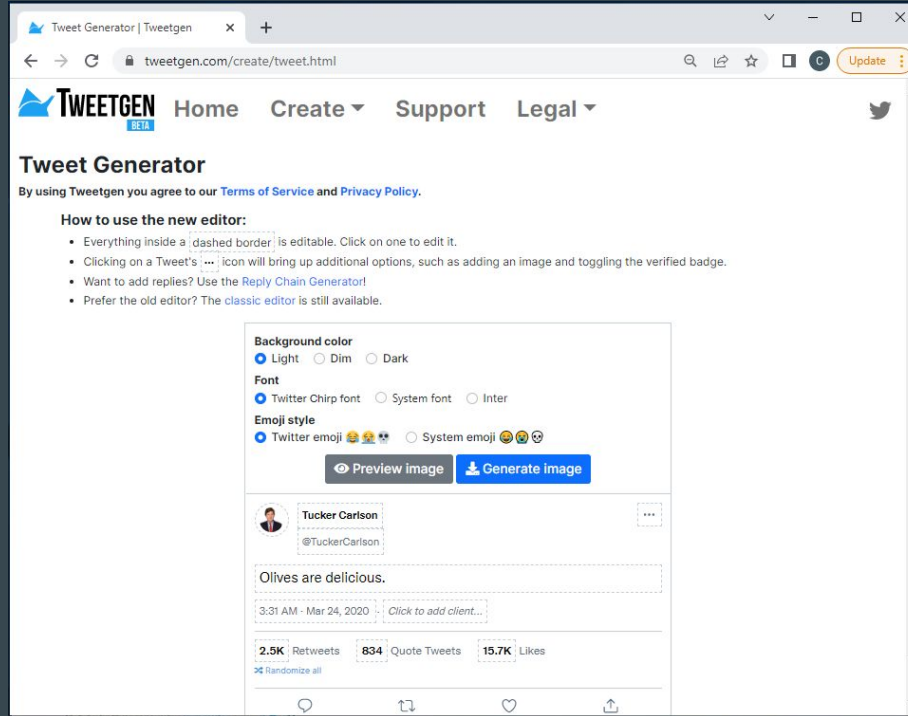
Goal: Validating attribution

https://www.snopes.com/fact-check/ted-cruz-texas-freezes-over/
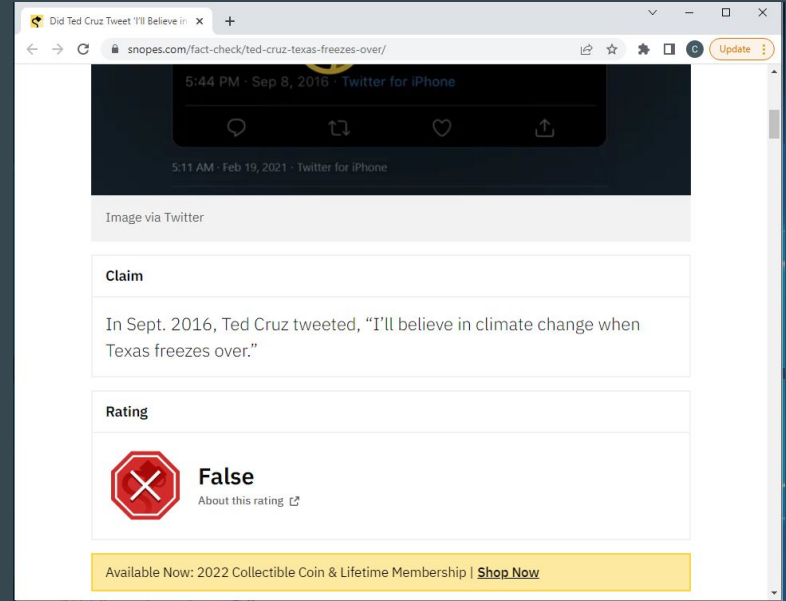
Given a screenshot of a tweet, we want to make sure that tweet was made by that person

Easily spread by credible people

A now [deleted tweet](deleted%20tweet) propagating the fake

Twitter screenshots are prevalent across all social media

Twitter: QTing would provide engagement

Instagram: Sharing posts across platforms

Caleb Bradford   @calebkbradford   ODU CS REU Final Presentation   2022-08-04

How can we verify attribution?

Twitter handle → **Haley Bragg** @haleybragg17

Tweet body → Disinformation on Instagram is literally not a problem.

Date and time → 11:25 AM · Jul 31, 2022

Engagement → **310** Retweets **214** Quote Tweets **4.6K** Likes

A lot of important metadata can be found in the screenshot itself

There is a good chance the answer in on the Web
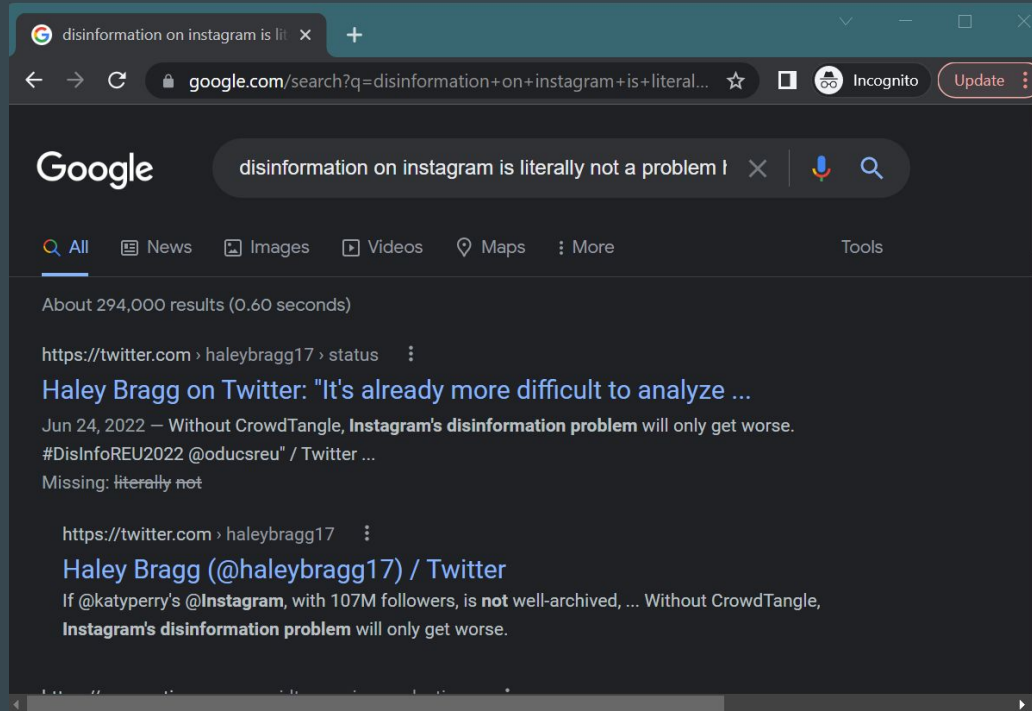
A common first instinct is to simply search Google for more information

Some tweets supposedly made by well-known accounts attract the attention of fact-checkers

Claim

U.S. President Donald Trump tweeted in November 2009 that "Obama's handling of this whole pandemic has been terrible!" and that he "would never let thousands of Americans die from a pandemic while in office."

Rating

False
About this rating ↗

False claim: In 2009 Trump tweeted that he "would never let thousands of Americans die from a pandemic"

By Reuters Staff   3 MIN READ   f t

Social media users are circulating a screenshot of what appears to be an old tweet from President Donald Trump (here , here , and here).

The image shows a tweet allegedly posted by Trump on November 23, 2009, during the swine flu outbreak. It reads: "Obama's handling of this whole pandemic has been terrible! As President, ALL responsibility becomes yours during a crisis like this, whether or not you're entirely to blame. John McCain, and for that matter myself, would never let thousands of Americans die from a pandemic while in office."

https://www.snopes.com/fact-check/2009-trump-tweet-pandemic/
Fact check from Snopes.com

https://www.reuters.com/article/idUSKCN2242AK
Fact check from Reuters.com

Caleb Bradford   @calebkbradford   ODU CS REU Final Presentation   2022-08-04

Using just the text of the alleged tweet, can we determine its veracity by querying search engines?

Caleb Bradford   @calebkbradford   ODU CS REU Final Presentation   2022-08-04

A Google query using just the body of a tweet
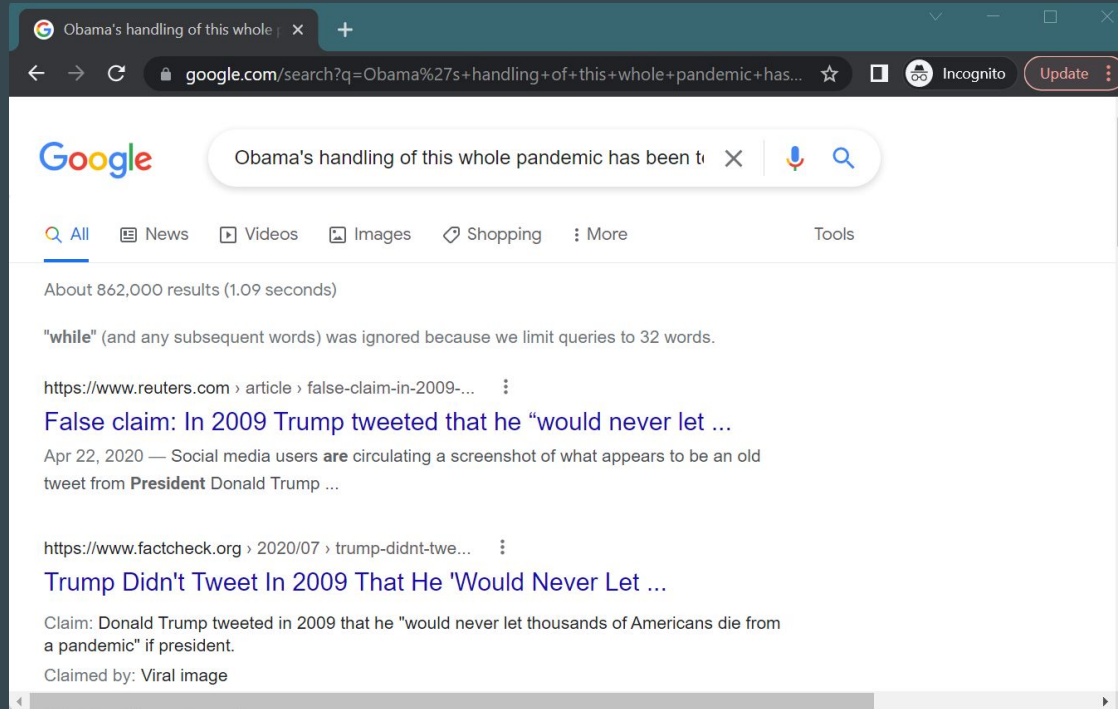
Live web version Archived version

Queries can also be made to sites with built-in search functions

Obama's handling of this whole pandemic has been terrible! As President, ALL responsibility becomes yours du

Search Results for "Obama's handling of this whole pandemic has been terrible! As President, ALL responsibility becomes yours during a crisis"

NEWS   COMPANIES   VIDEOS   PICTURES

1 Results   Sort by: Relevance ⌄   Date Range: Anytime   ⌄

False claim: In 2009 Trump tweeted that he "would never let thousands of Americans die from a pandemic"

... pandemic has been terrible! As President, ALL responsibility becomes... the swine flu outbreak.
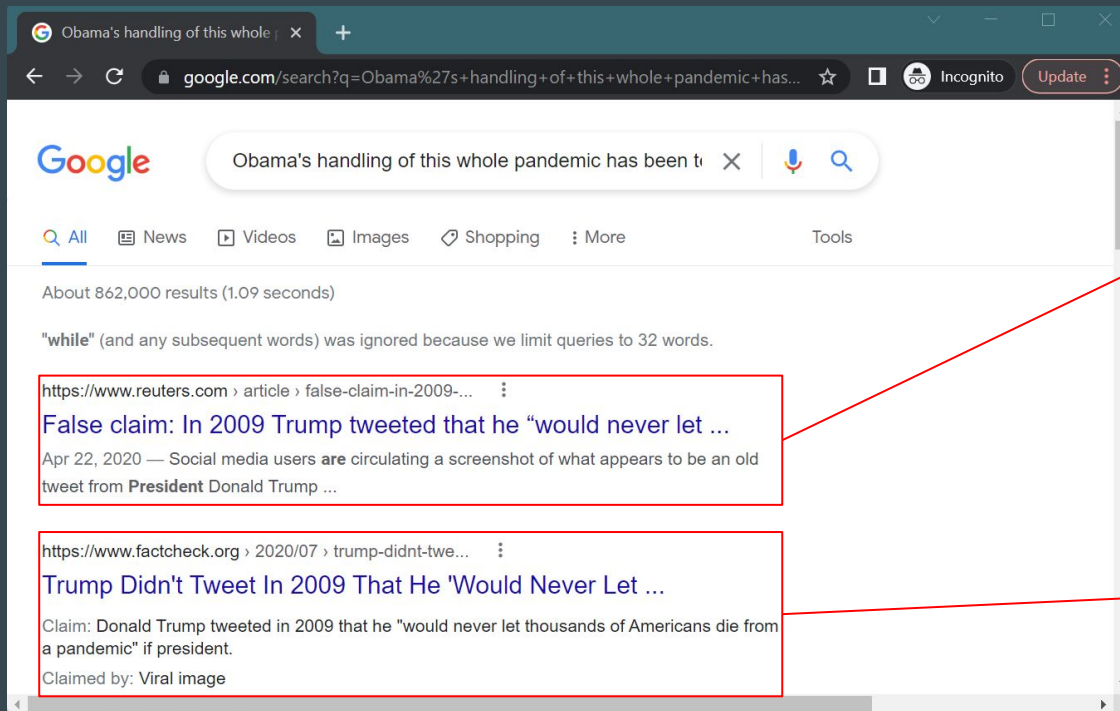It reads: "Obama's handling of this whole...
APRIL 22, 2020 10:51AM EDT

[Search query to Reuters](#)   [Archived version here](#)

Caleb Bradford   @calebkbradford   ODU CS REU Final Presentation   2022-08-04

Important metrics in the efficacy of a search query

Mean reciprocal rank = 1
Precision at 1 = 1

Mean reciprocal rank = 0.5
Precision at 1 = 0

Building a dataset of Snopes articles and search engine results

```json
"Article 1":{
    "URL": "https://www.snopes.com/fact-check/donald-trump-fake-tweet/",
    "Supposed author": "Donald Trump",
    "Tweet body": "My grandparents didn't come to America all the way from
    "Snopes query": {
        "Query URL": "https://www.snopes.com/?s=My+grandparents+didn%27t+co
        "Reciprocal rank": 0,
        "Precision at 1": 0
    },
    "Google query (site:snopes.com)":{
        "Query URL": "https://www.google.com/search?q=My+grandparents+didn%
        "Reciprocal rank": 1,
        "Precision at 1": 1
    },
    "Google query": {
        "Query URL": "https://www.google.com/search?q=My+grandparents+didn%
        "Reciprocal rank": 1,
        "Precision at 1": 1
    },
    "Reuters query": {
        "Query URL": "https://reuters.com/search/news?blob=My+grandparents+
        "Reciprocal rank": 0,
        "Precision at 1": 0
    }
}
```

[Full dataset available here](#)

Basic analysis of 30 data points found the following:
- **Snopes Search**
  - Mean reciprocal rank = 0.5500
  - Average P@1 = 0.5333
- **Google query**
  - Mean reciprocal rank = 0.4513
  - Average P@1 = 0.2667
- **Google query with site operator**
  - Mean reciprocal rank = 0.8667
  - Average P@1 = 0.8667
- **Search on Reuters for related article**
  - Mean reciprocal rank = 0.2167
  - Average P@1 = 0.2000

Caleb Bradford   @calebkbradford   ODU CS REU Final Presentation   2022-08-04

How can we automate the process of making queries given the text of an alleged tweet?

Web queries to fact-checking sites can easily be made programmatically

The same query to Reuters being made via Python script (tweet passed from input file CL argument)

Caleb Bradford   @calebkbradford   ODU CS REU Final Presentation   2022-08-04

HTML of articles can be scraped for their "truth rating"

Claim

U.S. President Donald Trump tweeted in November 2009 that "Obama's handling of this whole pandemic has been terrible!" and that he "would never let thousands of Americans die from a pandemic while in office."

Rating

False

About this rating ⧉

```
▼<div class="media position-relative"> flex
  ▼<div class="avatar-md avatar-lg-large mr-3"> == $0
    ▼<figure class="figure w-100 mb-0">
      <img class="figure-image img-responsive img-fluid w-100"
      width="100px" height="100px" src="https://www.snopes.com/
      tachyon/2018/03/rating-false.png" alt="False">
    </figure>
  </div>
```


Python script scrapes article for rating

Google does not appreciate non-humans making queries

A curl request to a Google query that returns an error page

Caleb Bradford   @calebkbradford   ODU CS REU Final Presentation   2022-08-04

Making Google queries is possible using software developed by Dr. Alexander Nwala

```
cs_cbrad022@ganymede:~$ scraper google "This is a google query"

googleSearchMain():

googleGetSERPResults() - start
        page: 1
        searchQuery: https://www.google.com/search?as_q=This%20is%20a%20google%20query
        derefURI(), sleep: 1

googleGetSERPResults() - end

printPayload(), src: google
1 https://support.google.com/docs/answer/3093343?hl=en
2 https://www.benlcollins.com/spreadsheets/google-sheets-query-sql/
3 https://www.youtube.com/watch?v=oCKvAcXTLZo
4 https://www.youtube.com/watch?v=oCKvAcXTLZo&t=0
5 https://www.youtube.com/watch?v=oCKvAcXTLZo&t=43
6 https://www.youtube.com/watch?v=oCKvAcXTLZo&t=117
7 https://www.youtube.com/watch?v=oCKvAcXTLZo&t=281
8 https://www.youtube.com/watch?v=oCKvAcXTLZo&t=507
9 https://www.youtube.com/watch?v=oCKvAcXTLZo&t=619
10 https://www.youtube.com/watch?v=oCKvAcXTLZo&t=666
11 https://www.youtube.com/watch?v=oCKvAcXTLZo&t=692
12 https://www.youtube.com/watch?v=bW6P2YvLyZg
13 https://www.youtube.com/watch?v=W8C9TcYt47M
14 https://codingisforlosers.com/google-sheets-query-function/
15 https://codingisforlosers.com/google-sheets-query-function/#1_The_basics_of_Google_Sheets_Queries
16 https://codingisforlosers.com/google-sheets-query-function/#6_Building_Pivot_Tables_with_Query
17 https://www.howtogeek.com/450465/how-to-use-the-query-function-in-google-sheets/
18 https://spreadsheetpoint.com/google-sheets-query-function/
19 https://developers.google.com/chart/interactive/docs/querylanguage
20 https://sheetsformarketers.com/google-sheets-formulas/query/
21 https://coefficient.io/how-to-google-sheets-query-function/
22 https://cloud.google.com/bigquery/docs/samples/bigquery-query-external-sheets-perm
```

Dr. Nwala's "scraper" tool in action

Using Nwala's work, broad searches for evidence becomes easier

Google script can find variety of fact-check articles

Caleb Bradford   @calebkbradford   ODU CS REU Final Presentation   2022-08-04

Future work

- Development of a more generic method of parsing fact-check articles for "truth rating"
    - Scraping raw HTML works, but needs to be specifically developed for every site we want to work with
    - The structure of a site is subject to constant change
- Building a more comprehensive dataset of known tweets
    - More tweets and their related queries
- Are there more effective ways to query the web for evidence?

Caleb Bradford   @calebkbradford   ODU CS REU Final Presentation   2022-08-04

Conclusions

- Tweet misattribution is a unique form of disinformation that is extremely prevalent on all social media
- Queries built using just the body of a well covered tweet have shown efficacy at finding them
- Tools developed and further work available here
  - https://github.com/calebkbrad/Attribution-Scraping
- Acknowledgements
  - Dr Michael Nelson
  - Emily Escamilla
  - Tarannum Zaki
  - Dr Alexander Nwala
  - My fellow cohort members

Caleb Bradford   @calebkbradford   ODU CS REU Final Presentation   2022-08-04