MICHIGAN STATE UNIVERSITY

Beta Presentation

Automotive Specific Dark Web Threat Intelligence

The Capstone Experience

Team GM

Anthony Ajlouny
Mitch Neelis
Kevin Kimmel
Ben Kelenski
Yuxin Li

Department of Computer Science and Engineering
Michigan State University

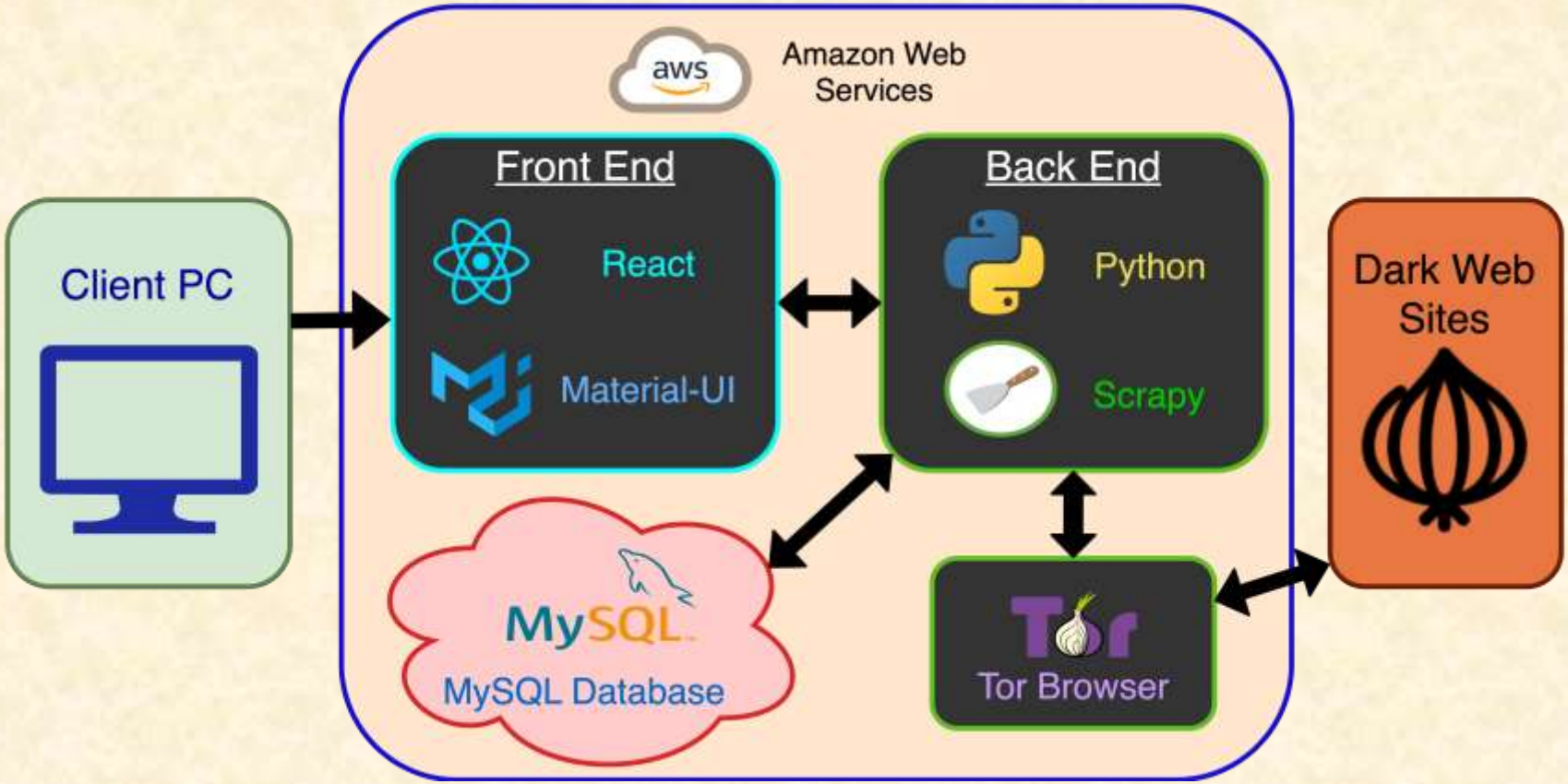
Fall 2020

From Students…
…to Professionals

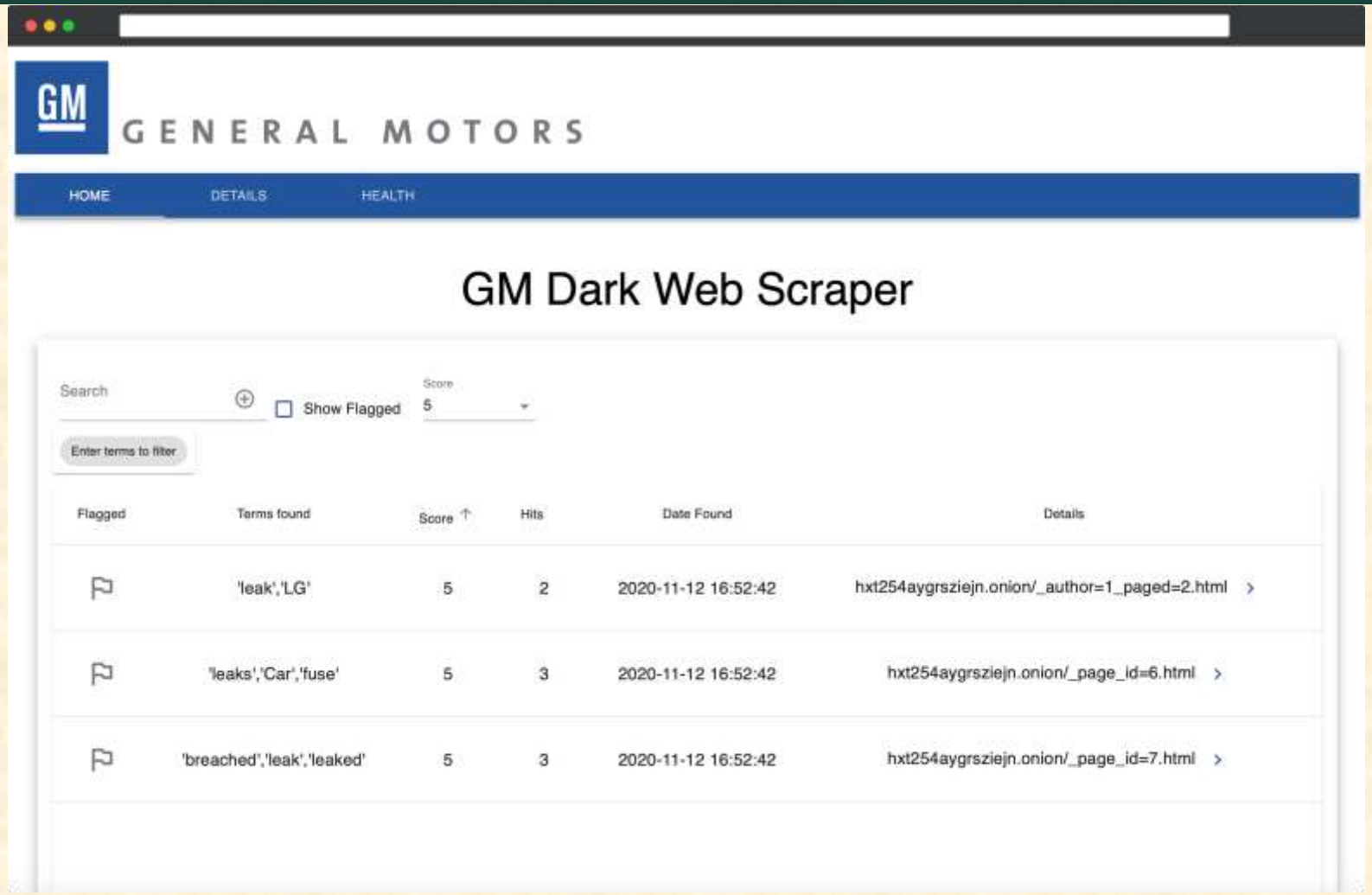
Project Overview

- The program scrapes through the dark web looking for defined key words
- Text found on the websites are scraped and saved in the database, including possible images that may exist
- The contents run through our scoring system to check for severity of information
- This info is pushed to the database, where can be viewed and sorted in the web app

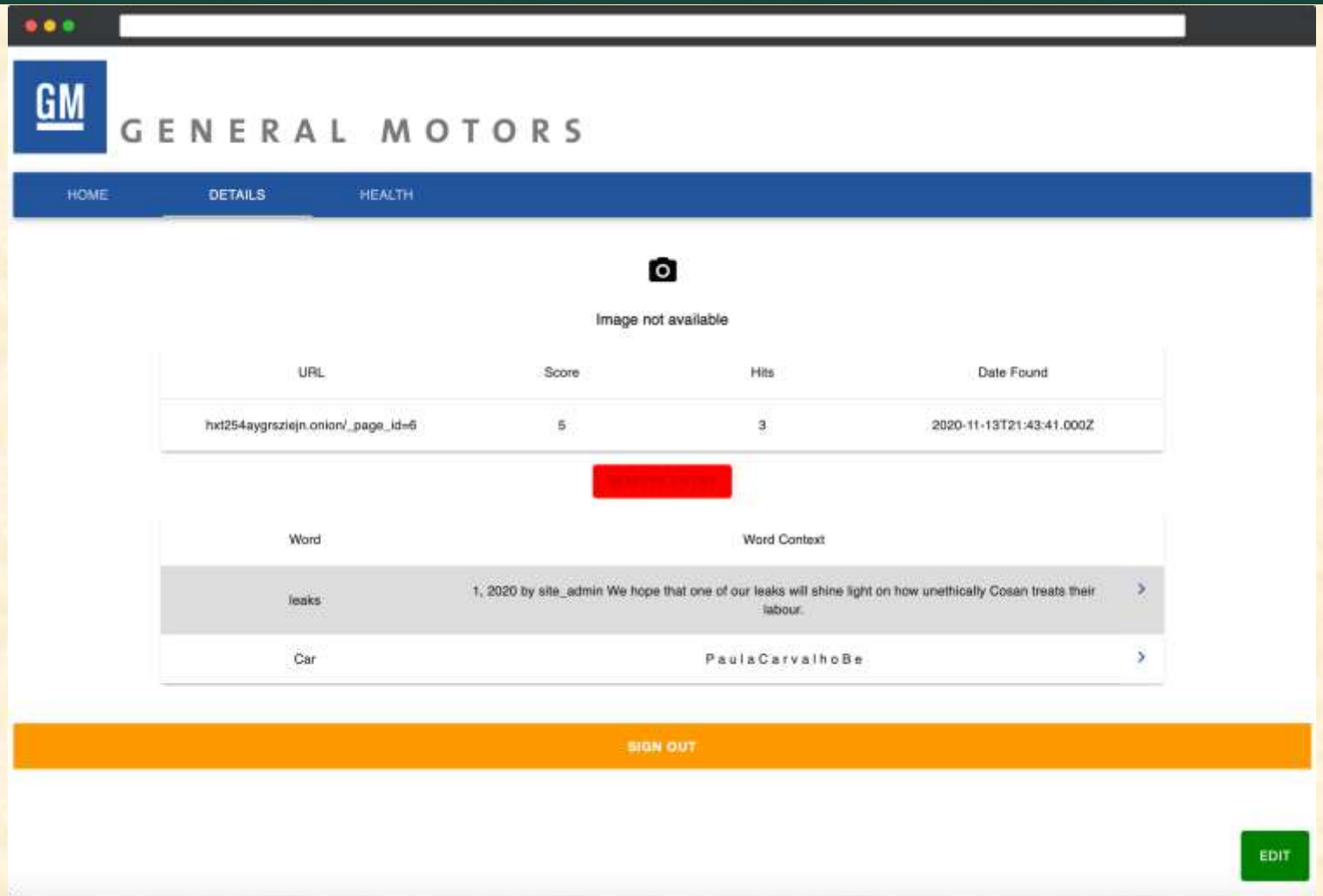
System Architecture



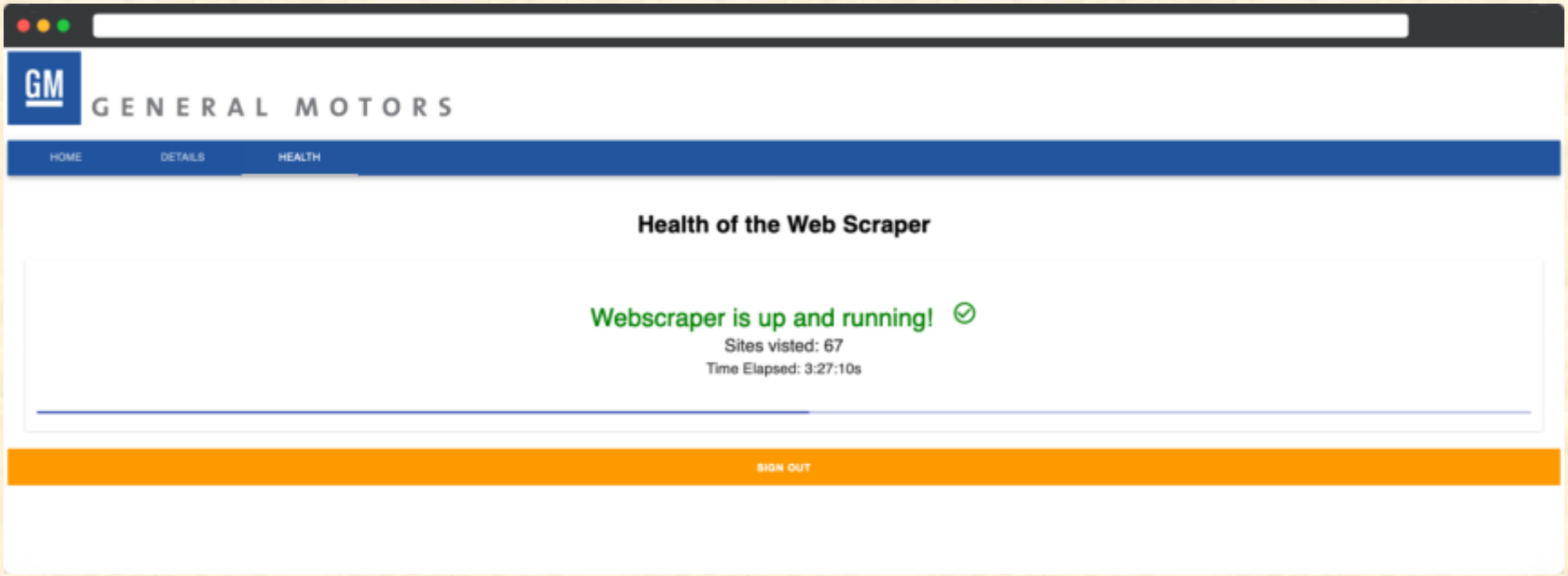
Scraper Main Page



Details Page



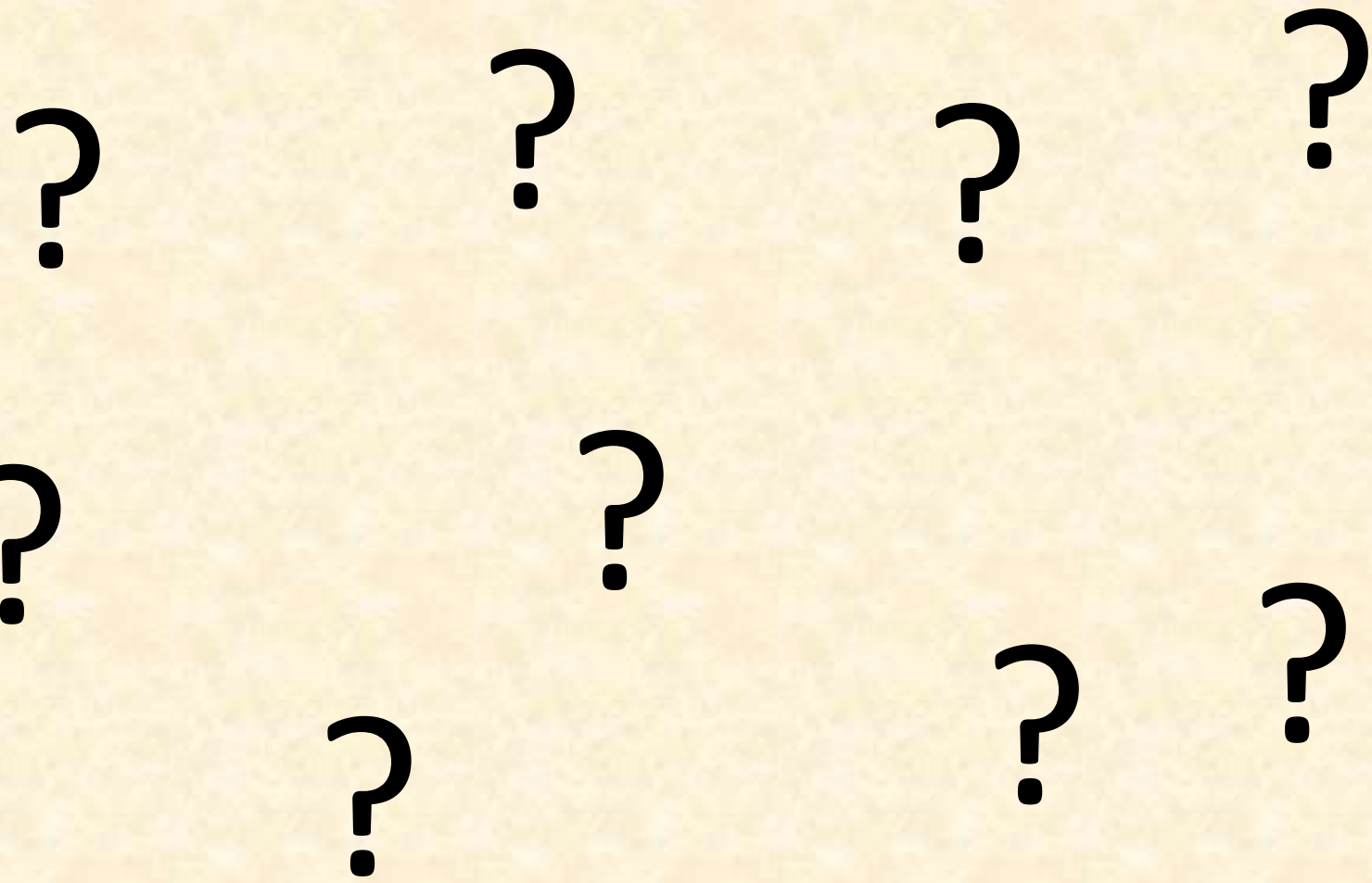
Scraper Health Page



Scrape Data Table

	ID	Score	URL	Terms found	Date Found	Favorite
0	73	3	hxt254aygrsziejn.onion/_page_id=6	{'term_1': {'term': 'breached', 'context': 'co...	2020-11-15 15:59:45.047723	False
1	12	2	hxt254aygrsziejn.onion/_p=382	{'term_1': {'term': 'breached', 'context': 'in...	2020-11-15 15:59:45.047723	False
2	1	2	hxt254aygrsziejn.onion/_p=303	{'term_1': {'term': 'breached', 'context': 'in...	2020-11-15 15:59:45.047723	False
3	45	2	hxt254aygrsziejn.onion/_p=205	{'term_1': {'term': 'breached', 'context': 'in...	2020-11-15 15:59:45.047723	False
4	4	2	hxt254aygrsziejn.onion/_cat=9	{'term_1': {'term': 'breached', 'context': '&r...	2020-11-15 15:59:45.047723	False
5	53	2	hxt254aygrsziejn.onion/_p=166	{'term_1': {'term': 'breached', 'context': 'Of...	2020-11-15 15:59:45.047723	False
6	109	2	hxt254aygrsziejn.onion/_cat=7	{'term_1': {'term': 'leak', 'context': 'July 5...	2020-11-15 15:59:45.047723	False
7	55	2	hxt254aygrsziejn.onion/_p=427	{'term_1': {'term': 'breach', 'context': 'to c...	2020-11-15 15:59:45.047723	False
8	65	2	hxt254aygrsziejn.onion/_p=136	{'term_1': {'term': 'leak', 'context': 'April ...	2020-11-15 15:59:45.047723	False
9	75	1	hxt254aygrsziejn.onion/_author=1	{'term_1': {'term': 'leak', 'context': '5.Post...	2020-11-15 15:59:45.047723	False
10	70	1	hxt254aygrsziejn.onion/_p=346	{'term_1': {'term': 'leak', 'context': '2, 202...	2020-11-15 15:59:45.047723	False
11	69	1	hxt254aygrsziejn.onion/_cat=26	{'term_1': {'term': 'leak', 'context': '0Orang...	2020-11-15 15:59:45.047723	False
12	67	1	hxt254aygrsziejn.onion/_p=142	{'term_1': {'term': 'breached', 'context': 'no...	2020-11-15 15:59:45.047723	False
13	63	1	hxt254aygrsziejn.onion/_p=392	{'term_1': {'term': 'leak', 'context': 'July 1...	2020-11-15 15:59:45.047723	False
14	0	1	hxt254aygrsziejn.onion/_p=362	{'term_1': {'term': 'leak', 'context': 'the th...	2020-11-15 15:59:45.047723	False
15	59	1	hxt254aygrsziejn.onion/_cat=5	{'term_1': {'term': 'leak', 'context': '5.Post...	2020-11-15 15:59:45.047723	False
16	78	1	hxt254aygrsziejn.onion/_cat=10	{'term_1': {'term': 'leak', 'context': 'August...	2020-11-15 15:59:45.047723	False
17	58	1	hxt254aygrsziejn.onion/_cat=11	{'term_1': {'term': 'leak', 'context': 'June 2...	2020-11-15 15:59:45.047723	False
18	57	1	hxt254aygrsziejn.onion/_p=202	{'term_1': {'term': 'leak', 'context': 'Symbol...	2020-11-15 15:59:45.047723	False
19	50	1	hxt254aygrsziejn.onion/_author=1_paged=2	{'term_1': {'term': 'leak', 'context': '9, 202...	2020-11-15 15:59:45.047723	False

What’s left to do?

- Debug Scraper Health Page
- Allow user to manually run the scraper
- Display possible images found for wider security coverage
- Neatly display full text file when prompted
- Debug parsing algorithm for URL's and terms
- Improve security on the web app

Questions?