MICHIGAN STATE UNIVERSITY

Alpha 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

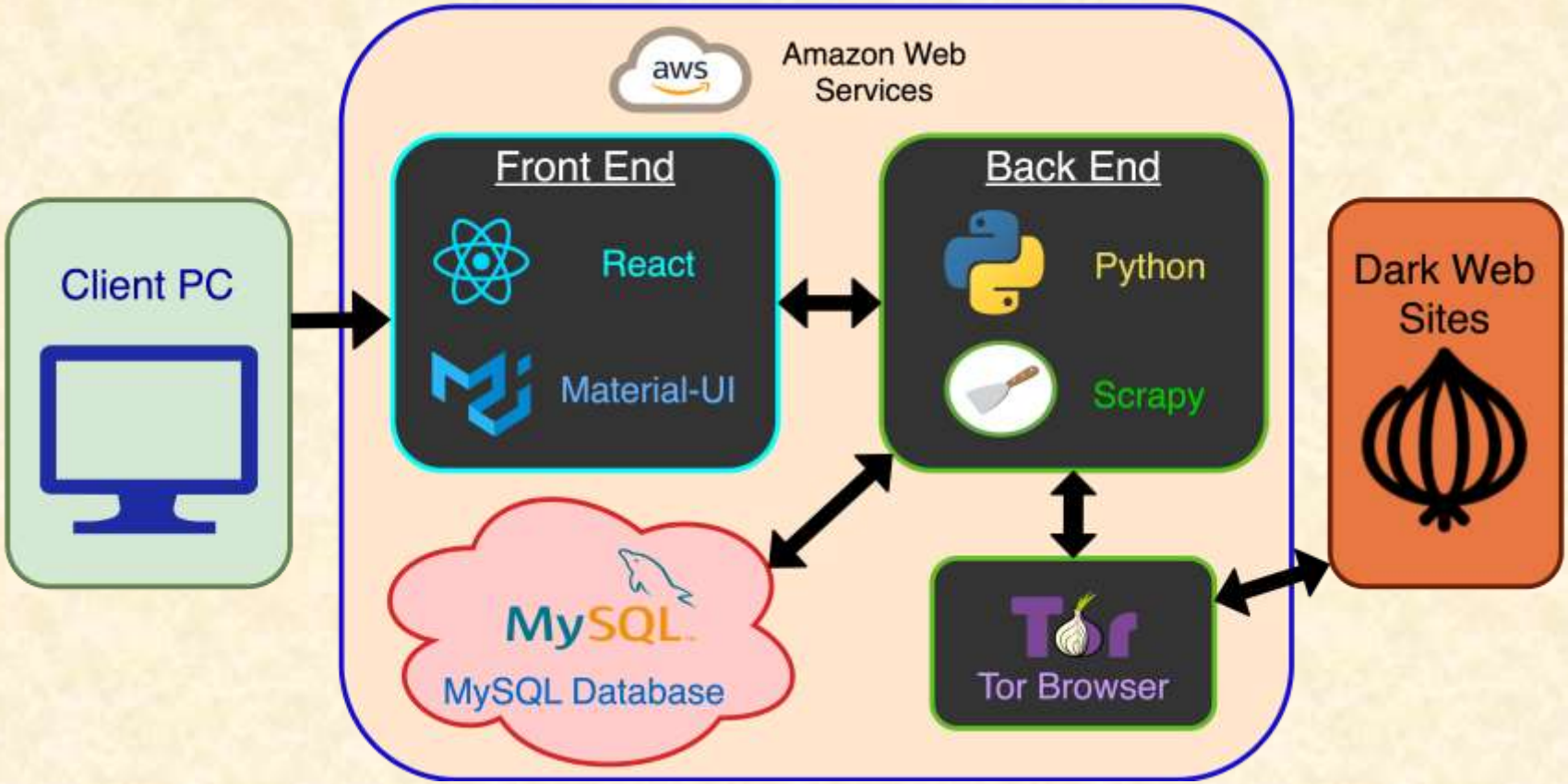
Fall 2020

From Students…
…to Professionals

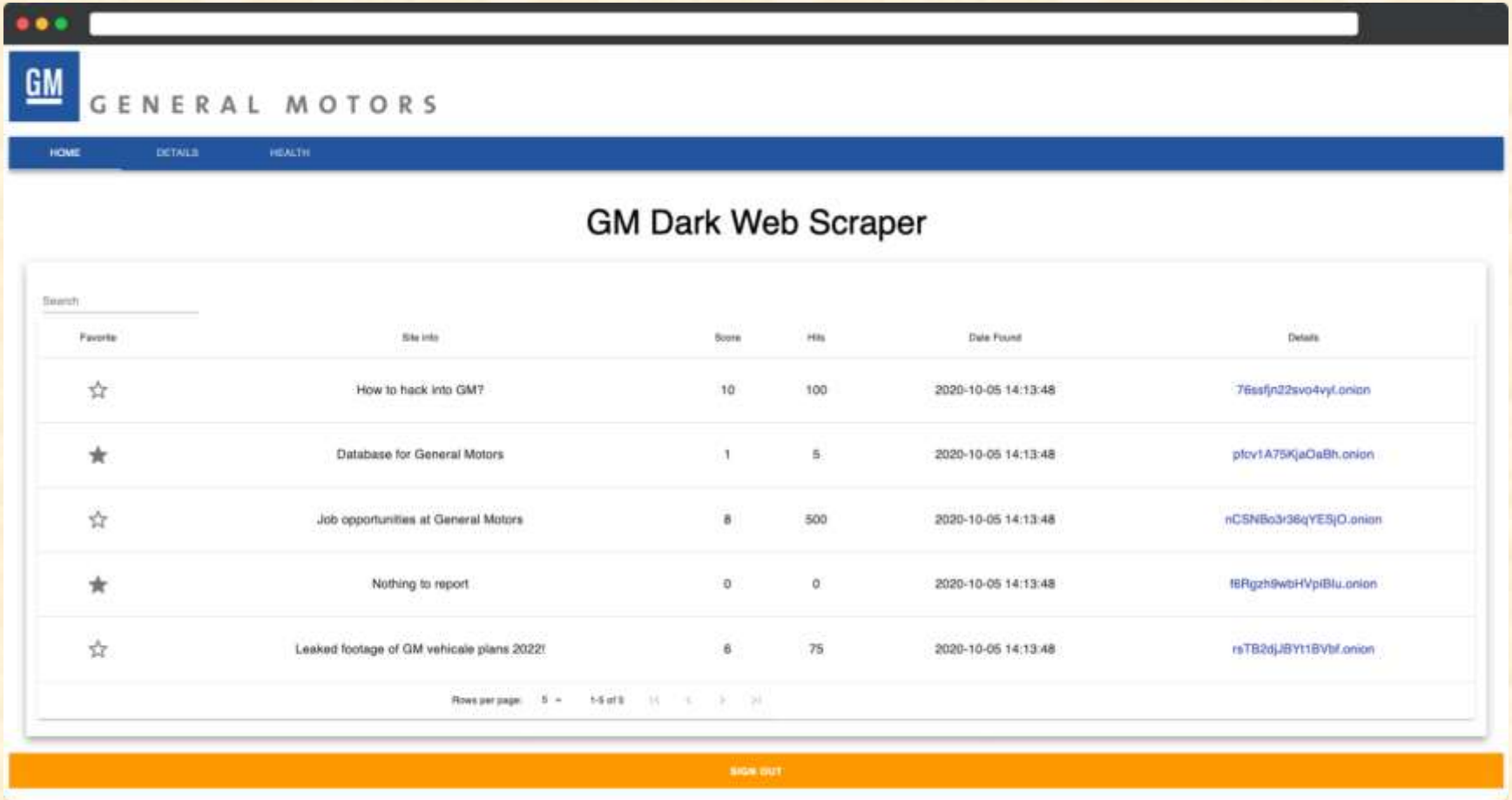
Project Overview

- The program scrapes through the dark web looking for defined key words
- If found, the program takes a screenshot and saves page contents
- 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 edited in the web app

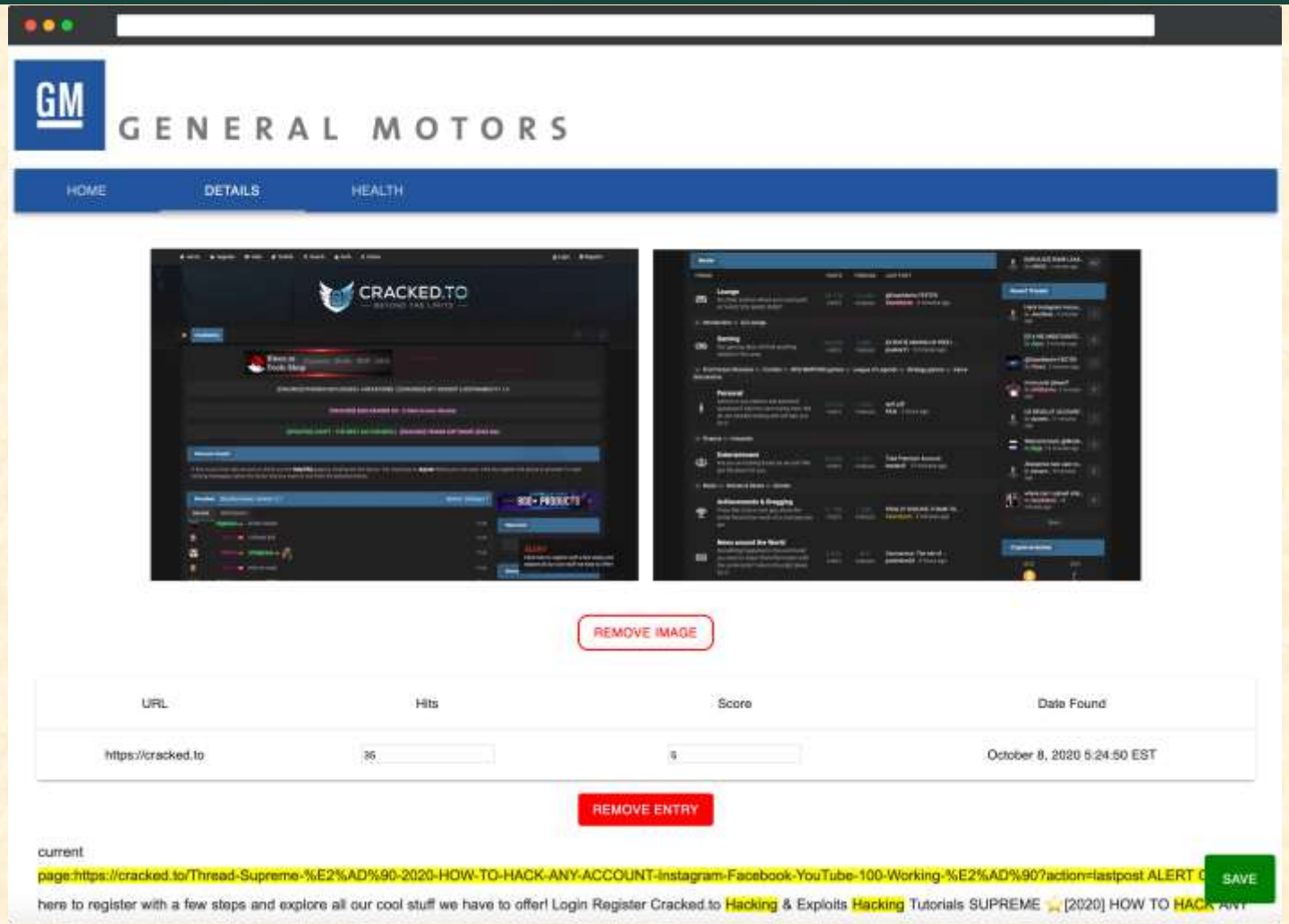
System Architecture



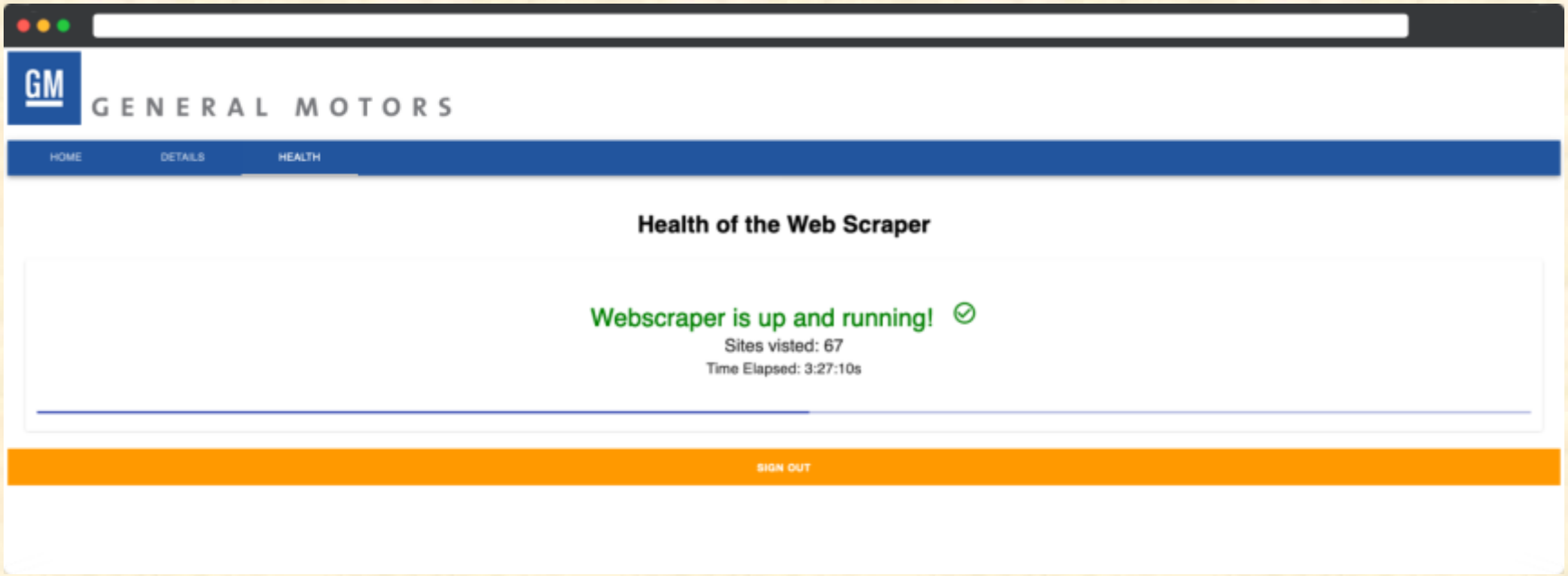
Main Page


GM
GENERAL MOTORS
HOME
DETAILS
HEALTH
GM Dark Web Scraper
Search
Favorite
Site Info
Score
Hits
Date Found
Details
How to hack into GM?
10
100
2020-10-05 14:13:48
76ssfjn22svo4vyf.onion
Database for General Motors
1
5
2020-10-05 14:13:48
ptcv1A75KjaOaBh.onion
Job opportunities at General Motors
8
500
2020-10-05 14:13:48
nCSNBo3r36qYE5jO.onion
Nothing to report
0
0
2020-10-05 14:13:48
f8Rgzh9wbHVpiBlu.onion
Leaked footage of GM vehicale plans 2022!
6
75
2020-10-05 14:13:48
rsTB2djJBY11BVbf.onion
Rows per page: 5
1-5 of 5
SIGN OUT

Details Page



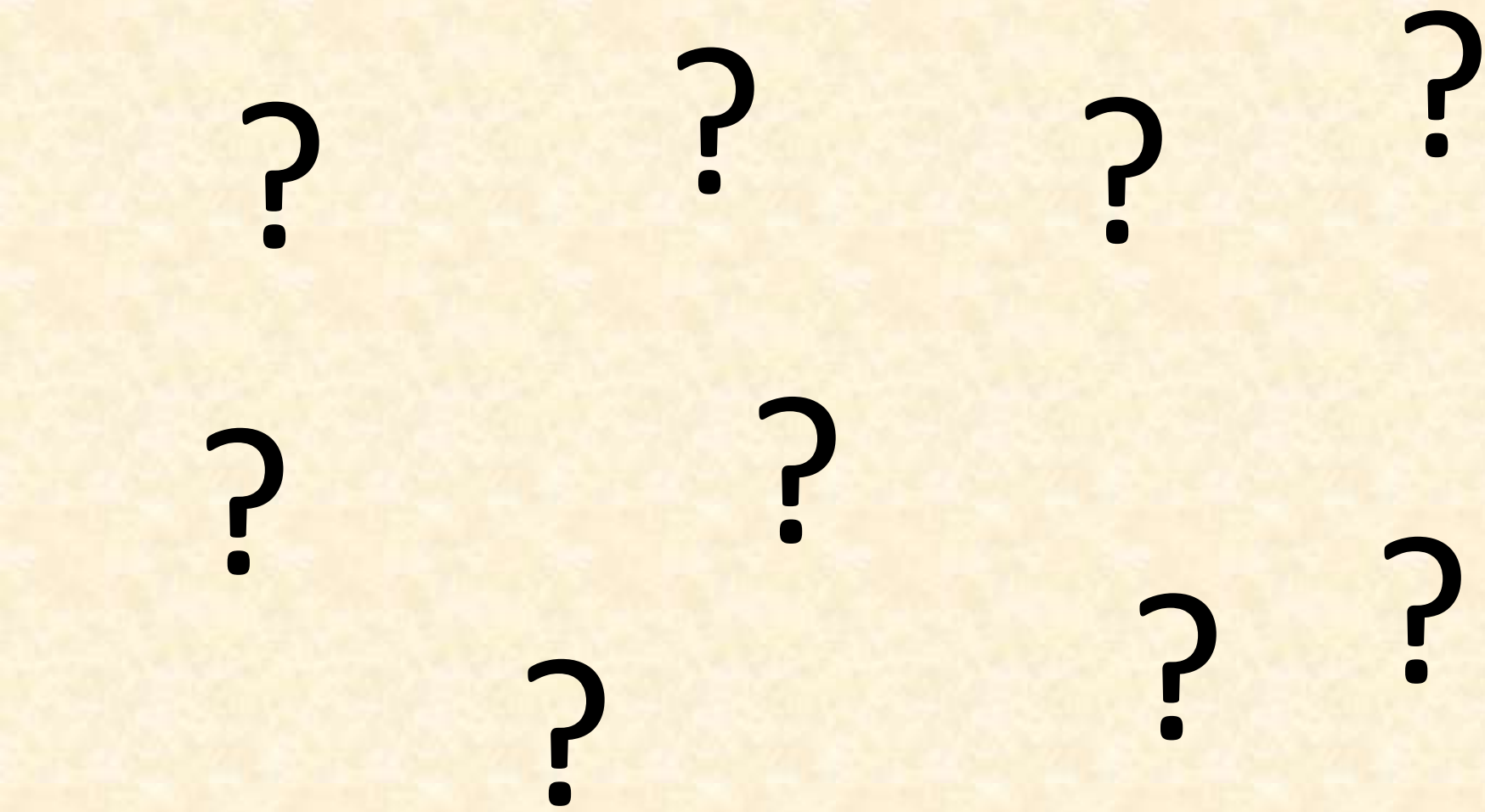
Scraper Health



Collected Data Table

	ID	Score	URL	Terms found	Context	Date Found	Favorite
0	73	3	hxt254aygrszieJn.onion/_page_id=6	[breached, leak, leaked]	[from corporate networks that were breached an...	2020-10-18 13:20:05.186433	False
1	12	2	hxt254aygrszieJn.onion/_p=382	[breached, leak]	[Kong and Colombo. MAS holding was breached...	2020-10-18 13:20:05.186433	False
2	1	2	hxt254aygrszieJn.onion/_p=303	[breached, leak]	[Kong and Colombo. MAS holding was breached...	2020-10-18 13:20:05.186433	False
3	45	2	hxt254aygrszieJn.onion/_p=205	[breached, leak]	[Kong and Colombo. MAS holding was breached...	2020-10-18 13:20:05.186433	False
4	4	2	hxt254aygrszieJn.onion/_cat=9	[breached, leak]	[Part 1-2. MAS holding was breached and 300 gb...	2020-10-18 13:20:05.186433	False
5	53	2	hxt254aygrszieJn.onion/_p=166	[breached, leak]	[Part 1-2. MAS holding was breached and 300 gb...	2020-10-18 13:20:05.186433	False
6	109	2	hxt254aygrszieJn.onion/_cat=7	[leak, leaks]	[is the seventh part of leak on COSAN.filelist...	2020-10-18 13:20:05.186433	False
7	55	2	hxt254aygrszieJn.onion/_p=427	[breach, leak]	[They allowed us to easily breach their networ...	2020-10-18 13:20:05.186433	False
8	65	2	hxt254aygrszieJn.onion/_p=136	[leak, leaks]	[is the third part of leak on COSAN. <a hre...	2020-10-18 13:20:05.186433	False
9	75	1	hxt254aygrszieJn.onion/_author=1	[leak]	[--></header>\n\n\t\n\t\t\t<div class="entry-s...	2020-10-18 13:20:05.186433	False
10	70	1	hxt254aygrszieJn.onion/_p=346	[leak]	[the third part of the leak on Aliansce Sonae....	2020-10-18 13:20:05.186433	False
11	69	1	hxt254aygrszieJn.onion/_cat=26	[leak]	[.entry-meta --> \n\t\t\t<!-- .entry-head...	2020-10-18 13:20:05.186433	False
12	67	1	hxt254aygrszieJn.onion/_p=142	[breached]	[Part 1-2. MAS holding was breached and 300 gb...	2020-10-18 13:20:05.186433	False
13	63	1	hxt254aygrszieJn.onion/_p=392	[leak]	[is a teaser for the leak on Indoco Remedies L...	2020-10-18 13:20:05.186433	False
14	0	1	hxt254aygrszieJn.onion/_p=362	[leak]	[leak. Second part of the leak includes compli...	2020-10-18 13:20:05.186433	False

What’s left to do?

- Implement web crawling code in AWS and connect it with the S3 bucket
- Format the incoming data to be neatly displayed on main page and details page
- Display stats on health screen
- Filtering by "favorite" option on web app

Questions?