



Games Behavior

CS 4624 Multimedia, Hypertext, and Information Access

Instructor: Dr. Edward A. Fox

Virginia Tech, Blacksburg, VA 24061

Authors: Daniel Lanigan, Renzo Neps, Liam McBride

May 9, 2024



Outline

1. Project Overview
2. Timeline
3. Deliverables
4. YouTube Scraper
5. Forum Scraper
6. Data
7. Visualizations
8. Future Work
9. Acknowledgements
10. References



1. Project Overview

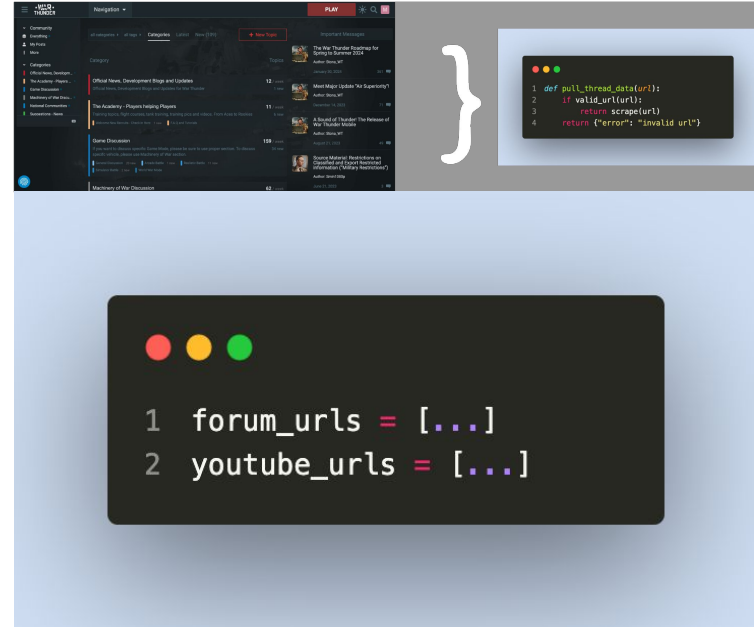
- Client
 - Dr. James D. Ivory
 - Virginia Tech Gaming and Media Effects Lab
- Problem: Unknown data on extremist behavior in online game communities
- Deliverables
 - Forum scraper
 - Cleaned data and analysis
 - Visualizations



Dr. James D. Ivory

2. Timeline: February

- Research scraping
- Develop blog scraper
- Understand/Modify previous scraper
- Determine media to scrape
- Test and begin scraping
- Report #1



February

March

April

May



Timeline: March

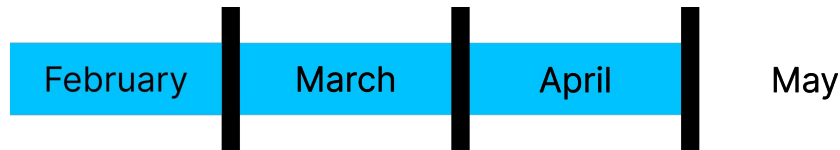
- Presentation #2
- Finish scraping
- Create storage of clean dataset
- Begin cleaning data
- Determine what we want to visualize





Timeline: April

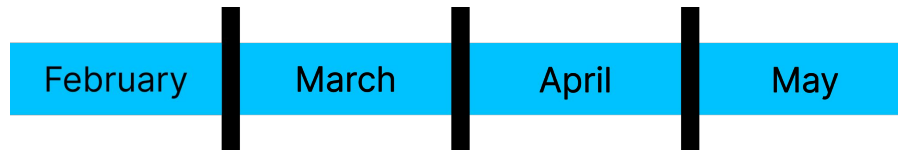
- Finish cleaning data
- Sentiment analysis with client
- Finalize visualization
- Begin/finish final report
- Final presentation





Timeline: May

- Report #2
- VTechWorks Submission





3. Deliverables

- YouTube Scraper Script
- Forum Scraper Script
- Raw Data
- Cleaned Data
- Visualizations

4. YouTube Scraper Script

```
function YoutubeScraper(iterations, search_query, activeSheet, offset)
{
  var next = null
  const start_date = formatDateString(new Date('January 1, 2013 00:00:00 -0500'));
  const end_date = formatDateString(new Date());
  for (let i = 0; i < iterations; i++) {
    if (next == null) {
      var search = YouTube.Search.list("snippet, id", {q: search_query, maxResults: 5, relevanceLanguage: "en", regionCode: "US", startDate: start_date, endDate: end_date})
    }
    else {
      var search = YouTube.Search.list("snippet, id", {q: search_query, maxResults: 5, pageToken: next, relevanceLanguage: "en", regionCode: "US", startDate: start_date, endDate: end_date})
    }
    var results = search.items.map((item) => [item.id.videoId, item.snippet.title, item.snippet.publishedAt])
    next = search.nextPageToken

    var ids = results.map((id) => id[0]).join(",")

    var stats = YouTube.Videos.list("statistics", {id: ids})

    var videoStats = stats.items.map((item) => [item.statistics.viewCount, item.statistics.likeCount, item.statistics.commentCount])

    activeSheet.getRange(1 + offset, 1, results.length, results[0].length).setValues(results)
    activeSheet.getRange(1 + offset, 4, videoStats.length, videoStats[0].length).setValues(videoStats)

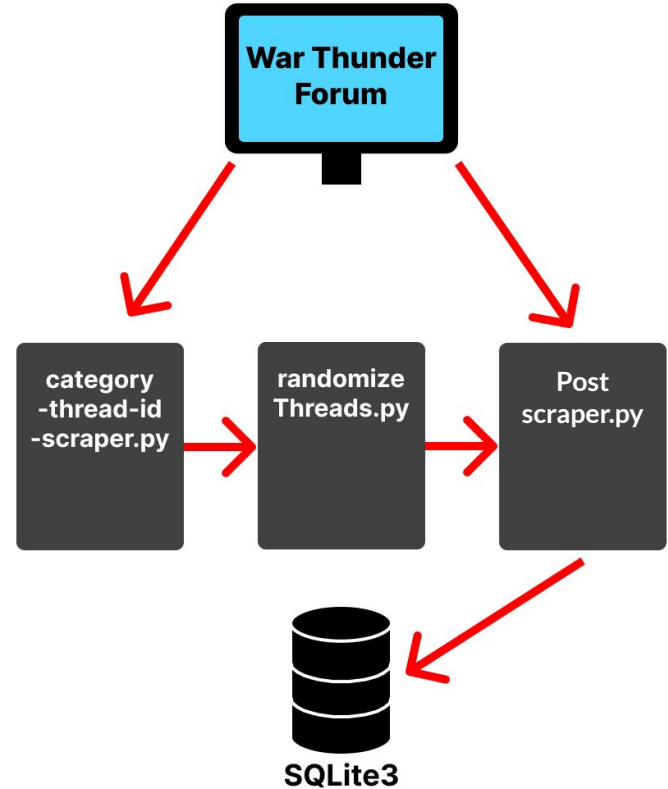
    /*var lVideoID = results.item.id.videoId.pop();
    var lTitle = results.item.snippet.title.pop();
    var lPublishedAt = results.item.snippet.publishedAt.pop();
    console.log(lVideoID + lTitle + lPublishedAt);
    //console.log(setValues(videoStats))*/
  }
}
```

```
function scrapeComments(x, y)
{
  var ss = SpreadsheetApp.getActiveSpreadsheet();
  var result=[['Name', 'Comment', 'Time', 'Likes', 'Reply Count', 'Reply Author', 'Reply', 'Published', 'Updated']];
  var vid = ss.getSheets()[0].getRange(x,y).getValue();
  var nextPageToken=undefined;

  for(let i = 0; i < 10; i++)
  {
    var data = YouTube.CommentThreads.list('snippet', {videoId: vid, maxResults: 100, pageToken: nextPageToken})
    nextPageToken=data.nextPageToken
    for (var row = 0; row < data.items.length; row++)
    {
      result.push([data.items[row].snippet.topLevelComment.snippet.authorDisplayName,
        data.items[row].snippet.topLevelComment.snippet.textDisplay,
        data.items[row].snippet.topLevelComment.snippet.publishedAt,
        data.items[row].snippet.topLevelComment.snippet.likeCount,
        data.items[row].snippet.topLevelComment.snippet.totalReplyCount, ',', ',', ',', '']);
      if(data.items[row].snippet.topLevelComment.snippet.totalReplyCount>0)
      {
        parent=data.items[row].snippet.topLevelComment.id
        var nextPageTokenRep=undefined
      }
    }
  }
}
```

5. Forum Scraper Script

- Files
 - category-thread-id-scraper.py
 - post-scraper.py
 - randomizeThreads.py



5. Forum Scraper Script cont.

```
liammcbride@Liams-MacBook-Air GamesBehavior24 % python category-thread-id-scraper.py
1
Latest Uncategorized topics - War Thunder - official forum
7
Latest Official News, Development Blogs and Updates topics - War Thunder - official forum
8
Latest National Communities topics - War Thunder - official forum
10
Latest The Academy - Players helping Players topics - War Thunder - official forum
11
Latest Welcome New Recruits - Check in Here topics - War Thunder - official forum
12
Latest F.A.Q and Tutorials topics - War Thunder - official forum
13
Latest Game Discussion topics - War Thunder - official forum
14
Latest General Discussion topics - War Thunder - official forum
15
Latest Arcade Battle topics - War Thunder - official forum
16
Latest Realistic Battle topics - War Thunder - official forum
17
Latest Simulator Battle topics - War Thunder - official forum
18
Latest World War Mode topics - War Thunder - official forum
19
Latest Machinery of War Discussion topics - War Thunder - official forum
```

Getting Categories

```
Latest General Discussion topics - War Thunder - official forum
Adding threads [0 - 30] 0
Adding threads [0 - 30] 30
Thread[420] has been recorded
Thread[100952] has been recorded
Thread[98214] has been recorded
Thread[99177] has been recorded
```

Getting Threads

```
liammcbride@Liams-MacBook-Air GamesBehavior24 % python randomizeThreads.py
Randomizing 100 threads to be scraped
From 100 threads there are 1149 total posts
liammcbride@Liams-MacBook-Air GamesBehavior24 % █
```

Randomizing Threads

```
liammcbride@Liams-MacBook-Air GamesBehavior24 % python post-scraper.py
Scraping thread: 1352
Running scrape on posts [0 to 30]
Running scrape on posts [7 to 37]
Scraping thread: 1780
Running scrape on posts [0 to 30]
Running scrape on posts [13 to 43]
```

Scraping Posts

6. Raw Data

- Forum Data
 - SQLite3 database
- YouTube Data
 - CSV file
 - SQLite3 database

```
▼ Youtube
  > capstone_sheets_War_Thunder
  > capstone_sheets_War_Thunder_Advice
  > capstone_sheets_War_Thunder_Gameplay
  > capstone_sheets_War_Thunder_Renzo
  ▼ capstone_sheets_War_Thunder_Tankenstein
    📄 _2gPTot9OXE.csv
    📄 _7nEzjSf-ng.csv
    📄 _v6ZJ9UuINA.csv
    📄 _y47ja0ITnl.csv
    📄 _YzpXfrLU90.csv
    📄 --Rj0gc2RI4.csv
    📄 2LTVuh2qG08.csv
    📄 3Mf38leHSZw.csv
    📄 4-BQW2pO5-Q.csv
    📄 4oCiEtzG9VA.csv
```

	name	comment	timestamp	likes	replies	video_id
	Filter	Filter	Filter	Filter	Filter	Filter
1	@cylerner	The intro to this ...	2024-04-05T04:5...	0	0	6hphDYXIs34
2	@SlipKnot7866	I NEED TO KNOW...	2024-04-03T02:4...	0	1	6hphDYXIs34
3	@littlesneez9002	Aurora XV 54	2024-04-04T03:4...	-1	-1	6hphDYXIs34
4	@sammy2021	been rewatching ...	2024-03-30T19:0...	0	2	6hphDYXIs34
5	@sammy2021	@@user-...	2024-04-03T20:2...	-1	-1	6hphDYXIs34
6	@user-gu8qi4me8x	How are you doin...	2024-04-03T20:0...	-1	-1	6hphDYXIs34
7	@IcyMan143	Goosebumps	2024-03-24T05:5...	0	0	6hphDYXIs34
8	@jermany8019	<a href="https://...	2024-03-24T05:5...	1	1	6hphDYXIs34
9	@littlesneez9002	<<Fire away ...	2024-04-04T03:4...	-1	-1	6hphDYXIs34

6. Cleaned Data

- CSV files
 - YouTube
 - Forum

```
CleanedYouTubeComments.csv
1 6hphDYXls34,The intro to this trailer really reminded me of the OG cod games like Bla
2 6hphDYXls34,I NEED TO KNOW THE MUSIC &gt;:O
3 6hphDYXls34,Aurora XV 54
4 6hphDYXls34,been rewatching these trailers with the NVIDIA ai HDR thing and OMFG
5 6hphDYXls34,@user-gu8qi4me8x If you have an NVIDIA GPU and the Latest drivers its on
6 6hphDYXls34,How are you doing this?!
7 6hphDYXls34,Goosebumps
8 6hphDYXls34,"<a href="https://www.youtube.com/watch?v=6hphDYXls34&t=105">1:45</
9 6hphDYXls34,&lt;&lt;Fire away Coward!&gt;&gt;
10 6hphDYXls34,And than there is me raging at 2.3 thinking ill never get back to top ti
11 6hphDYXls34,@-sixy- I have to specify it a bit better i meant back to top tier. I u
12 6hphDYXls34,youll get there someday
13 6hphDYXls34,I managed to perfectly sync the music with the free bird riff and it perf
14 6hphDYXls34,Parece una pelcula
15 6hphDYXls34,No me canso de ver estos trilers ..cada vez mas perfectos la msica gr
16 6hphDYXls34,BGM name?
17 6hphDYXls34,Still the best trailer I keep coming back too!
```

```
CleanedForumComments.csv
1 25630,"I will not get tired until everyone realizes how current system is buggy and make
2 25630,"Or play a different vehicle while you wait for your 1st game to end if u believe
3 25630,"Yeah, game wants us to grind but also punishes us for doing exactly that. But som
4 25630,very useful bug to be aware of. honestly I had felt I was losing RP at times but r
5 25630,"If there was a RP gain reduction like ten percent for all vehicles, there would b
6 25630,It will be added after your next battle is complete. People always think this is
7 25630,Its not a bug. It hust gets added to the next completed battles module research
8 25630,"haha, tell me you didnt read without saying you didnt read. " ,,,,,,,,,,
9 25630,I did read. And my point remains. You dont lose rp ,,,,,,,,,,,,,,,,,,,,,,
10 25630,Your point doesnt have a proof and result of wishful thinking. Back to back matche
11 25630,"It wont show in the post battle results, those are just for the battle.
12 25630,My aim is not convincing the people who are too afraid of admitting they knew wron
13 25630,There is actually a report going about about the lost RP from quitting matches ear
14 25630,"Can you give me a name of the topic or even a link so I can add my input? Because
```



7. Visualizations

- War Thunder Forum Data
- YouTube Data
- Previous Team's Dictionary

Aafia Siddiqui
Abbas Hamideh
Abdel Aziz Rantisi
Abdul Alim Musa
Abu Bakr al Baghdadi
Abu Musab al-Zarqawi
Abu
Abu Sayyaf
ISIS Philippines
ACAB
Accelerationism
Accel
Act for America
Active Clubs
R.A.M
Rise Above
Adam Green
Adrenaline control
Uhuru
Adrenochrome
Mujahideen

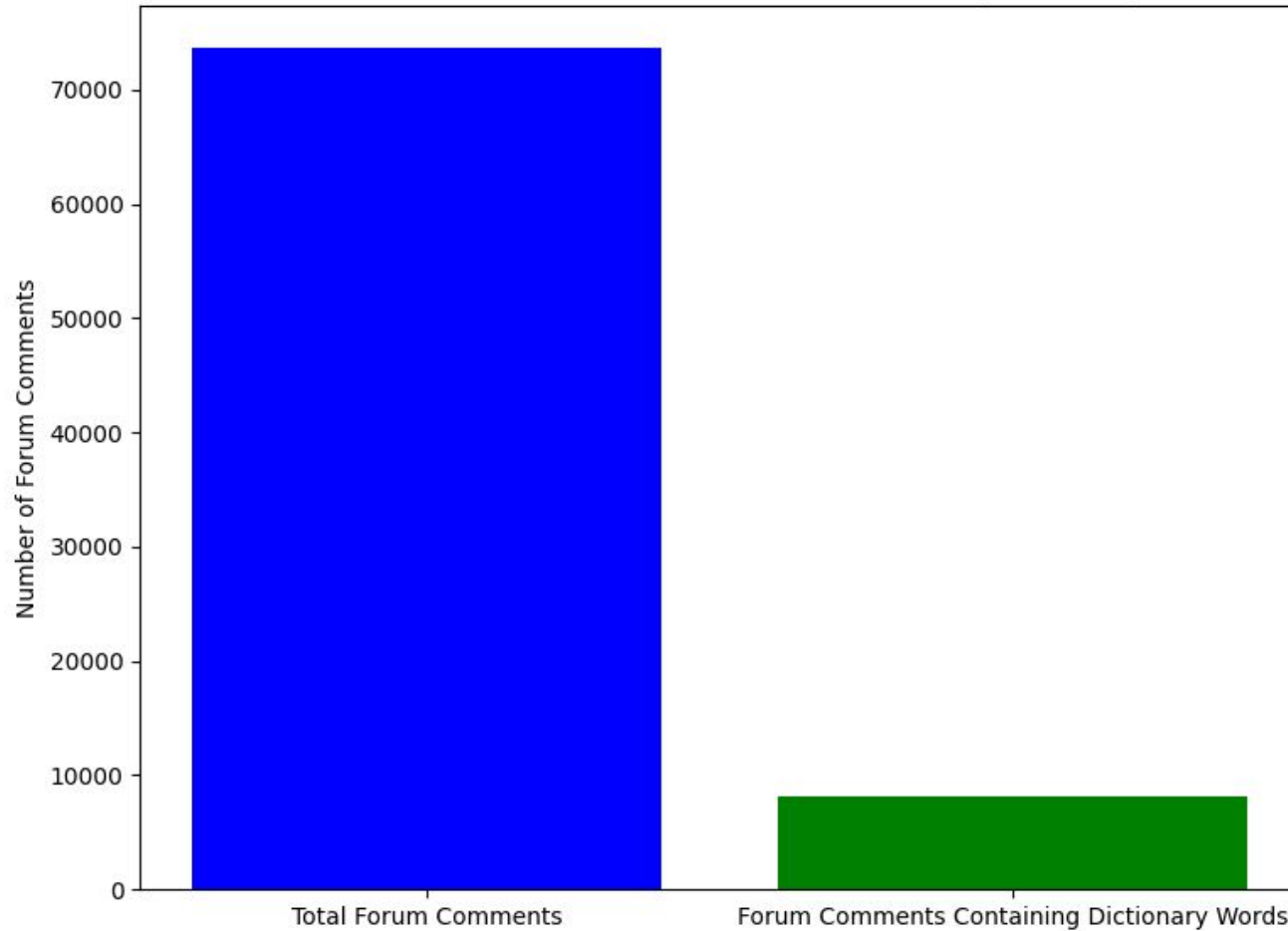
Extremist Dictionary

"Hate on Display", 2024, Anti-Defamation League, Accessed April 2024

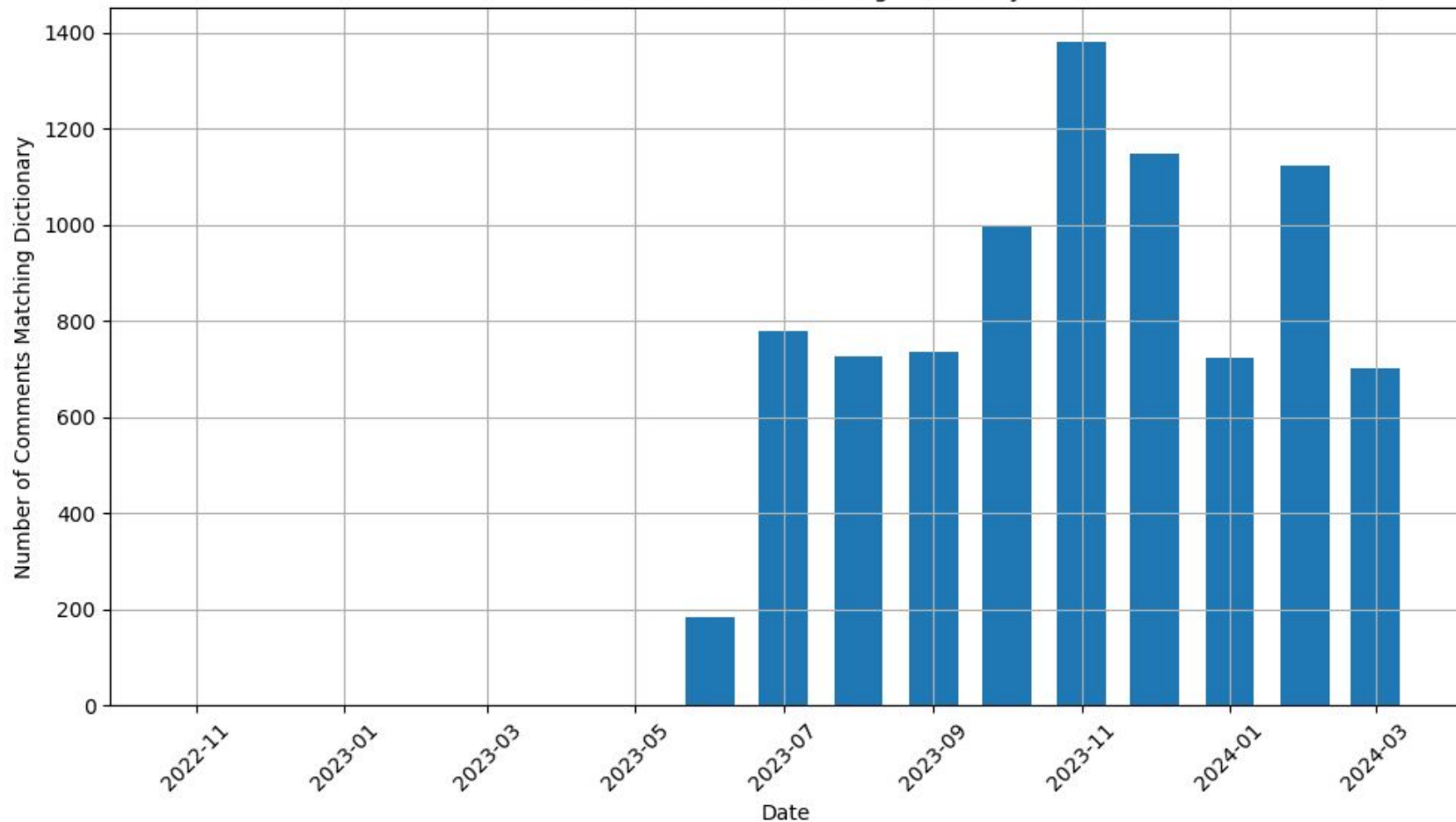
<https://www.adl.org/resources/hate-symbols/search?f%5B0%5D=topic%3A1703>

War Thunder Forum Visualizations

Total Forum Comments vs. Forum Comments Containing Dictionary Words



Forum Comments Matching Dictionary



Document Leak

Eurofighter
Typhoon DA7
fighter jet

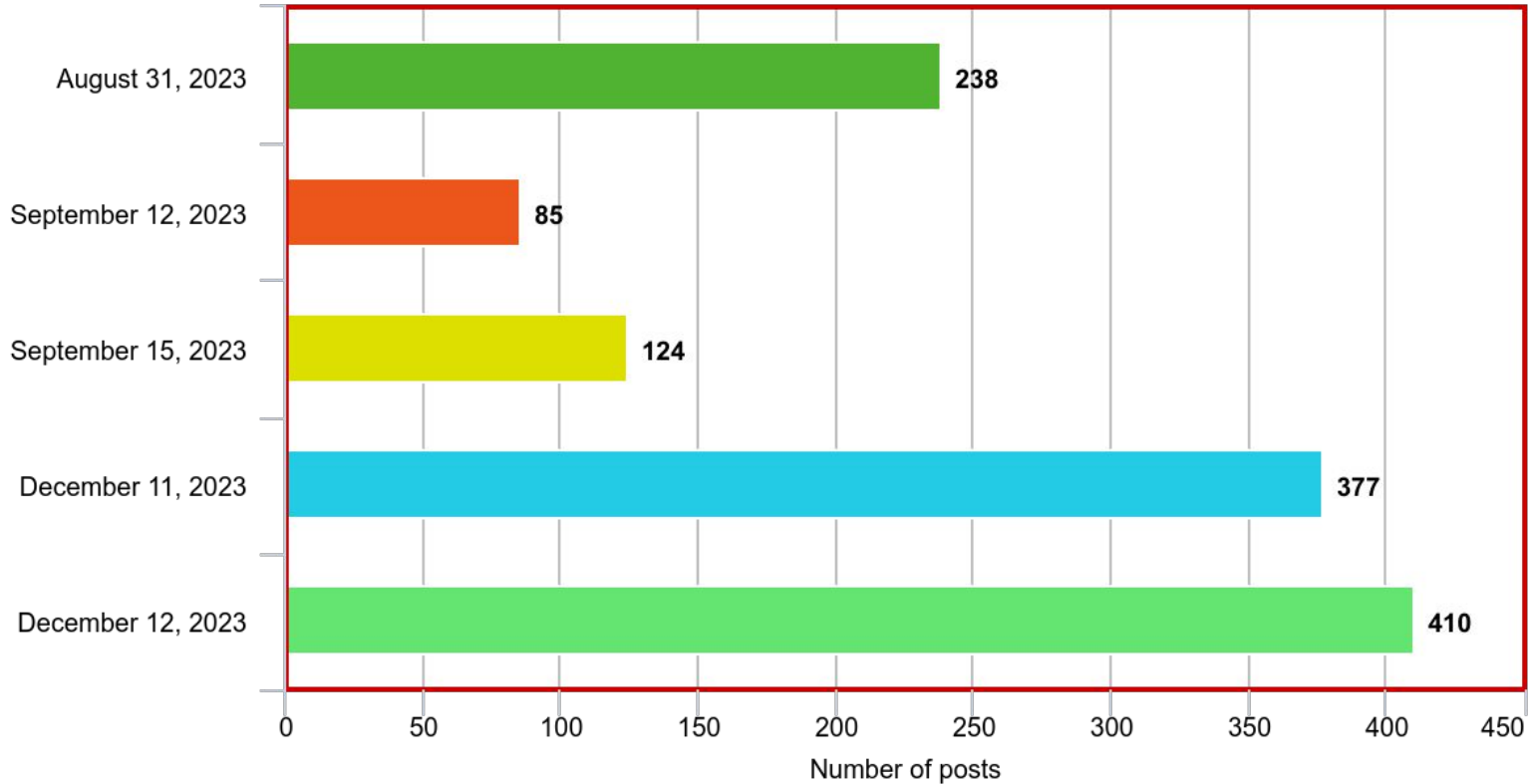
Lockheed F-117
Nighthawk stealth
attack aircraft

Boeing AH-64D
Apache helicopter

Norinco VT-4 tank

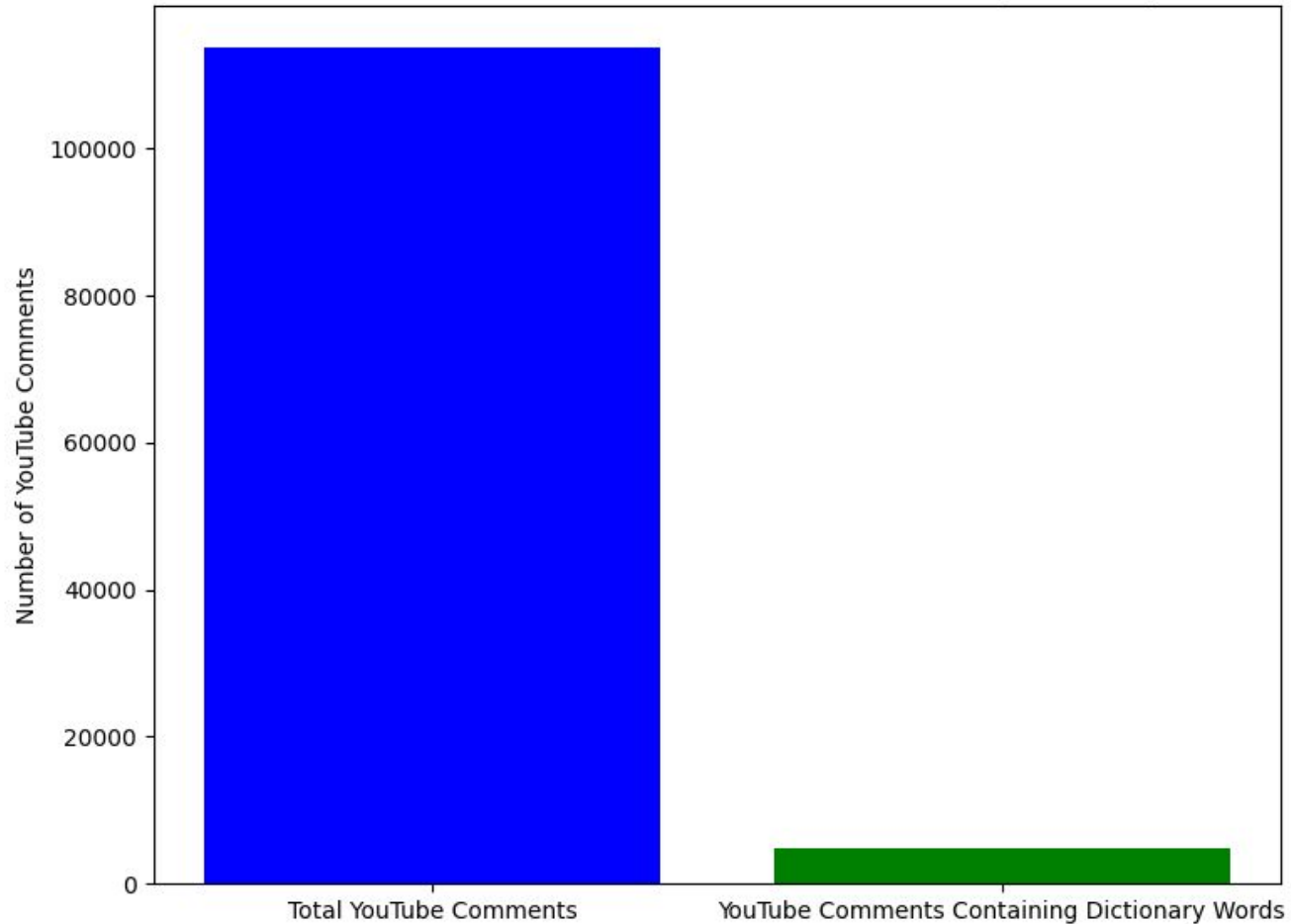
M2A2 Bradley
IFV

Number of posts on each respective leak date
Our scraped data from the War Thunder Forms

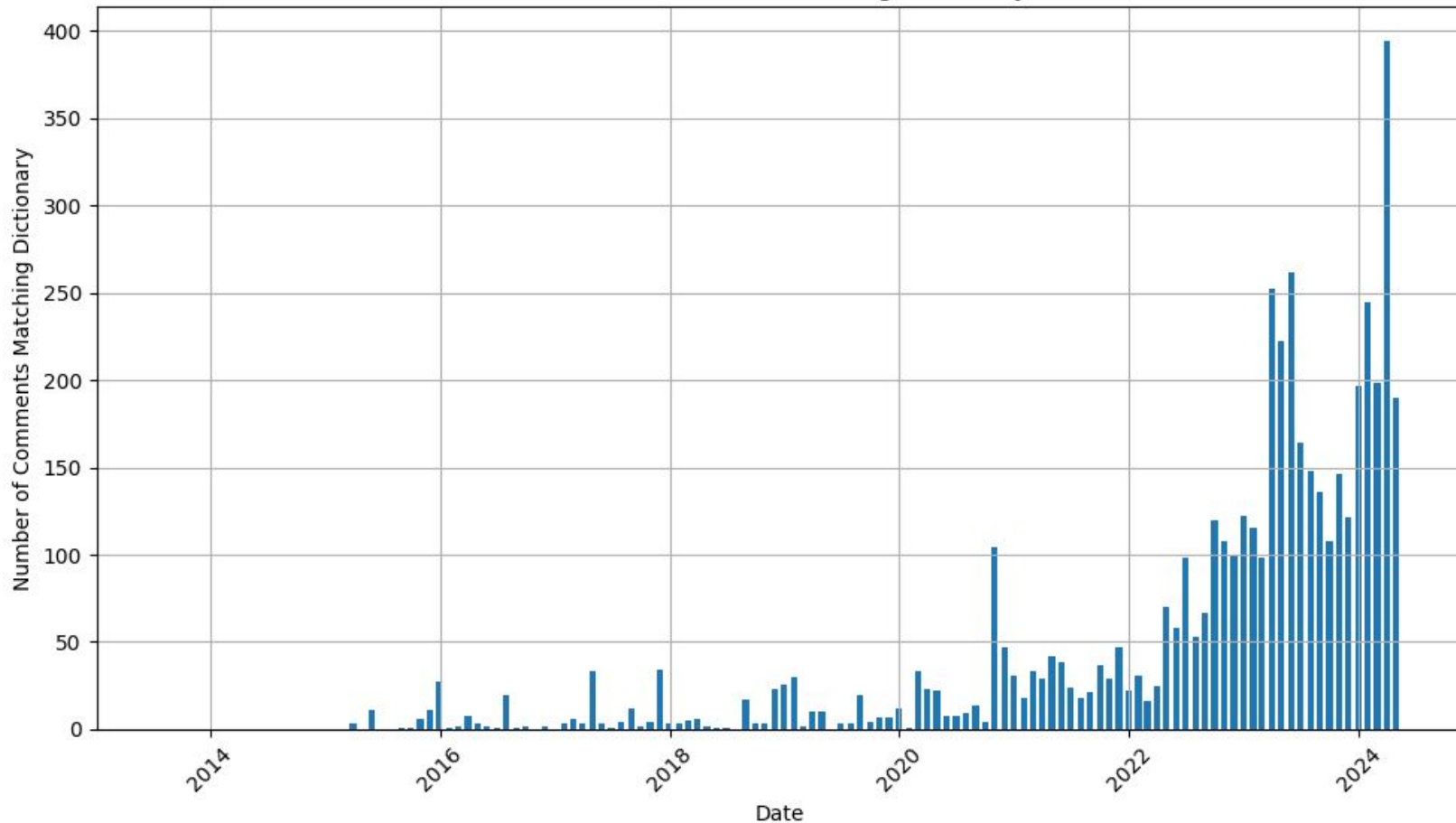


YouTube Visualizations

Total YouTube Comments vs. YouTube Comments Containing Dictionary Words



YouTube Comments Matching Dictionary





8. Future Work

- Scrape Old War Thunder Forum

9. Acknowledgements

- Client:
 - Dr. James D. Ivory
 - Michael Senters (Grad Researcher)
- Professor:
 - Dr. Edward A. Fox



Professor Fox



Dr. James D. Ivory



10. References

- James D. Ivory, College Of Liberal Arts and Human Sciences, Accessed January 2024
<https://liberalarts.vt.edu/departments-and-schools/department-of-english/faculty/Jimmy-ivory.html>
- Niko Olson, Mauricio Lovelace, Earl Carter, “Gaming and Political Extremism”, 2023, CS4624 team term project, Accessed January 2024
<http://hdl.handle.net/10919/115018>
- Gaijin Games Kft, War Thunder Forums, 2024, Accessed January 2024
<https://forum.warthunder.com/>
- Randie Pathirage, “YouTube Comment Scraping Made Easy with Apps Script”, May 16, 2023, Medium, Accessed March 2024
<https://medium.com/@randiepathirage/youtube-comment-scraping-made-easy-with-apps-script-a66acfdb5ffd>



10. References P2

- Soham, Script to export all sheets in a Google Sheet to CSV, Stackoverflow, Accessed March 2024
<https://stackoverflow.com/questions/28707169/google-spreadsheet-download-all-the-sheets-at-once/28711961#28711961>
- Google, Google Sheets Official API Documentation, April 3, 2024, Accessed March 2024 <https://developers.google.com/sheets/api/guides/concepts>

Questions