safeBook: A Tool to Increase User Autonomy on Facebook

Aayusha Shrestha
Computer Science, NYUAD
aayusha.shrestha@nyu.edu

Advised by: Prof. Yasir Zaki, Prof. Talal Rahwan

ABSTRACT

Targeted advertisements on Facebook have been a topic of contention as they are both intrusive and misleading. As users' feeds get occupied by advertisements and sponsored content, users may feel that they are losing a sense of autonomy: they are no longer able to control what they see. Facebook has been notorious for not only exchanging user information with third-party apps, but also for countering ad-blocking apps with its own anti ad-block mechanisms.

Facebook advertising and suggested content have not only raised concerns about user privacy, but they have also hampered user experience by showing users content they do not want to see. While multiple ad-blocking mechanisms have been made available to users in the past, they have failed due to their continuous feud with Facebook's anti-ad-detection mechanisms. For this project, we aim to take a novel approach to combat Facebook ads by using machine learning to predict the structure of the DOM and identify and block ads before it is rendered on the user's end. We also aim to explore topic-based filtering for Facebook posts and introduce users to a new way of dealing with unwanted content. We hope to explore the changes in user perception about privacy, security, usability, and autonomy on Facebook by enabling users to rid their feeds of unsolicited ads and use Facebook for its intended purpose – socializing.

KEYWORDS

online advertising, human-computer interaction, ad block, user privacy, topic-filtering

This report is submitted to NYUAD's capstone repository in fulfillment of NYUAD's Computer Science major graduation requirements.

جامعة نيويورك أبوظبي
NYU ABU DHABI

Capstone Project 2, Spring 2022, Abu Dhabi, UAE

Reference Format:
Aayusha Shrestha. 2022. safeBook: A Tool to Increase User Autonomy on Facebook. In *NYUAD Capstone Project 2 Reports, Spring 2022, Abu Dhabi, UAE*. 11 pages.

1 INTRODUCTION

Facebook, with its 2.7 million active user base, has gathered a fair amount of criticism regarding its impact on myriad aspects ranging from psychology and user behavior to cybersecurity and marketing. While Facebook's global reach is commendable in its ability to bring together people from all over the world, its shortcomings in maintaining user privacy, security, and safety deserves scrutiny. The proliferation of targeted Facebook advertisements demonstrates how the company is restricting user autonomy and compromising their security.

In 2020 alone, Facebook generated 97.9% of its global revenue from advertising [16]. As Facebook sustains itself mainly on this revenue, users' feeds get more occupied by sponsored content. This compromises user autonomy: not only are users no longer able to control what they see, but their privacy is also invaded as their personal information and behavior on the platform becomes accessible to third party advertisers and companies.

We conducted a user survey to understand user perceptions of privacy, security, and relevance of posts on Facebook (See section 5.2). Unsurprisingly, our survey results demonstrate that users are concerned about not only the compromised privacy, but also about the degraded user experience due to ads: 74.5% of participants have considered deleting or have deleted their Facebook accounts due to lack of trust in the way Facebook handles their private data and 71.9% of participants agreed that removing ads would improve their experience on Facebook.

The number of users who are dissatisfied with their Facebook experience is extremely high. If and when users become concerned about privacy and usability on Facebook, they may attempt to regain their autonomy by changing their privacy settings or using ad blockers. This doesn't work for two reasons. Firstly, Facebook deploys various mechanisms

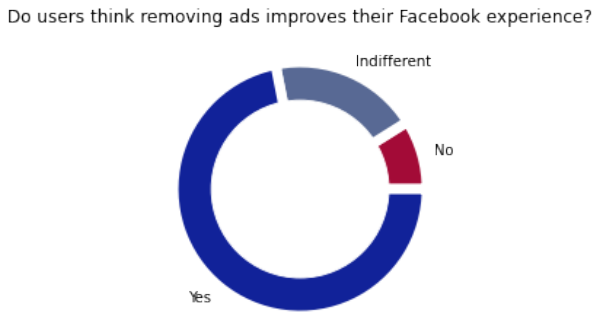

Figure 1: Do users think removing ads improves their Facebook experience?

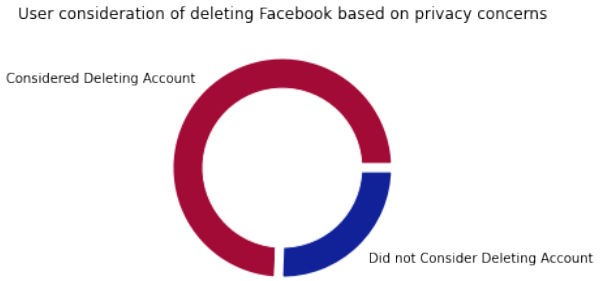

Figure 2: User consideration of deleting Facebook based on privacy concerns

to evade ad blockers, making them ineffective. Secondly, while it may seem as though users can limit ads on Facebook, Facebook only allows users to control how it uses data from third-party apps and sources for targeted ads. Users however, cannot stop Facebook from providing ads based on their interaction on the platform. Additionally, even though Facebook has features to hide a certain advertisement, it is not that the user is able to get less ads; one ad is simply replaced by another.

Both the intrusive nature of ads and the potential misinformation ads may circulate has given rise to concerns about user safety and autonomy. In the past, there have been expressions of frustrations with Facebook through everything from artistic expressions [12] to public backlash. In the past decade, Facebook's role in enabling Cambridge Analytica to influence presidential elections in 2016 and its role in impacting the mental health of its users as revealed by a whistleblower in 2021, has received global attention as people demand transparency from the company.

More and more people feel that organizations like Facebook "curtail users' freedom to choose their exposure to ads due to forced exposure," which "threaten(s)" users' freedom to use Facebook at their discretion [20]. This has prompted ad avoidance, anger, and frustration towards advertisers and platforms that enable them, all hindering the user experience.

This paper also explores the usability and autonomy in terms of relevance of posts for users. Our survey accounts for user perceptions of reporting unwanted content and the effectiveness of such mechanisms. 54.5% of participants claimed that reporting unwanted content on Facebook did not improve their Facebook experience, meaning that they still encountered unwanted posts.

Recognizing the overwhelming invasion of privacy and security of Facebook users along with the compromised user experience stemming from the proliferation of unwanted content, this project aims to build tools to give back users autonomy and freedom as they navigate through social media by deploying user-specified filters and blockers on Facebook. The project is designed for the use of Facebook on a web browser as we intend to leverage the Document Object Model (DOM) Tree structure. The DOM is the "object-oriented representation of the web page" where all properties, methods, and events of a webpage are organized as objects that can be accessed and modified with scripting languages [11].

```
1  <?xml version="1.0"?>
2
3  <body name = "body root">Body root
4      <div name="child 1">This is child 1 like
5          <div name="child 2">This is child 2,
6          nested inside child 1
7              <div name="child 4">This is child 4,
8              nested inside child 1 and 2</div>
9          </div>
10     </div>
11     <div name="child 3">This is child 3
12         <p name="para">Paragraph nested in child 3</p>
13     </div>
14 </body>
15
```

Figure 3: Sample XML for DOM Tree Visualization

This project has two components. The first component is a topic-based filter that relies on Natural Language Processing tools to identify the topic of a post. The second component is the ad-blocking mechanism. This project intends to make use of the DOM structure to identify patterns in the way Facebook writes their posts. We aim to gather the DOM structure of sponsored and unsponsored posts to feed into a machine learning model, which will predict whether a post is sponsored or unsponsored based on the DOM structure. Once the model detects an ad based on the structure, it will block ads before it is rendered on the users end.

Since social media is evidently an intersection of technology with human behavior and psychology, we want to take an HCI-centered approach – the driving force behind this

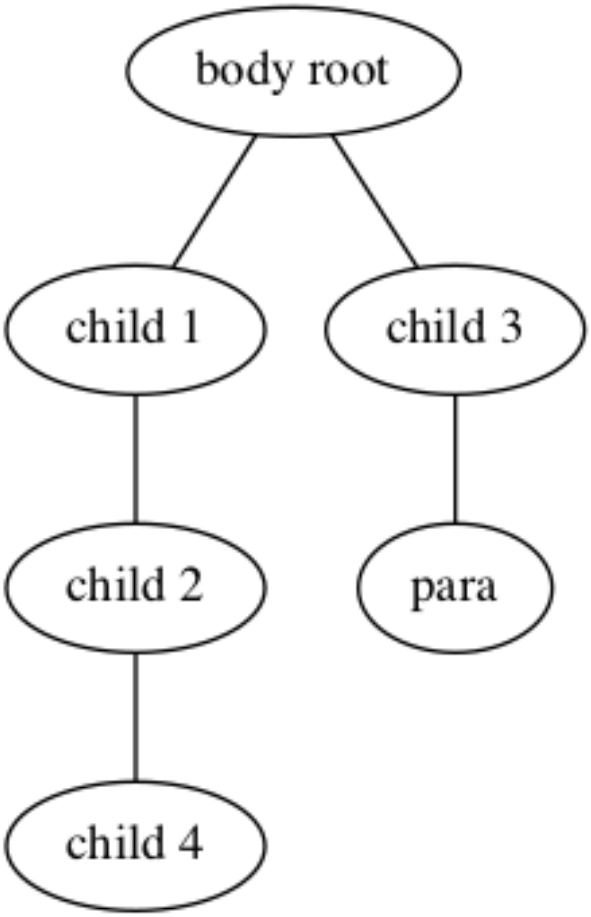

Figure 4: DOM Tree Visualization of Sample in Figure 1

research is to make users' feeds more relevant to them and enable them to regain their autonomy on social media platforms. We hope to enable users to use social media platforms like Facebook for its intended purpose – socializing.

2 RELATED WORK AND LITERATURE REVIEW ON AD BLOCKING

Online advertising is not limited to just social media platforms. Advertising across myriad online platforms have attracted global attention, and a lot of work has been done exploring advertising mechanisms along with ad-blocking and anti-ad blocking tools. In this section, we will be reviewing past literature on ad blocking mechanisms along with industry tools available in the market.

The most common ad blocking mechanism identified was the use of filter lists. Filter lists are mainly open source, and contain a set of matching rules that enable blocking of an advertisement based on its URL. Most ad blockers allow users to use one or multiple filter lists based on their preference while some embed filter lists as a part of their implementation [19]. Filter lists essentially regulate what the browser can fetch and they block requests that match their blacklist. The decisions of the filter lists are used by ad blockers to either remove the element encasing the URL from the DOM structure, or replacing the element with a white or gray box that covers the advertisement. While this is the most widely adopted mechanism used by ad blockers, anti ad-blocking tolls have been able to easily identify and block ad blockers based on filter lists. Snyder et. al [14] study the ways advertisements evade filter lists such as Easy List . They found that "90.16% of the resource blocking rules in EasyList provide no benefit to users in common browsing scenarios," [14] and that filter lists are often hard to maintain, and add a lot of deadweight that hinder page load times and ultimately, the user experience. Additionally, filter lists are more susceptible to being evaded: many advertisers simply change their URLs, remove ad keywords and image dimensions from their URL, or redirect advertisement resources to the first party rather than resourcing them from a third-party to avoid being detected by filter lists. For example, the URL https://c.betrad.com/geo/ba.js?r170201 was blocked by the EasyList rule ||betrad.com$third-party. The resource was moved to a new domain, c.evidon.com, to avoid being blocked [14].

Another common approach and complement to filter lists is the use of source code to identify ads or sponsored content. This method traverses through the website's source code, identifies keywords such as "sponsored" or "ad" and blocks or removes the element containing such words. However, this method can again be easily evaded by advertisers and platforms by DOM obfuscation. This exploits the limitations of element-based filtering based on the source code. For example, Tramèr et al. [18] illustrate how Facebook makes use of DOM obfuscation on a regular basis by breaking up words such as "sponsored" into something incomprehensible. We also found that Facebook adds filler words to make keywords undetectable, and even hides the sponsored tag when there is any interaction with it(such as hovering over the word "sponsored"). Additionally, Facebook also "embeds hidden ad-disclosure honeypots within regular user posts in an effort to deliberately cause site-breakage for ad-block users" [18]. Anti-ad blocking software also uses over-segmentation, where a large number of elements are injected into the DOM "to overwhelm an ad-blocker's classifier with inputs and exhaust its resources."

The paper AdVersarial: Perceptual Ad Blocking meets Adversarial Machine Learning discusses the use of and shortcomings within perceptual ad blocking. Perceptual ad blocking has been gaining popularity in recent years. Leveraging

the legal constraints that advertisers are expected to operate in, perceptual ad blocking works on visual cues to distinguish ads from regular posts or content. Advertisers, as per disclosure standards, are required to make ads recognizable to users, so they often carry logos or words to signal ad content. This mechanism uses computer vision combined with machine learning to identify ads as a human would. While this mechanism may be harder to evade for anti-ad blocking software, several attacks have been effective in tricking perceptual ad blockers. Similar to the obfuscation of URLs discussed earlier, anti ad-blocking software can simply change the visual display of ads. For example, the ad disclosure text and images can be rendered as noisy versions with perturbed logos to evade ad blocking [18]. Additionally, advertisers can also embed ad signals on regular posts to make ad blockers ineffective.

At present, there are numerous ad blockers available, mostly as browser extensions, which use one or a combination of the mechanisms discussed above. We will be reviewing a few ad blockers currently available in the market.

(1) Ublock Origin [8]
Ublock Origin is a browser-based ad blocker that uses two approaches: it relies on available filter lists and also crawls the html code of websites for keywords such as "sponsored." Ublock comes with its own set of filter lists, but it also allows users to add their own custom filter lists. As discussed previously, such mechanisms are easily evadable, through HTML obfuscation of keywords and using aliases for URLs. As advertisers continue changing their advertising mechanisms, ublock has been required to constantly update their algorithm, and has been ineffective for periods when advertisers bring forth significant changes [6].

(2) AdBlock Plus [2]
AdBlock Plus uses a combination of filter lists and perceptual ad blocking. It has preset filter lists and options for the user to add their own lists. It's Sentinel Project [18] uses web page screenshots and employs deep learning mechanisms to detect ads from the rendered web pages. It relies on ad disclosures and is, again, vulnerable to evasion techniques like obfuscation and image fragmentation. However, being vulnerable to evasion techniques is not what this ad blocker is scrutinized most for. It has an "acceptable ads" criteria, where it whitelists advertises who pay AdBlock Plus to have their ads reach users and be avoided by the blocker.

(3) Ad-highlighter [1]
Ad highlighter also uses perceptual ad blocking to identify ad disclosures like AdBlock Plus. In addition to identifying texts and images, it detects links to ad-policy pages to identify ads. Ad-Highlighter detects the AdChoices logo by comparing each image in a page to a template using average hashing: for each image, a hash is produced by resizing the image to a fixed size and setting the ith bit in the hash to 1 if the ith pixel is above the mean pixel value. An image matches the template if their hashes have a small Hamming distance [18]. It is also susceptible to being tricked by obfuscation, especially as it "assumes that all img tags in the DOM are shown as is, thereby ignoring potentially complex CSS transformations applied when rendering HTML. This can cause the downstream classifier to process images with unexpected properties" [18].

(4) Percival [10]
Percival also uses perceptual ad blocking by analysing the browser's image rendering. It segments the page into frame and uses image classification to identify ads. It also takes the structure of the web page source code, looks for images within elements, decodes the image, and blocks it from rendering if it's an ad. Based on issues within other ad blockers, it is possible that if the source code contains ad-like images within regular posts, but the images are not rendered, Percival may block the regular post based on its algorithm. Additionally, Percival is also limited as it covers up the image of the ad, but does deal with any text associated with it. So, the ad image is hidden, but there remains a blank space and ad text left on the user's feed, occupying space.

3 AD WARS - THE CASE OF FACEBOOK

Throughout the past decade, Facebook has been subject to interrogation and scrutiny by regulators and privacy advocates about its privacy practices [17]. It has also been accused of "discrimination, inciting social division, micro-targeting, single-house-based targeting and disclosure of personal data to advertisers" [9]. While Facebook has been under fire for the way it handles sensitive user information, it has also received criticism on the forced nature of advertisements. Users tend to perceive advertisements, even if they are "pertinent" as Facebook claims them to be, to be an interruption to their goal in using Facebook at their own discretion [20]. Additionally, ads that are too pertinent may even seem creepy to an extent that users feel they are constantly under surveillance, as if Facebook is "listening" to them [**?**]. The invasive and interruptive nature of Facebook ads has led to a backlash from the user base: a study [9] on psychological reactance

to exposure to ads concluded that ads evoke negative emotions such as anger, being threatened, and restricted, making them avoid ads and perhaps even platforms that advertise excessively.

As of 2020, over 20% of a user's Facebook feed was taken up by ads. In order to protect the user experience, ad blockers have been at an arms race with Facebook as both parties keep updating their algorithms and mechanisms to try to go undetected by the other party. Facebook uses ad obfuscation techniques such as "making the markup of ads sufficiently similar to that of regular Newsfeed posts that the two could not be distinguished by filter-list-based ad blockers" [15], making it impossible for ad blockers based on CSS selector based techniques to use it.

Facebook is more dangerous than other advertising platforms as it has access to a large amount of personal data, which users make available unknowingly. Facebook collects user activities such as likes and ad clicks, and combines the information with the user's demographic information so that it can facilitate targeted ads for the advertisers. This is then used in advertising and retargeting of ads from other platforms, cluttering the user's feed with ads.

While Facebook does have privacy settings that allow the user to moderate who gets access to their information such as home address, work place, political inclination, gender, "such privacy controls are not meant to restrict the access of the service provider to this information or the subsequent processing of this information for different purposes" [9]. In addition to that, most users may not be aware of how their activity is tracked, and may be giving advertisers more to work with based on their likes, comments, and time spent on something.

Additionally, the privacy settings made available to users are getting more and more vague as well as complicated. This lack of transparency further exacerbated user autonomy on being able to take back control of their social media experience. The ad settings do not "clearly distinguish how users give consent or exercise their right to withdrawal of consent and the right to object to data processing activities" [9]. Facebook provides an option for users to opt out of giving information to third-party apps, but it also implies that users will not be able to opt out of ads on Facebook itself. Nor can users opt out of Facebook tracking their activity within the platform. Additionally, even if a user decides to hide "sensitive information" such that it may not be used for advertising, Facebook has influence over other apps and can easily infer the hidden information through third party data and use it for providing ads [9].

As the content structures on Facebook feeds have become more complex, so have the settings regarding them. Users are no longer able to fully be aware of, find, and navigate their feed settings. Feed settings are crucial in maintaining user satisfaction on online platforms as they allow more autonomy of the user's end. A study found that "were unaware of their feed settings, and had difficulty navigating and understanding feed settings, especially ad settings" [13]. This further illustrated how user freedom is being restricted through multiple mechanisms.

4 RESEARCH GOALS

Upon reviewing existing tools and past literature on ad-blocking, it is evident that the most common approaches to ad-blocking is the use of text scraping, filter lists, and image recognition. Since all these methods have been evaded by image and text obfuscation, this project is based on a novel approach that uses machine learning on the DOM structure itself.

The primary goal of this project is to study the DOM structure of Facebook and identify patterns in the structure so that machine learning algorithms can be deployed to predict the DOM structure if and when Facebook decides to change it. We want to leverage the fact that there will always be a DOM tree structure that Facebook operates in and want to explore if there are any patterns in the way they structure and modify it. We hope to investigate the differences in the structures of a sponsored post as opposed to a post by someone or something a user follows. Our goal is to identify where and how ads are placed within the DOM structure, extract the images and text associated with it for further analysis, and delete both the image and text associated with it so that the user's feed will have more content they signed up for instead of unsolicited advertisements.

Due to a lack of past work on topic-based filtering on Facebook, we hope to establish an understanding of user perceptions about the topic-based filtering and annoyance regarding specific topics. We aim to further facilitate autonomous control for users by introducing filters based on topics. We currently achieve this by implementing a Natural language processing tool to allow users to filter out content about certain topics.

Our next step is to conduct user studies to study user attitudes towards a "cleaner" Facebook. It will be interesting to evaluate changes in user patterns and engagement with Facebook when ads are filtered out. Our goal is to increase user satisfaction and productivity on the platform.

We will be evaluating the success of our tool by measuring the difference between the number of ads on users' feeds, the maintenance of page load time with our tool, and the qualitative analysis of user perceptions towards privacy, autonomy, and usability on Facebook when they use our tool. We hypothesize that our tool will be able to declutter users' Facebook's feeds of ads and increase usability and user satisfaction.

While our goal will be to develop a tool that goes undetected by Facebook in order to ensure that our tools aren't disabled, we will make the tool open source. We hope that we will garner public interest, and perhaps also the interest of Facebook so that they can use the results of our studies to make further developments in their advertising policies and orient their goals towards users and not advertisers.

5 METHODS AND FINDINGS

We adopted multiple methods and approaches to navigate this project. We created a new Facebook profile for the purpose of this project and followed multiple pages and users and also added other dummy profiles as friends. We created posts and engaged with the dummy account's feed by scrolling, liking posts, and searching for items. as well to make the experience as authentic as possible.

(1) Findings about Ad Wars
In the initial stages of the project, we spent time manually investigating the Facebook code. We found that Facebook does indeed obfuscate words such as "sponsored" to make ads undetectable through text finding and comparison, but what was more interesting was that Facebook seems to inject the word "sponsored" into regular posts. We believe that this is done to make ad blockers inefficient, as they would block posts users might actually want to see.
Additionally, during the course of our project, we also confirmed that Facebook does indeed change the structure of its DOM by introducing new patterns and non-element nodes in its DOM structure.
We also found that at least 20% of one's Facebook feed is taken up by sponsored content. This does not include suggested content. On the higher end, as much as 40% if a user's Facebook feed is taken up by ads. However, it is important to note that these numbers are based on the Facebook profiles we used for testing purposes, and that the number may vary on an actual user's account.

(2) Survey to understand user needs and sentiment
(a) Survey methods and goals
We conducted an online survey using Qualtrics [5] and Prolific [4] to gain insight about users' perception about privacy and security on Facebook, their experience with ads and suggested content, and their attitudes towards the potential of introducing ad based filters. Our findings signal that most users are strongly opposed to advertisements and suggested content on Facebook and many users believe that deploying ad blockers and topic filters will improve their experience on Facebook.
Participants answered up to 20 questions (depending on their survey flow which was condition-based), which consisted of multiple choice questions, text response questions, and Likert scale. We asked questions about the degree to which participants found ads and suggested content useful or annoying, if they expect their experience to improve when ads, suggested content, and content on topics they don't wish to see are removed, and their experience with reporting content on Facebook. We did not ask for demographic information as we extracted that from prolific using the user's prolific ID. The full survey can be found in Appendix A.
Participants were recruited using Prolific's recruitment resources. We sampled 500 participants and ensured that there was a balanced sample in terms of gender. We also pre-screened participants to ensure that our participants were Facebook users. We recruited participants from the United States as we wanted to capture the changes in perceptions of privacy and security on Facebook after the 2018 Cambridge Analytica Exposure and the Whistleblower accusation in 2021. The U.S. was chosen as its citizens were directly impacted by the Cambridge Analytica scandal and would be more aware about it.
(b) Findings
(i) On Trust, Privacy, and Security

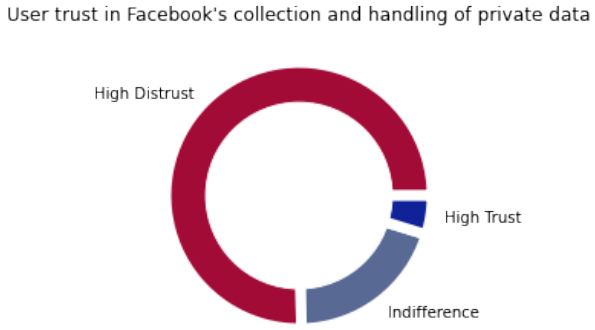

Figure 5: User trust in Facebook's collection and handling of private data'

The survey findings reflected a deep distrust with both Facebook and third party ad-clockers and a resentment for ads. According to our survey, 75.25% of respondents expressed that they did not trust Facebook in terms of privacy as they believed that Facebook shared private user data with third parties. This distrust and privacy concerns has made

74.46% of respondents consider deleting their Facebook accounts.

(ii) On Sponsored and Suggested Content
The participants of our survey expressed not just privacy concerns, but also annoyance on Facebook. More than two-thirds of the respondents reported that they found ads highly annoying and more than half of the respondents reported that they found suggested content highly annoying. This annoyance most likely stems from the irrelevance of ads to users based on the fact that on average, participants have made less than 2 purchases based on ads when most of them have been on Facebook for over 10 years.
Apart from expressing annoyance, the responses also expressed a concern regarding the ulterior motives of Facebook advertisers as 52.47% of participants reported that they think suggested and sponsored content on Facebook aims to influence social and political views of users. Due to a combination of both compromised usability and privacy concerns, 67.3% of participants claimed that ads take up too much space on their newsfeeds and 71.88% of participants agreed that removing ads would improve their user experience on Facebook. Even though participants expressed the desire to get rid of ads, they also expressed distrust regarding the way ad blockers deal with their private data. Almost three quarters of participants reported that they think ad blockers also collect private information and that they do not trust the way they handle this information.

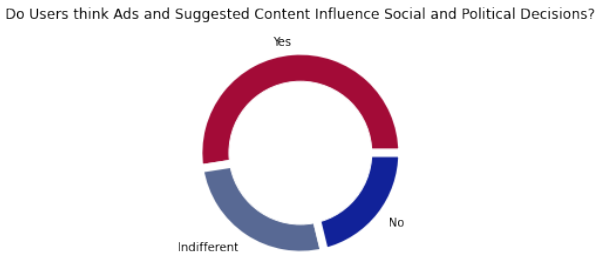

Figure 6: Do Users think Ads and Suggested Content Influence Social and Political Decisions?

(iii) On Topic-Based Content
We asked participants if they have ever reported content they did not like: 54.45% of participants who responded that they reported content expressed disappointment in the mechanism as reporting did not improve their Facebook experience. Over one-third of the participants even reported that their Facebook experience was hampered by posts on topics they did not wish to see. While topic-based filters have not yet been implemented and are in no way as prevalent as ad-blockers, participants seemed keen to try out a topic-based filter as 46.34% of participants strongly believed that using a topic-based filter would improve their Facebook experience. 80.4% of participants were not opposed to trying a new tool that blocks posts based on a user-specified topic filter.

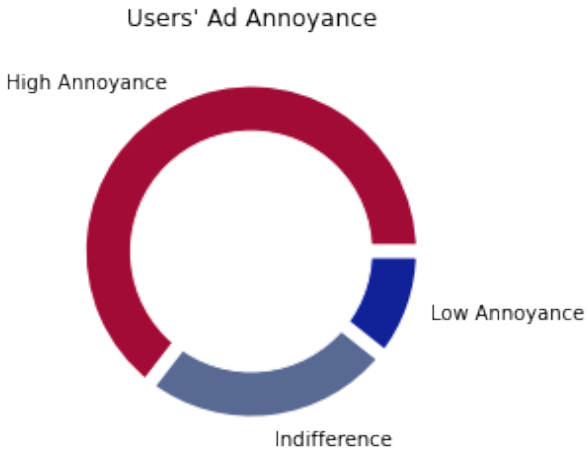

Figure 7: Users' Ad Annoyance

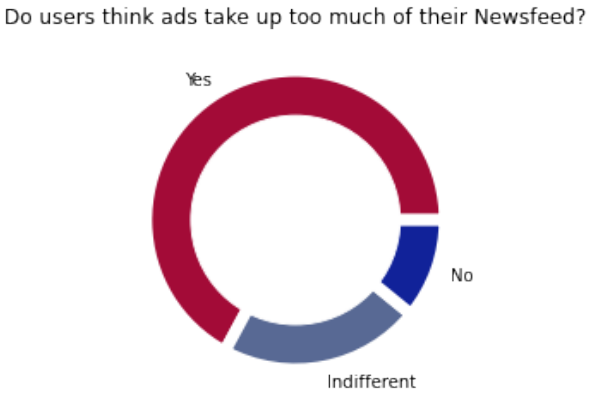

Figure 8: Do users think ads take up too much of their Newsfeed?

(3) Visualizing DOM trees to identify patterns
One of the first steps we took was to explore patterns in the DOM structures of Facebook posts. Upon exploring the DOM structure, we were able to identify that each post on Facebook is labeled as a "FeedUnit", and

1 in 3 participants reported that their Facebook experience is hampered by posts on unwanted topics

Figure 9: Are users' Facebook experience hampered by posts on unwanted topics?

from this "FeedUnit", we are able to extract the content of the post. The extraction of contents was useful in implementing the topic-based filter (See section 5.5).
Since we were not focused on the content but on the structure at this stage, we made use of Selenium for Python and Python ElementTree XML API to scrape Facebook's HTML structure and successfully visualize the DOM tree structure of Facebook posts. We achieved this in the following manner: First, we automated the login and scrolling of the Facebook newsfeed as the posts need to be interacted with for our scraper to detect them successfully. We then looked up each individual "FeedUnit" and extracted its XML structure and stored it as a web element.
We take each "FeedUnit" element, traverse its content, and store it in an in-order list, where each node is an XML web element represented by its XML tag and attributes. Storing the information as web elements and in an in-order list allows us to preserve all the information, including the hierarchies in the structure and the unique attributes of each node.
We then traversed the in-order list of web elements and created edges to establish relationships between nodes to prepare it for visualization: each node was connected to another node through an edge. We made use of the PyGraphViz interface to visualize the tree.
While there seemed to be patterns in the DOM trees of Facebook posts, they were far too complex for us to assume based on human perception, which is why we shifted to a Machine Learning based approach.

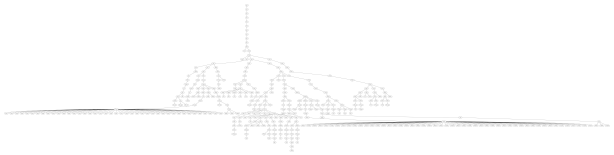

Figure 10: Visualization of a single FeedUnit, i.e. a single Facebook post (see Appendix)

(4) Data gathering for Machine Learning Model After we gathered the DOM structure of each Facebook post and visualized it, we realized that it was not possible for the human eye to detect patterns in the DOM structure and so we started collecting labeled data for the Machine Learning algorithm to learn from.
Our goal here was to gather the DOM structure of posts and categorize them as "sponsored" or "unsponsored". We achieved this goal by using a combination of computer vision and image recognition along with web browser automation.
As we illustrated earlier, we cannot rely on textual analysis to detect sponsored content. We decided to use image recognition instead of textual analysis as image recognition to identity sponsored content works in the short term until Facebook decides to obfuscate the "Sponsored" tag on its posts. We wanted to leverage the AdChoices requirement that mandates the "Sponsored" tag on Facebook posts and used this tag to identify sponsored content. Since we only needed this method for extracting data, a short-term image recognition approach was adopted as it would not be impacted by the longer run risks of image obfuscation.
We launched Facebook using Selenium and took a screenshot of the entire newsfeed. We compared the newsfeed against the "sponsored" tag and the three dot menu of each post to identify which posts were sponsored based on the proximity of a three dot menu to the sponsored tag. We kept a record of all posts based on the three dot menu and labeled sponsored posts.
After having a record of which posts are sponsored, we extracted the HTML DOM tree of the newsfeed and extracted the structure of each post. We compared this list of DOM trees to the list of labels, and stored the DOM tree of each post into either a "Sponsored" folder or an "Unsponsored" folder as a text file. We now have DOM trees of individual Facebook posts categorized based on their sponsored status.
(5) Topic Filter
For the topic filter, we took a textual-analysis-based approach. We extracted the textual content from each post element and pre-processed the text by getting

rid of random characters placed for obfuscation and words like “Like”, “Share”, “Comment”. It should be noted that the textual information we gathered is of the post only, which includes alt text, but does not include text present in the comments section.

Our initial approach was to identify high frequency words in a post. For this, we further pre-processed our text by removing stop words (articles, prepositions, pronouns, conjunctions, etc). We then identified words with the highest frequency. However, this did not give us accurate information about the posts we were analyzing. It also did not give us words generic enough to be considered a topic or genre.

We switched to a different approach and adopted the use of TextRazor [7], a Natural Language Processing API, to get the subject of each post. We took the pre-processed text and fed it into the TextRazor classifier, which gave us topics that aligned with the International Press Telecommunications Council Media Topic NewsCodes [3]. However, the TextRazor classifier gave a very detailed classification with thousands of genres. For the initial stages of the project, we decided to limit the genres and topics to 17 topics as per the broadest IPTC Media Topic NewsCodes (see appendix for details). We achieved this by utilizing regular expressions in Python to extract only the broad topics and avoid the detailed subtopics.

After we labeled each post with a standard newscode, we asked the user what topics they wished to block and compared the user’s desired label against the labels of all posts. We made use of Selenium to inject our own HTML into Facebook’s HTML and blocked posts whose label matched with the label the user wished to block. It is important to note that we didn’t remove the posts completely, but rather placed a color block over the post. This is to show the user how much of their newsfeed is occupied by unwanted content and also to allow the user to access the post if they want to. The colored block can be removed by the user by clicking on a checkbox on the blocked post. This is to ensure that the user has the autonomy to decide whether they want to see a post. Users can choose to hide and unhide content as they please.

6 CHALLENGES

During the course of this project, we faced multiple challenges and expect some more challenges as we implement the project further. Here are a few encountered and expected challenges:

(1) The changing DOM structure of Facebook posed an initial challenge as it required a lot of code restructuring. However, as Facebook kept altering its DOM structure, we kept implementing more checks and made more generalizations in parsing the DOM structure. This only strengthened our approach as it is now robust and constantly improving to accommodate changes in the Facebook DOM.
(2) As mentioned earlier, Facebook constantly changes its ad signals, such as the Sponsored label and the ad choices logo. This posed a minor challenge for us as we used image recognition techniques to label posts as sponsored or unsponsored before collecting their DOM information. For the purpose of this project, image recognition is only used for data collection, and so we kept updating our sample image every time Facebook changed the labels, so that we could successfully match the sponsored label. Once we gather enough data for the Machine Learning algorithm to learn from, the image obfuscation will not impact the project as we will stop relying on it.
(3) A future challenge we must keep in mind is the lack of user trust in ad blockers. This could entail that there is less trust in not just ad blockers, but also in any third-party application such as our intended topic-filter. Our survey showcased that 67.6% of users think ad blockers collect private data from users and 73.2% of these users do not trust the way ad blockers handle their private information. In order to garner user trust in our products, we must make them aware about the open source nature of our program and illustrate how their information is not stored by us.

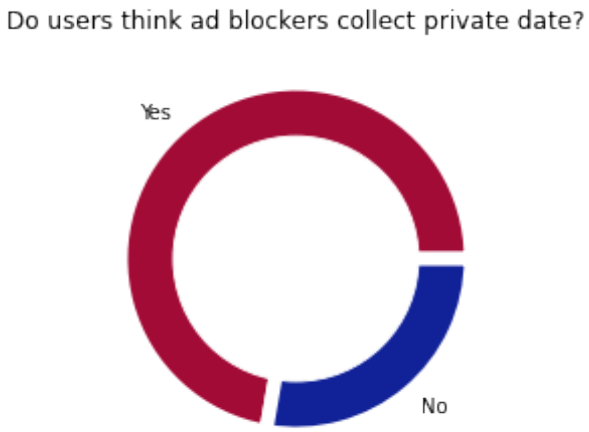

Figure 11: Do users think ad blockers collect private date?

Since our tool involves a lot of DOM scraping, users may feel uncomfortable as their information, such as

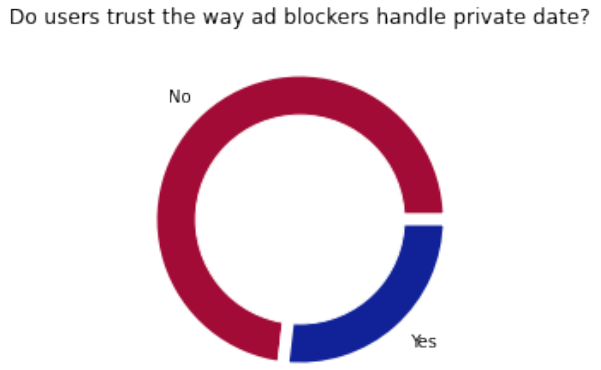

Figure 12: Do users trust the way ad blockers handle private date?

pages and posts on their newsfeed, will be scraped. One way to overcome this is to replace any text or image we gather from the DOM with random characters and placeholders to protect user information and privacy. This way, we will only be gathering the structural DOM while preventing the collection of any information associated with the user.

7 NEXT STEPS

Our next goals for this project are to conduct user tests for our topic-filter tool and implement a Machine Learning algorithm based on the labeled data we have gathered during this project. Before we involve users in testing, we also need to anonymize their data by replacing any identifiers with random characters and placeholders as mentioned in section 6.3.

8 CONCLUSION

This research aims to understand user needs and increase user satisfaction on Facebook by removing unsolicited and intrusive content and allowing users to take back control of their social media space. In order to achieve an ad-free and user-centric Facebook, we aim to introduce topic-based filter and use novel techniques based on ML prediction of DOM structures instead of filter lists and computer vision, as these methods have been evaded by anti-ad-blocking software in the past. We hope to use this study to understand user perceptions about safety, privacy, autonomy, and usability on Facebook and send signals to Facebook itself about the dire need for it to correct its intrusive advertising mechanisms.

REFERENCES

[1] [n.d.]. Ad Highlighter. https://chrome.google.com/webstore/detail/ad-highlighter/micfanbigkmkoeadjclckplnjhioedka?hl=en

[2] [n.d.]. Adblock Plus. https://adblockplus.org/

[3] [n.d.]. Media Topics (en-GB)- IPTC NewsCodes controlled vocabularies. https://www.iptc.org/std/NewsCodes/treeview/mediatopic/mediatopic-en-GB.html

[4] [n.d.]. Prolific. https://www.prolific.co/

[5] [n.d.]. Qualtrics. https://www.qualtrics.com/uk/free-account/

[6] [n.d.]. r/Adblock - Adblock not working on Facebook? https://www.reddit.com/r/Adblock/comments/bewjdd/adblock_not_working_on_facebook/

[7] [n.d.]. TextRazor - The Natural Language Processing API. https://www.textrazor.com/

[8] [n.d.]. UBlock Origin. https://ublockorigin.com/

[9] Sourya Joyee De and Abdessamad Imine. 2020. Consent for targeted advertising: the case of Facebook. *AI & SOCIETY* 35, 4 (Dec. 2020), 1055–1064. https://doi.org/10.1007/s00146-020-00981-5

[10] Zainul Abi Din, Panagiotis Tigas, Samuel T. King, and Benjamin Livshits. 2020. {PERCIVAL}: Making In-Browser Perceptual Ad Blocking Practical with Deep Learning. 387–400. https://www.usenix.org/conference/atc20/presentation/din

[11] MDN Web Docs. [n.d.]. Introduction to the DOM - Web APIs | MDN. https://developer.mozilla.org/en-US/docs/Web/API/Document_Object_Model/Introduction

[12] Ben Grosser. [n.d.]. Safebook. https://chrome.google.com/webstore/detail/safebook/cbicnijdcimabdbbpneihlmljcipmdmh?hl=en

[13] Silas Hsu, Kristen Vaccaro, Yin Yue, Aimee Rickman, and Karrie Karahalios. 2020. Awareness, Navigation, and Use of Feed Control Settings Online. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376583

[14] Peter Snyder, Antoine Vastel, and Benjamin Livshits. 2020. Who Filters the Filters: Understanding the Growth, Usefulness and Efficiency of Crowdsourced Ad Blocking. *arXiv:1810.09160 [cs]* (May 2020). http://arxiv.org/abs/1810.09160 arXiv: 1810.09160.

[15] Grant Storey, Dillon Reisman, Jonathan Mayer, and Arvind Narayanan. 2017. The Future of Ad Blocking: An Analytical Framework and New Techniques. *arXiv:1705.08568 [cs]* (May 2017). http://arxiv.org/abs/1705.08568 arXiv: 1705.08568.

[16] H. Tankovska. [n.d.]. Facebook's advertising revenue worldwide from 2009 to 2020. https://www.statista.com/statistics/271258/facebooks-advertising-revenue-worldwide/ Source: Facebook Annual Report 2020, page 66.

[17] The New York Times. 2018. Mark Zuckerberg Testimony: Senators Question Facebook's Commitment to Privacy. *The New York Times* (April 2018). https://www.nytimes.com/2018/04/10/us/politics/mark-zuckerberg-testimony.html

[18] Florian Tramèr, Pascal Dupré, Gili Rusak, Giancarlo Pellegrino, and Dan Boneh. 2019. AdVersarial: Perceptual Ad Blocking meets Adversarial Machine Learning. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security (CCS '19)*. Association for Computing Machinery, New York, NY, USA, 2005–2021. https://doi.org/10.1145/3319535.3354222

[19] Craig E. Wills and Doruk C. Uzunoglu. 2016. What Ad Blockers Are (and Are Not) Doing. In *2016 Fourth IEEE Workshop on Hot Topics in Web Systems and Technologies (HotWeb)*. 72–77. https://doi.org/10.1109/HotWeb.2016.21

[20] Seounmi Youn and Seunghyun Kim. 2019. Understanding ad avoidance on Facebook: Antecedents and outcomes of psychological reactance. *Computers in Human Behavior* 98 (Sept. 2019), 232–244. https://doi.org/10.1016/j.chb.2019.04.025

9 APPENDIX

Here are the visualizations of the DOM Structures of some FeedUnits, i.e. Facebook posts

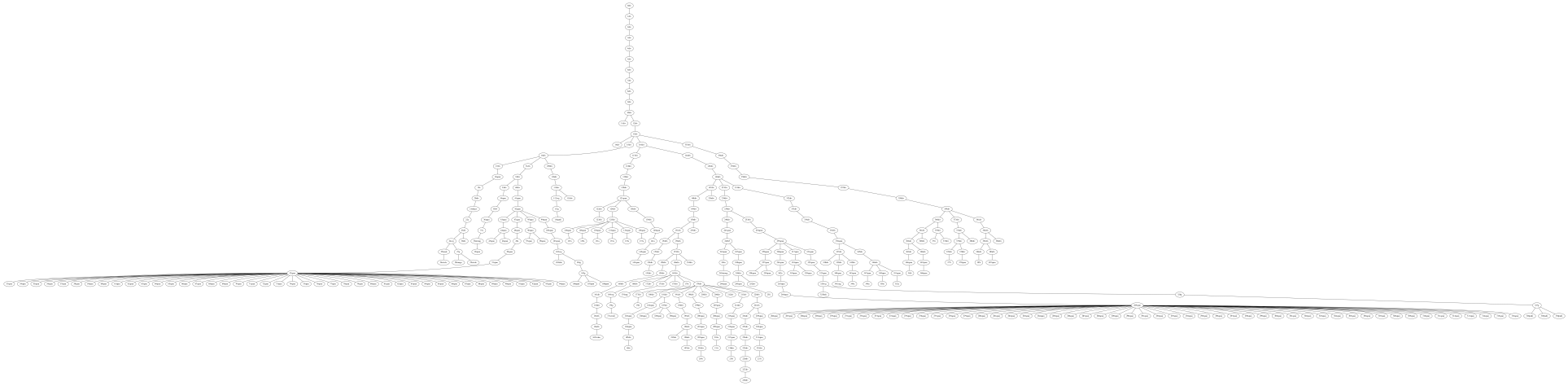

Figure 4: Visualization of a single FeedUnit, i.e. a single Facebook post

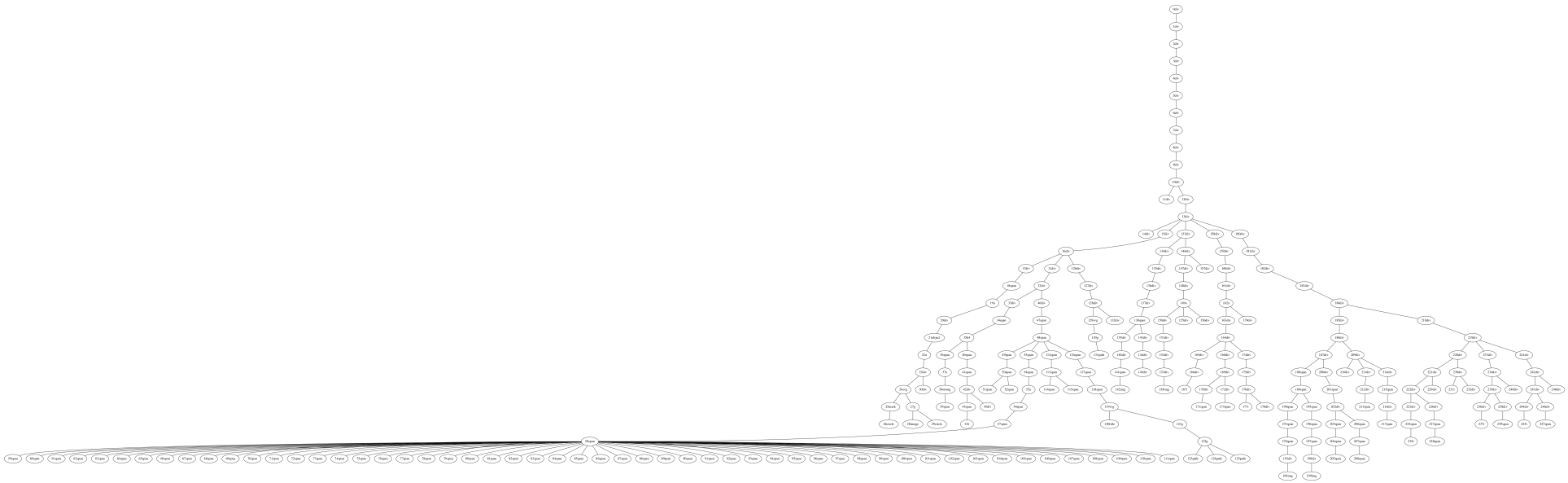

Figure 5: Visualization of a single FeedUnit, i.e. a single Facebook post

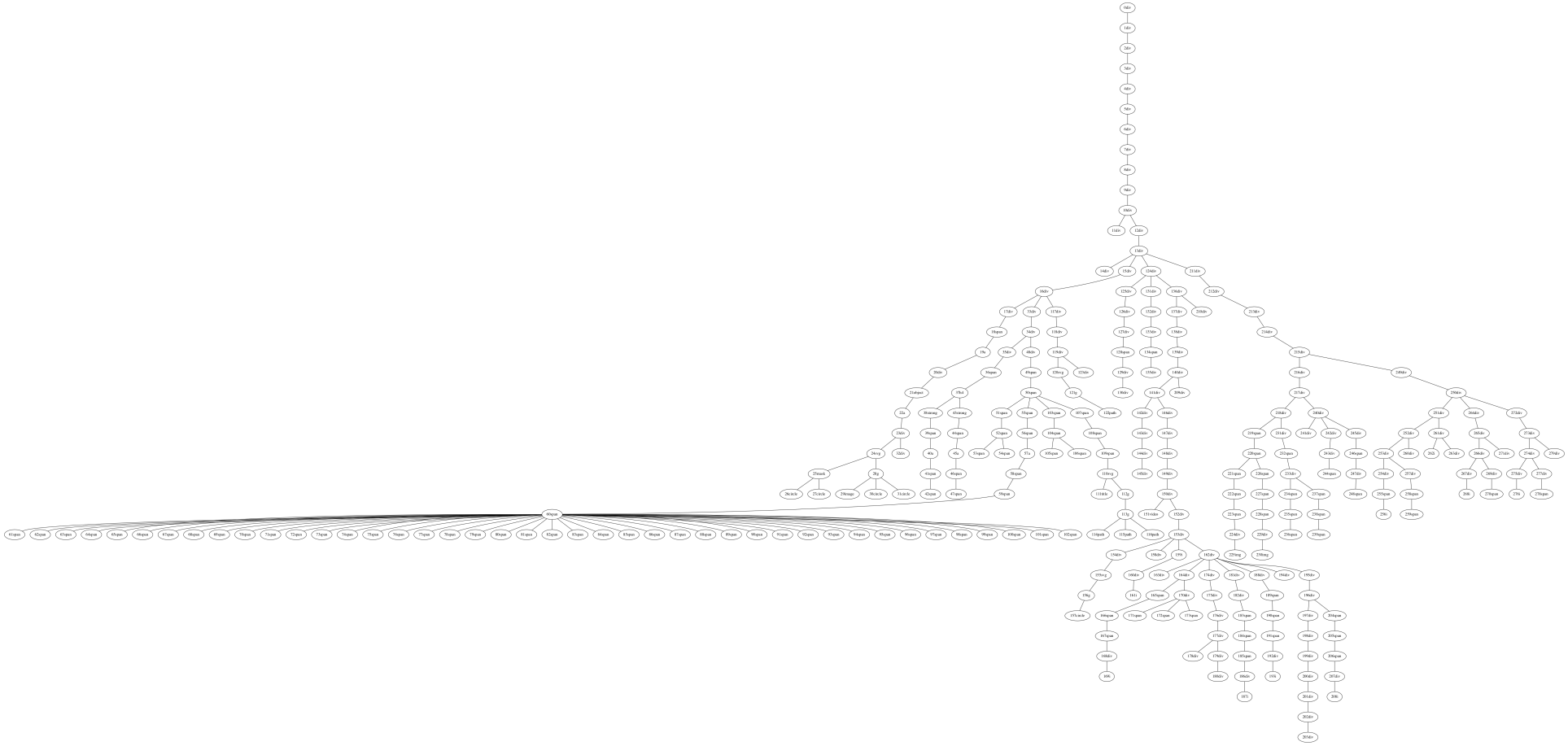

Figure 6: Visualization of a single FeedUnit, i.e. a single Facebook post

FACEBOOK USAGE

1. How long have you had a Facebook account for?
    a. 1 - 3 years
    b. 4 - 6 years
    c. 7 - 9 years
    d. 10 - 12 years
    e. 13 - 15 years

2. How many friends do you have on Facebook?
   Note: To find out the number of friends you have, go to your Facebook profile page, and look under the Friends section.
    a. user text input

3. On average, how many hours do you spend on Facebook per day?
    a. 0 - 2 hours
    b. 3 - 4 hours
    c. 5 - 6 hours
    d. 7 - 8 hours
    e. 9 - 10 hours
    f. More than 10 hours

TRUST, PRIVACY, AND SECURITY

1. To which extent do you trust Facebook not to share your private data with third parties?

   1 2 3 4 5

   Complete distrust Complete trust

2. Have you ever **stopped using** your Facebook account due to lack of trust in the way Facebook handles your private data?
   a. Yes, and I never used it again
   b. Yes, for a while, but then I started using it again
   c. No, but I considered it
   d. No, and I haven't even considered it

SPONSORED AND SUGGESTED CONTENT

Note: The terms "**ad**" and "**suggested content**" are described as follows:

- **Ad** refer to content that the advertisers pay for, in order for it to appear on your feed). Facebook clearly labels such content as "sponsored".
- **Suggested content** refers to content that is identified by Facebook's algorithm as something you might be interested in.

1. When Facebook decides which **ads** and **suggested content** to display in your feed, do you think this decision is made randomly, or based on your user history and personal information?
    a. Randomly
    b. Based on user history and personal information
    c. A mix of both

2. Specify the degree to which you find **ads** annoying?

1	2	3	4	5
Not annoying at all				Extremely annoying

3. Specify the degree to which you find **suggested content** annoying?

1	2	3	4	5
Not annoying at all				Extremely annoying

4. Specify the number of times you made a purchase based on an **ad**?
   [INSERT NUMBER HERE]

5. How often are you able to recognize whether a certain content is an **ad**?

1	2	3	4	5
Never				All the time

6. How often are you able to recognize whether a certain content is a **suggested content**?

1	2	3	4	5
Never				All the time

7. Specify the degree to which you agree with the statement that **ads** and/or **suggested content** on Facebook aim to influence your social and/or political views?

1	2	3	4	5
Completely disagree				Completely Agree

8. Specify the degree to which you agree with the statement that **ads** take up too much of your newsfeed?

1	2	3	4	5
Completely disagree				Completely Agree

9. How frequently do you encounter **ads** that you would consider to be inappropriate?

1	2	3	4	5
Never				extremely frequently

10. Have you ever installed an **ad** blocker for Facebook?
    a. Yes
    b. No

11. How effective do you think **ad** blockers are on Facebook?

1	2	3	4	5
Not effective at all				Extremely effective

12. Do you think **ad** blockers collect your private data, and if so, do you trust the way they handle it?
    a. Yes, they **collect** private data, and I **trust** the way they handle it
    b. Yes, they **collect** private data, and I **do not trust** the way they handle it
    c. No, they **do not collect** private data

13. Specify the degree to which you agree with the statement that removing **ads** enhances your user experience?

1	2	3	4	5
Completely disagree				Completely Agree

TOPIC-BASED CONTENT

1. Have you ever reported content you did not like on Facebook?
    a. Yes
    b. No

2. On a scale of 1 to 5, to what degree was your Facebook experience improved by reporting content?

1	2	3	4	5
Improved to great extent				Did not improve at all

3. Specify the degree to which your Facebook experience is hampered by posts on topics you don't want to see?

1	2	3	4	5
Not at all				Extremely

4. On a scale of 1 to 5, how likely are you to use a topic-based filter to filter posts on Facebook?

1	2	3	4	5
Extremely Unlikely				Extremely Likely

5. Specify the degree to which you agree with the statement that removing posts based on a topic enhances your user experience?

1	2	3	4	5
Completely disagree				Completely Agree