DEPARTMENT OF INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Seminar "Data Analytics for Cybercrime and Undesirable Online Behaviors"

Investigating Pegasus' Impact on the Online Activity of Journalists Targeted with Illegal Surveillance – an open-source data analysis.

Kumaravel Rajan - 03738617

Daniel Bücheler - 03695926

DEPARTMENT OF INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Seminar "Data Analytics for Cybercrime and Undesirable Online Behaviors"

Investigating Pegasus' Impact on the Online Activity of Journalists Targeted with Illegal Surveillance – an open-source data analysis.

Author:	Kumaravel Rajan - 03738617
	Daniel Bücheler - 03695926
Supervisor:	Prof. Dr. Jens Großklags
Submission Date:	13th March, 2022

I confirm that this seminar ”data analytics for cybercrime and undesirable online behaviors” is my own work and I have documented all sources and material used.

Munich, 13th March, 2022

Kumaravel Rajan - 03738617

Daniel Bücheler - 03695926

Contents

1. Introduction

Pegasus is a smartphone spyware developed by the Israeli technological firm NSO group [1]. It was first discovered in August 2016 when Emirati activist Ahmed Mansoor received a message promising "secrets" regarding torture happening in the prisons of the United Arab Emirates by following the given link [2]. This alerted Mansoor who forwarded the message to The Citizen Lab of the University of Toronto [2].

Upon investigation The Citizen Lab described it as the "most sophisticated cyber weapon ever discovered" [2]. It was found that three previously unknown zero-day vulnerabilities in iOS were exploited upon clicking the link [3]. If Mansoor had clicked on the link, all the data on his phone including text messages, photos, emails and location data would have been transmitted to the attacker [2]. To an activist who regularly speaks to sources on the condition of anonymity this would have been devastating.

The official stand of NSO group is that it provides "authorized governments with technology that helps them combat terror and crime." NSO also confirmed that "The company sells only to authorized governmental agencies" [4]. In spite of these claims Pegasus has repeatedly been found to be used by the client governments of the NSO group to repress dissident voices such as journalists and social activists [5].

Following this, in 2019 one of the biggest messaging platforms WhatsApp reported that Pegasus had exploited a vulnerability in the application to launch zero-click attacks. Using this vulnerability Pegasus could be installed on the target device by simply calling it on WhatsApp. The attack was successful even when the call was not answered. The success of Pegasus lies in the developer team's ability over the years to keep finding zero-day vulnerabilities to facilitate remote installation of the malware on their victim's devices, even as more and more vulnerabilities became known and were fixed by Apple.

In 2020, around 50,000 phone numbers believed to have been victims of Pegasus surveillance were leaked to the French NGO Forbidden Stories and the UK NGO Amnesty International. A team of journalists belonging to 17 media organizations spread worldwide took part in the analysis of the 50,000 phone numbers. It was given the name "The Pegasus Project" [6]. On July 18, 2021 The Pegasus Project began publishing results of the analysis. It was found that the list contained opposition politicians, activists and around 200 journalists from nearly two dozen countries [8, 7] along with 14 head of states [9]. The Guardian named 38 journalists in Morocco, 48 journalists in Azerbaijan, 12 journalists in the United Arab Emirates and 38 journalists in India as victims of Pegasus [10].

As part of the revelations, Forbidden Stories posted detailed profiles of 44 victims of Pegasus on their website [11]. Our project is focused on these 44 individuals. We aim to visualize how Pegasus victims reacted to the fact that they were being spied upon by analyzing their online activity.

2. Related Work

The title „The Pegasus Project" describes a coordinated investigative effort by 17 news organizations and over 80 journalists [12, Q&A Box *What is the Pegasus Project?*]. As such, this investigation has produced a large number of reports covering different aspects of the topic and focusing on different clusters of targeted journalists (e.g. in [14, 15, 13]).

Forensic investigation of phones on the list of purportedly targeted numbers was conducted by Amnesty International's Security Lab and The Citizen Lab (University of Toronto). Amnesty published an extensive report on their forensic methodology [16], which was peer-reviewed by The Citizen Lab [17]. Of 67 forenscially investigated phones, 23 were successfully infected and another 14 showed signs of attempted infection [18].

Marczak and Scott-Railton of The Citizen Lab have also published several reports with new findings about Pegasus targetings and NSO's exploit infrastructure. In 2016, Pegasus was discovered for the first time after a failed attempt to infect Saudi-Arabian activist Ahmed Mansoor's phone [19]. Between 2017 and 2019, Marczak, Scott-Railton *et.al.* published an extensive series on targeting of politicians, journalists and activists in Mexico [20, 21, 22, 23, 26, 25, 24, 27]. From 2016 to 2018, they scanned the entire Internet to reveal the Pegasus attack infrastructure on a global scale [28]. The results from this analysis are used to identify possible attack vectors (e.g. suspicious messages) as belonging to Pegasus, e.g. by comparing links in the message with domain names known to be part of the NSO infrastructure.

In 2020, they reported on the first instance of a *zero-click* infection, i.e. completely in the background without any user interaction. Previously, only exploits that require some action by the user (e.g. clicking a link) had been known. In many instances, infection attempts were only detected because victims were suspicious of messages they received. For example, in early 2020 they reported on a (probably failed) infection attempt against the *New York Times*' Beirut Bureau Chief Ben Hubbard [29], who was suspicious of text messages he received. Less than two years later, they report that Hubbard was successfully infected using a zero-click vulnerability [30].

The Citizen Lab is not only reporting on Pegasus malware, but also on infections with other commercial malware [32, 31]. Some journalists were even infected with two different spywares by two different government actors at the same time [30].

Woodhams describes the increasing targeting of independent journalists with surveillance and lists some of the possible consequences [33]: For one, the information collected from the victim's phone regarding ongoing investigations and sources can be used to interfere with their work, e.g. by bringing up charges against and / or arresting the individual [33]. It may also discourage potential sources from contacting journalists due to fear of their identity becoming known not only to the journalist, but also to the Pegasus operator [33].

There are at least two cases where the Pegasus spyware is directly associated with the

murder of a journalist: Only two days after the murder of Mexican journalist Javier Valdez Cárdenas, his wife and colleagues were targeted with malicious text messages designed to infect them with Pegasus [33, 29]. Similarly, the killing of Saudi-Arabian journalist Jamal Khashoggi at the Saudi consulate in Istanbul coincides with his friend and confidant Omar Abdulaziz' targeting with Pegasus: Just the day before Khashoggi's murder, The Citizen Lab published a report on Abdulaziz' targeting [29] and Abdulaziz filed a lawsuit against NSO group [33, 34]. Khashoggi and Abdulaziz had been communicating for several months beforehand using an encrypted messaging app [33, 29], which would have been compromised by Pegasus.

In addition to facilitating detainment or physical violence, surveillance leads to self-censorship when journalists choose not to report on certain topics or not to focus on some issues out of fear for surveillance and repercussions [33]. This „terrorizing effect" [33] is created by the mere presence of tools like Pegasus in the hands of oppressive regimes and does not require the direct targeting of the respective individual.

The Digital Violence Platform ([35], background info at [36]) provides an interactive visualization of digital targeting, physical targeting and contextual events. It allows the user to navigate a 3D-space visualizing timelines of events for different journalists, including incidents such as failed or successfull Pegasus infections, intimidation, arrest, murder and more. The platform also offers video investigations to showcase the stories of particular targets. It also has an interactive graph of NSO group's corporate network, shedding light on the complex network of corporate affiliations that finances the spyware [36].

3. Data collection and methodology

As stated earlier, the basis of our analysis is a list of journalists as mentioned on the website of Forbidden Stories who were targeted by the Pegasus malware [37]. The list consists of 44 journalists from 11 different countries.

For each journalist, there is a short profile page with further information about that individual's work and their targeting with Pegasus. Specifically, the timeline of their selection for targeting, confirmed targeting (e.g. suspicious messages) and confirmed infections are described. In addition, the profile pages also give a brief overview of important investigations conducted by the respective journalists and previous or ongoing harassment, prison sentences, exile etc. these journalists have had to face.

The data collection process described below was conducted for all 44 affected journalists listed by Forbidden Stories. The journalist profiles were very helpful for the analysis, as it presented a starting point and it usually pointed to some article or post by the individual.

3.1. Data sources and scraping tools

For each journalist, we collected activity data from three sources: *Twitter*, *Facebook* and their main media outlet(s). The latter was usually a newspaper website, but could also be a blog or some other type of dedicated online presence. This type of data will be referred to as *Publishing House* or *News Outlet*.

3.1.1. Parsing Twitter and Facebook

Parsing the Twitter and Facebook data was facilitated via the use of python web scraping libraries. Since Twitter is a much more public medium by design than Facebook, parsing Twitter data was comparatively easier.

The Twitter data for the Pegasus victims was parsed using `snscrape` [38]. The Forbidden Stories list facilitated finding the Twitter usernames of victims which were then passed on to snscrape to retrieve all posts made by the victim.

Retrieving the Facebook data was trickier. Facebook does not allow a free hand when it comes to data parsing. The python library we used, `facebook-scraper` [39], required either of our Facebook login cookie in order to parse the given journalist's page correctly. Scraping more than a predefined amount of data per minute often lead to the corresponding user account being blocked. Custom exception handling code was necessary to completely automate this process.

3.1.2. Parsing publishing houses

Collecting the data from each journalist's publishing house was most tedious, as newspaper websites and blogs don't provide a unified interface or even API for us to access. Finding the relevant publishing house website was easy for most journalists, as important articles were linked on each journalist's Forbidden Stories profile page. Usually, we could already see that an individual had worked at one or sometimes two newspapers in the past.

Most newspaper websites provide „author pages", i.e. an overview page of all articles written by a specific person. After finding the URL to this page, we'd use Python's Requests library [40] to load pages and Beautiful Soup [41] to extract data from the raw HTML. Some pages had special anti-bot measures (e.g. provided by Cloudflare) and required use of the Cloudscraper library [42] to retreive the raw HTML. To collect more than ten or 20 articles, it was usually necessary to reverse-engineer the site's pagination URL scheme (e.g. an author page would be at `http://example.com/author/abc` and the second page of that author's articles would be at `http://example.com/author/abc?page=2`).

Some newspaper websites use AJAX technology to dynamically load additional content without reloading the page. In these cases, we identified the relevant end point and necessary parameters using a web browser's development tools and queried the AJAX API directly. In some cases, this was significantly easier than working with the RAW page HTML.

The number of news websites where these author-specific pages don't function properly is astonishingly high. Some newspaper websites, on the other hand, don't provide such pages at all, and thus could not be parsed in this manner. In a few cases (for very small news websites), it was instead feasible to simply parse the entire news site.

Where possible, we collected the following data points for each article: Title, date of publication, URL, author and word count.

3.1.3. Counting number of times the keyword "Pegasus" was mentioned

Once we had the online activity data from all the data sources (social media platforms and publishing house(s)) we searched all collected posts and articles for mentions of the keyword "Pegasus". The approach here being that the impact on a journalist's life of finding out that they could be a possible Pegasus victim would be directly proportional to the number of times the keyword "Pegasus" is mentioned in their online posts. As we analyzed activities of journalists from eleven countries, we deliberately did not include terms such as „surveillance" or „spyware" in this keyword search, because those would be different for each language. The „Pegasus" keyword, on the other hand, is a proper name and thus the same across all languages.

3.2. Data analysis and plotting with Pandas

Data analysis was conducted using the Python-based data analysis tool Pandas [43]. The results from data collection were stored in three JSON files per journalist (one each for Twitter, Facebook and Publishing House activity). If some of the data is unavailable for some

journalists (e.g. if they don't have a Twitter account), the analysis of that part is simply omitted.

During analysis, the three JSON files are loaded into Pandas DataFrame objects, re-indexed according to the date of publication, and grouped by month. For each group, the data is then aggregated: For Facebook and Twitter, only the number of tweets / posts per month is counted. For the publishing house data, the number of articles is counted and their average length for each month is calculated.

The three aggregated datasets are then joined into a single dataframe and plotted using Matplotlib [44]. Additionally, the twelve-month moving average of the overall activity (i.e. total number of articles, facebook posts and tweets) is calculated to smooth out seasonal variations and provide a measure of the overall trend (increasing / decreasing activity). For visualization of the activity counts, we chose a stacked bar chart with individual colors for Facebook, Twitter and Publishing House activity. The twelve-month moving average is plotted as a dashed line on the left axis, the publishing house length data (if available) as a red line on the right axis. The number of times "Pegasus" was mentioned in posts is shown as a simple bar graph in red on the left axis.

A sample generated graph is shown in Figure 4.7.

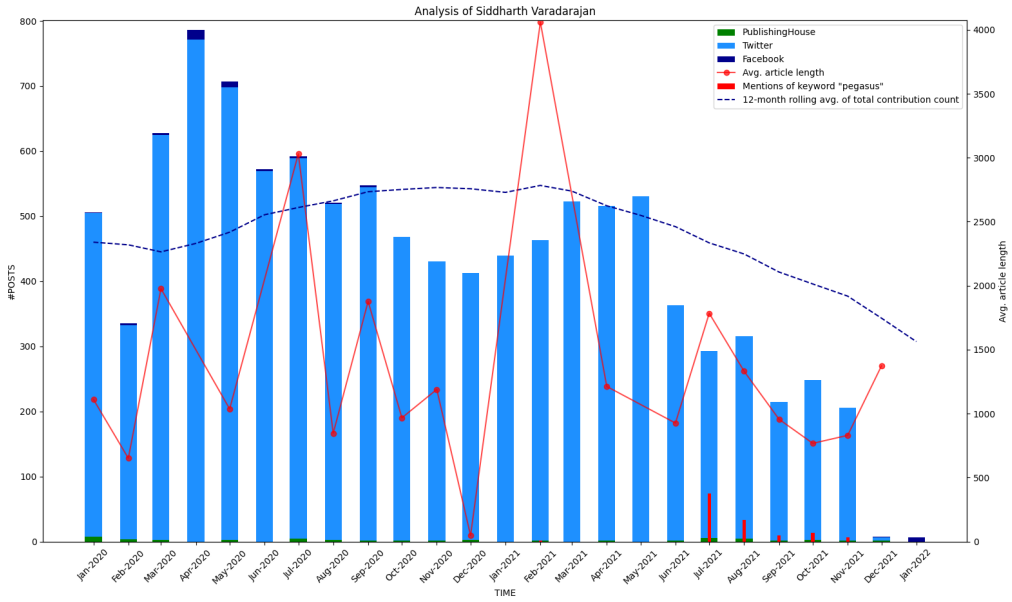

Figure 3.1.: A sample generated graph - Publication statistics for Siddharth Varadarajan, 2020-2021

3.3. Analysis of connections among journalists

In the course of the project work, we noticed certain patterns emerge from the 44 journalists listed on Forbidden Stories. There were multiple cases where starting from a Pegasus infected journalist we could trace their colleagues / ex-colleagues / family members who had also been targeted with Pegasus. These connections were traced by manually going through the profile pages of different journalists on Forbidden Stories and noting down the connections. To our surprise, we found enough material there to make this a secondary part of our seminar. The results of this analysis are shown in 4.2.

3.4. Issues and failed approaches

Availability of data for some journalists

As already mentioned above, not all data points could be collected for all journalists. One reason was that the resources we parsed did not exist for some journalists or we could not find them (e.g. no Twitter account). Some of the journalists featured by Forbidden Stories work for TV stations, and thus there are no articles on a newspaper website for us to collect.

In other cases, scraping of publishing house data failed due to broken websites. Especially the pagination system was broken in many sites (i.e. we could parse the latest ten or twelve articles by a specific author, but no more); other sites did not offer author-specific pages at all. Another newspaper (for Indian journalist Jaspal Heran) publishes partly in Punjabi, a language spoken in Pakistan and India [45]. As the website does not load the articles as text, but as image files (!), optical character recognition (OCR) of Punjabi text would be necessary to automatically process this newspaper's articles. While this would certainly be an interesting project, it is significantly out of scope of the seminar.

Sentiment Analysis

In addition to a metadata analysis of publications (i.e. frequency and length), we initially planned to conduct a sentiment analysis on the contents of tweets, posts and articles. We hoped to find out if being subjected to surveillance has an impact on the journalists' emotions and style of writing.

As visualized in Section 4.2, we are dealing with journalists from 11 countries publishing content in many different languages. To cover at least the majority of languages, we would need to perform sentiment analysis in English (Indian journalists), Spanish (Mexican journalists), Arabic and French (both for Moroccan journalists). While there are many libraries for English (e.g. TextBlob [46]) and some for Spanish (e.g. [47]), we expect this task to be extremely difficult for Arabic and less common languages like Azerbaijani or Hungarian.

As newspaper articles and, to some extent, Tweets and Facebook posts on official profiles are explicitly written and edited with a target audience in mind, we are also unsure about the insights to be gained. For this reason, we have decided to omit the sentiment analysis and instead focus on collecting more metadata for a larger number of journalists (initially,

we planned to do the full data collection of publishing house & facebook data only for a few journalists whose Twitter data seemed promising).

4. Results

This chapter will go through the results of the publication analysis for a number of interesting cases. Towards the end of the chapter, we also present the graphs representing the connections between journalists in each country.

4.1. Analysis of online activity of selected journalists

4.1.1. Carlos Ketohou

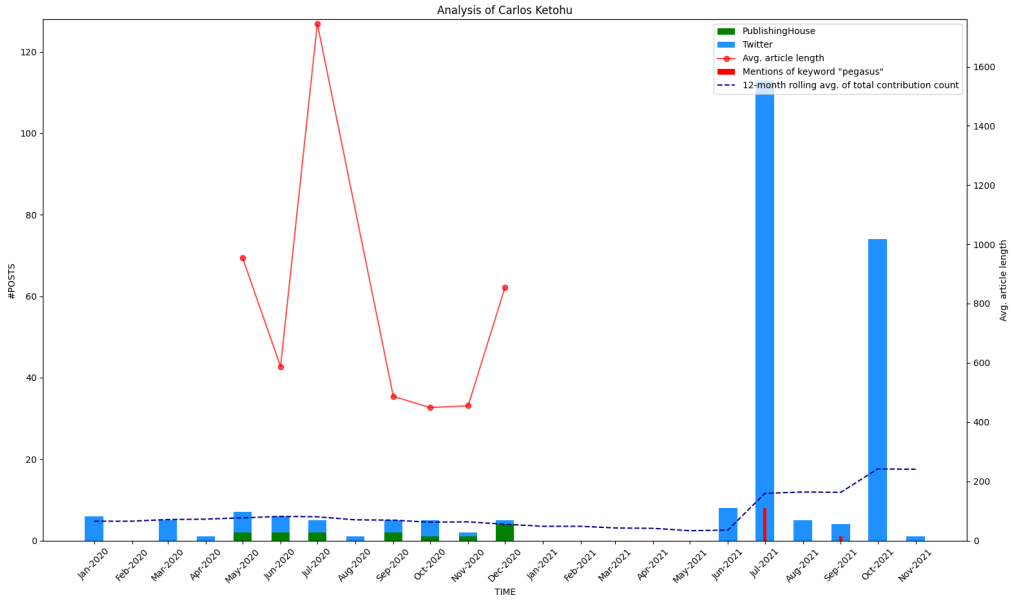

Figure 4.1.: Publication statistic for Carlos Ketohou, 2020-2021

Carlos Komlanvi Ketohou is a Togolese journalist who mostly reports on political issues and is „a vocal defender of the freedom of press in Togo and frequently denounces attacks against it". [48]. He was the founder and director of the independent magazine „L'Indépendant Express" from 2007 until the Togolese authorities closed the paper in 2021 [48]. He was targeted with Pegasus in 2017 and 2018, but his phone could not be analyzed to confirm the infection [48]. In December 2020, he was arrested for „defaming the government" following a report accusing Togolese ministers of stealing golden spoons during a government

reception [48]. He was forced to reveal the passwords to his email and social media accounts during the interrogation about the article and his sources [48].

The graph in figure 4.1 shows very clearly the duration of his prison sentence: While he wasn't exactly a heavy Twitter user in the months leading up to his arrest, he continuously published a small amount of tweets. The publication data from L'Indépendant Express also shows that he is publishing quite regularly, even though our data might not be exhaustive at this point: Direct parsing of author-specific pages was not possible for L'Indépendant Express because all articles were posted under a generic „author account" and the author of each individual article was simply added to the text. This is why our scraping tool might not have detected all of Carlos Ketohou's articles as coming from him.

Of course, also his publications on L'Indépendant Express's website stopped during his prison sentence. However, in June, his Twitter use is already slightly higher than in the months before his arrest (even though the first Tweet since December 2020 occurs on June 21st). In July he posted a staggering number of tweets, amounting to a more-than-tenfold increase compared to his previous highest number of tweets in a month. Even more strikingly, this is immediately followed by a drop to levels seen before in August and September 2021, and up again in October.

His Twitter use during this time may not be causally related to the Pegasus Project revelations in late July 2021, but the correlation is certainly interesting: The data suggests that, as his publishing house has been closed, he turns to Twitter as a means of communications in late June and throughout July 2021. As he does not seem to have been involved in the Pegasus Project investigations, he would most likely only have found out about his targeting with the revelations in late July. Throughout the next two months, his Twitter activity is again very low, at about the same levels as before his arrest (when he was, unlike now, still publishing articles in L'Indépendant Express). The reason for this drop in Twitter activity is unclear, but it is possible that he was taken aback by finding out that he was targeted with a digital surveillance trojan and, as a result, showed reduced activity (e.g. due to indimidation, taking measures to protect himself etc.). Ketohou did not respond to our request for his comment on this subject.

4.1.2. Lenaïg Bredoux

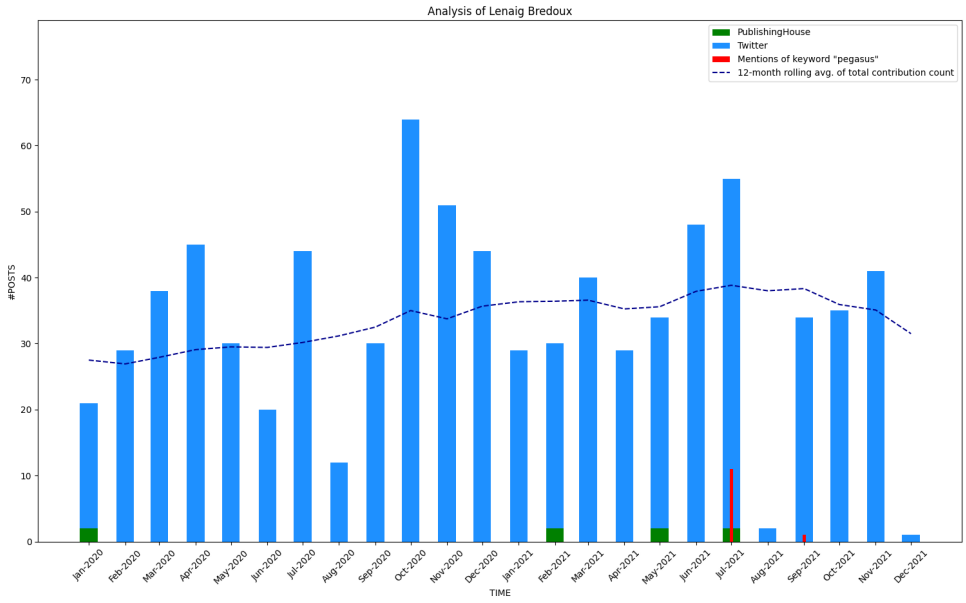

Figure 4.2.: Publication statistics for Lenaïg Bredoux, 2020-2021

Lenaïg Bredoux is a French journalist who works at the investigative news site Mediapart [49], whose founder and director Edwy Plenel was also targeted [50]. She was targeted with Pegasus in 2019 and 2020 [49]. A forensic analysis by Amnesty International's Security Lab confirmed the infection and found that a vulnerability in iMessage was used to gain access to her phone [49, 1].

In 2015, Bredoux reported on diplomatic tensions between France and Morocco [49], which is most likely how she ended up getting targeted with Pegasus. She has also written several investigative articles on sexual harassment and sexual assault allegations since 2016 [49].

Her activity graph for the last two years shows that she has been using Twitter quite constantly, but with varying intensity: From just over 10 tweets in August 2020 to more than 60 in October 2020 (when she was appointed as Mediapart's „gender editor" [49]). In August 2021, her tweet count dropped to a level not seen since mid-2019, when she did not post a single tweet in one month.

Again, this is exactly correlated with the revelations about the Pegasus malware in late July 2021. However, it is likely that Bredoux knew that she was under surveillance some time before the revelations: She already tweeted extensively about Pegasus during the days of the Pegasus Project's publication and also wrote that „The @FbdnStories consortium warned us several weeks ago that @edwyplenel [...] and myself have been targeted by the Pegasus software [...]" (translated from [51]). Her large number of Tweets in July 2021 can be explained by the extensive coverage of Pegasus, posting Mediapart articles, retweeting and participating

in discussions about the topic. Even though the relatively low number of mentions of the „pegasus" keyword (narrow red bar in the graph) suggests that only a small part of her twitter activity was related to Pegasus, manual inspection shows that most of her tweets during that time were. However, many tweets don't mention the keyword „pegasus" directly but talk about surveillance in general, or simply reply to an ongoing discussion.

Especially the very low number of tweets in August 2021 is striking, even considering that her twitter activity in August of the previous year was much lower than average (looking even further back, August has a relatively small number of tweets in most years). In August 2021, she only posted on Twitter twice, whereas she posted twelve times in August 2020 and 30-40 times per month on average.

Her rolling twelve-month average of posts also shows an interesting behavior around the time of the Project Pegasus revelations: It had been steadily increasing for the entire duration of the plot until July 2021 (in fact, this development goes back to 2018, when her Twitter use was lowest). After July, her activity starts to decrease notably (even though her Twitter use picks up again after almost no activity in August).

Both of these patterns – the extremely low activity in August 2021 and the decrease of the 12-month rolling average after July 2021 – don't have to be connected to the Pegasus Project revelations, but it is certainly striking that both of these anomalies correlate with the Pegasus findings.

Responding to our request for her comments, Bredoux responded (translated from French): *Hello Thank you for your message and your interest. I will disappoint you: the very significant drop in August is due to four weeks of vacation. It is true, however, that the Pegasus revelations and the legal consequences in France had exhausted me and that I had further cut off all social networks during this period. Have a good evening*

4.1.3. Sushant Singh

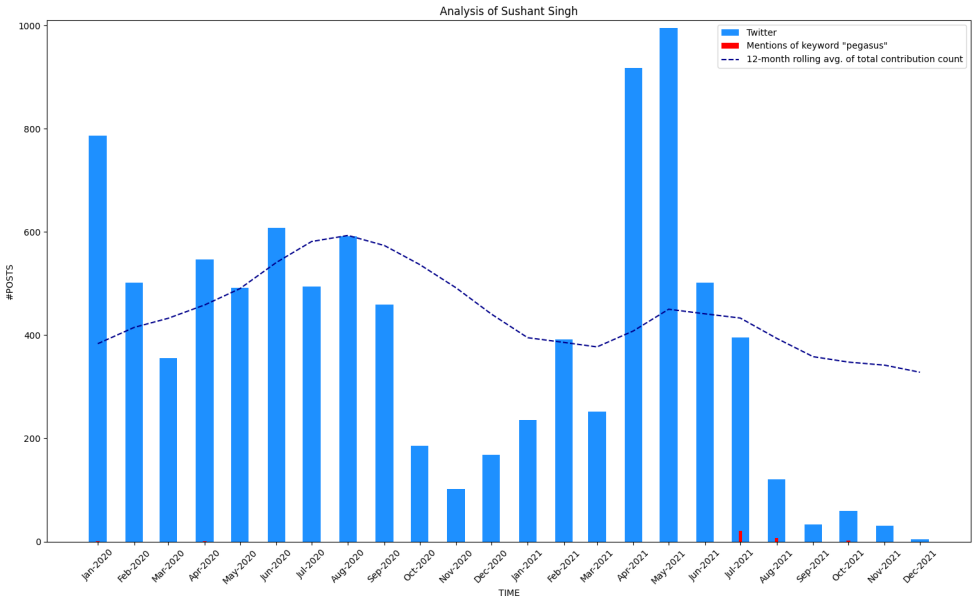

Figure 4.3.: Publication statistics for Sushant Singh, 2020-2021

Sushant Singh is an Indian journalist and international affairs analyst [52]. He was infected with Pegasus in 2021 and his phone remained under surveillance up until shortly before the Pegasus Project published its findings [52]. The infection with Pegasus was confirmed by forensic Analysis of his phone [1]. When analyzed, the phone was running the most recent version of iOS [52], thus it is very likely that a Zero-Day vulnerability in iOS was exploited to install Pegasus on his phone.

Before working as a journalist, Singh was part of the Indian military and served as a UN peacekeeper [52]. He does not work for a single news outlet, but publishes articles and analyses in multiple papers including *The Caravan*, *Scroll.in* and *Foreign Affairs* [52]. The total number of articles he writes is very small and even these articles are spread out among many publication's websites. For this reason, analyzing his publishing house contributions was not possible.

However, he is a heavy Twitter user, posting up to almost a thousand tweets per month. His activity was on a high level before October 2020, when it dropped for a few months (to still 100-200 tweets per month), and then increased again. In April and May 2021, his post count reached a peak before it dropped very rapidly. In June and July, the post number was still within the 12-month average and the levels seen in the months before. Singh did mention the keyword „pegasus" in a fair number of tweets in July, but these do not account for a major share of his activity.

However, from August until November the post count reached unprecedented low. The

parsing of Twitter data was done in mid December 2021 and thus the activity data of this month isn't complete. It was already lower than average at this time in 2020, but the level seen in 2021 is even far below that. This can also be seen in the 12-month average, which is decreasing noticably during this period.

As with Lenaïg Bredoux, the strong reduction in Twitter activity begins at the same time as the Pegasus revelations in late July. However, in his case the decrease isn't as significant as with Bredoux, and his Twitter activity varied more strongly in the year before (c.f. the 12-month rolling average between August 2020 and Jan 2021). If the activity reduction was caused by the Pegasus Project findings, the impact was thus not as strong for him as for Bredoux.

4.1.4. Ferdinand Ayité

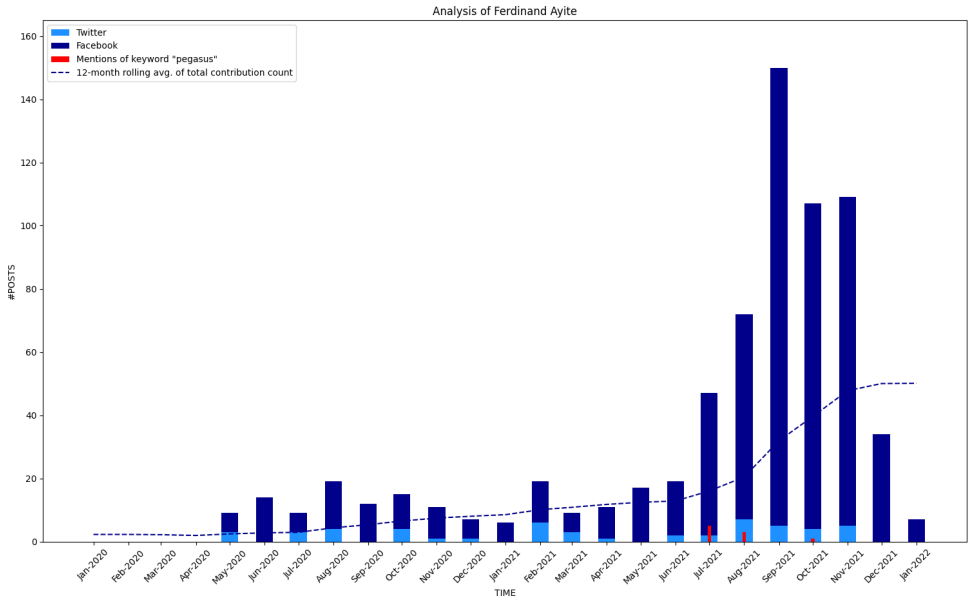

Figure 4.4.: Publication statistics for Ferdinand Ayité, 2020-2021

Ayité heads the Togolese investigative bi-monthly newspaper L'Alternative [53]. His newspaper is known to be critical of the ruling Gnassingbé family which has led the country for 54 years [53].

Ayité was targeted with Pegasus spyware in 2018 [53]. It is interesting to note that in the same year he alleged the electoral commission in Togo of corruption [53]. Ayité also is active in Nubuéké, a citizen movement which has been repressed by authorities [53]. He also said that in 2017 and 2018 as retaliation to him covering protests seeking the resignation of President Faure Gnassingbé, electricity supply to his newspaper office was cut several times [53].

More recently in March 2020, L'Alternative was suspended for two months upon the French ambassador complaining about an article about French President Macron [53]. A few months later the publication broke a scandal named "Petrol gate" concerning the misappropriation of 764 million Euros by a couple of civil servants [53]. Though Forbidden Stories was unable to analyze Ayité's phone [53], considering the themes of his journalistic activity it isn't a surprise that the government might have been interested in tracking him with the help of Pegasus.

The activity graph of Ayité shows that his activity increased significantly from July 2021: Earlier, he usually posted less than 20 posts per month in total on Twitter and Facebook. From July to September 2021, his activity has increased massively - from around 60 posts in July 2021 to 80 posts in August 2021 to an all time high of 150 posts in September 2021. It decreased again after that, but still remained at levels higher than before (the post count for January 2022 is not accurate, as our scraping finished in early January).

It seems that Ferdinand Ayité wasn't demotivated or intimidated by finding out he was targeted with Pegasus, but instead even more determined and more active than before. This was also confirmed by his reply to our request for his comment: *Hello sir, Yes quite. The fact of having been a victim of this spyware is a motivation to denounce those responsible. Of course I talked a lot about Pegasus and continue to talk about it ever since On social networks.*

4.1.5. Taoufik Bouachrine

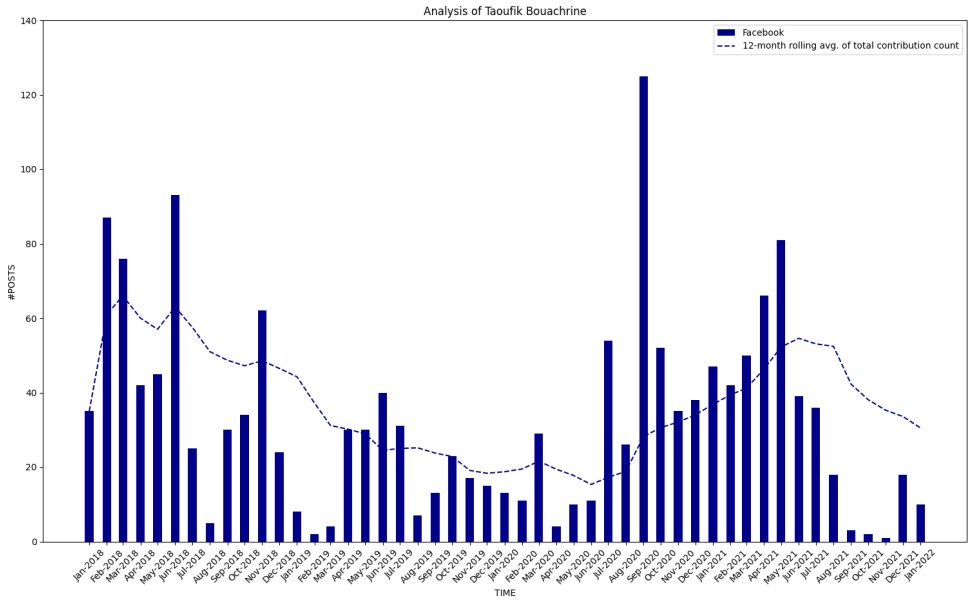

Figure 4.5.: Publication statistics for Taoufik Bouachrine, 2018-2021

Bouachrine is a Moroccan journalist and editor at Akhbar al-Yaoum, an Arabic daily that he founded in 2007 [54]. It had a reputation of being one of the last remaining newspapers critical

of the regime. In it Bouachrine regularly criticized the Makhzen (authorities) in Morocco [54]. He is currently serving a 15-year prison sentence on counts of "rape and attempted rape", “abuse of power for sexual purposes” and “human trafficking” [54]. He was targeted with Pegasus from 2017 to 2018 and his wife Asmae Bouachrine was targeted from 2017 to 2019 [54].

It was in the same time frame when he is estimated to have been targeted with Pegasus that he was arrested in February 2018 at his newspaper’s office in Casablanca [54]. The court proceedings brought to light the fact that Bouachrine’s office had been bugged with cameras and these allegedly showed forced sexual encounters between him and his employees [54]. Several subordinates of Boauchrine appeared as plaintiffs but their trustworthiness took a hit when at least 2 of them denied having been raped by the editor [54].

The graph shown here is the result of parsing Taoufik’s Facebook page which is currently being administered by his wife Asmae Bouachrine. Since most of the posts on the page by his wife dealt with questioning the shaky evidence based on which her husband had been jailed and posting updates about her husband’s state of mind while in prison we felt parsing this facebook page would show the psychological effect of being a victim of illegal surveillance. Correspondingly, we can see from the graph that the number of Facebook posts have been continuously growing starting from July 2020 until June 2021. This observation is facilitated by focusing on the 12-month rolling average metric. However, quite interestingly from July 2021 onwards (the time of Pegasus revelations), the 12 month rolling average sees a continuous decline in the page’s online activity. The graph suggests Asmae Bouachrine might have been reluctant to post online after the Pegasus revelations. However, since she did not respond to our request for her comments we cannot be certain that being named in the Pegasus revelations was the only factor contributing to the reduction in her online activity.

4.1.6. Vijaita Singh

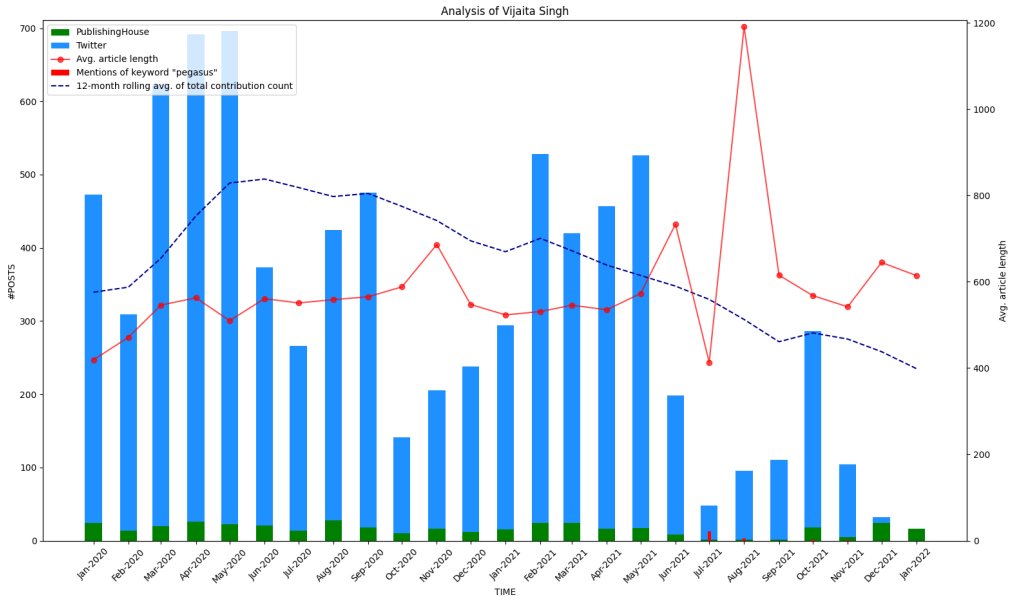

Figure 4.6.: Publication statistics for Vijaita Singh, 2020-2021

Vijaita Singh is a news reporter for The Hindu, a much respected daily in India [55]. She has covered controversial subjects like the activities of Maoists in central India, terrorism and the developments on the Kashmir issue [55]. Her phone was targeted with Pegasus in 2019 [55] - a time period in which she was covering multiple high profile terrorism cases [55]. Amnesty Internaional's security lab conducted a remote analysis on her phone and confirmed an attempted infected via a malicious SMS link in June 2019 [55]. However, it was not possible to find out if the infection was successful [55].

From her graph we can easily spot historically low levels of activity in the months post the Pegasus revelations in July 2021. July, August and September 2021 have seen the lowest number of posts by Vijaita since the beginning of the year 2020. It can also be seen that the keyword "Pegasus" has been mentioned multiple times by Vijaita in months July, August, October and November 2021. This shows that she was well aware that she might be a victim of the spyware and wanted to spread awareness about it to her followers. In the month of August 2021 it can be seen that she has written only one article for the Hindu but its length is close to twice the usual article length written by Vijaita. Upon visiting the said article we found out that it was a rather detailed book review and was not connected to Pegasus. We however cannot be certain that being a Pegasus victim was the only influencing factor in Vijaita reducing her online activity post July 2021 since she did not respond to our request for her comments.

4.1.7. Siddharth Varadarajan

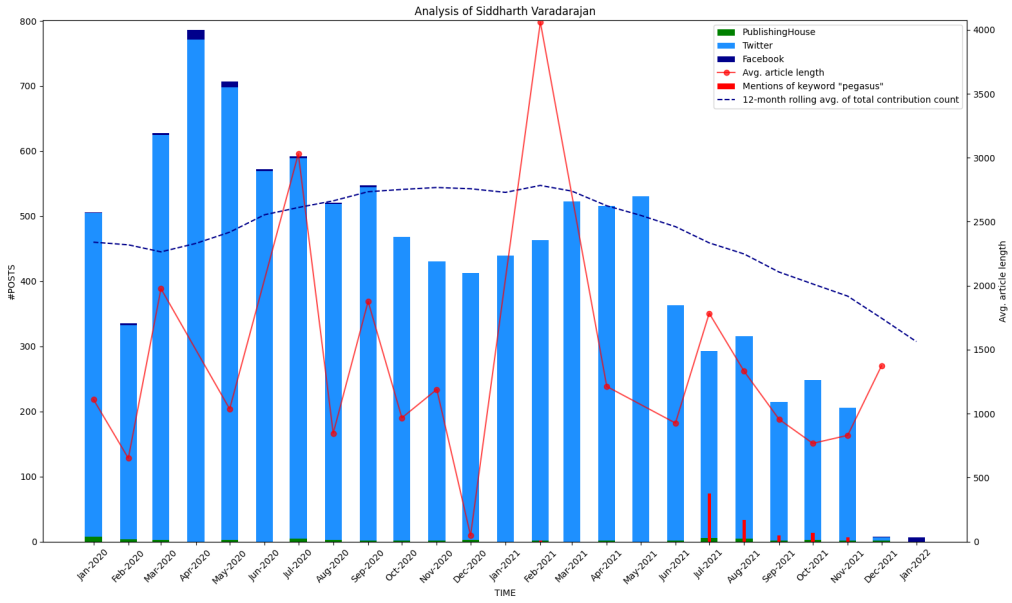

Figure 4.7.: Publication statistics for Siddharth Varadarajan, 2020-2021

Siddharath Varadarajan alongside MK Venu and Sidharth Bhatia founded The Wire in 2015 [56], the publishing house from which multiple journalists were targeted for Pegasus surveillance. In fact, he is among the at least five employees at The Wire who have been subjected to illegal surveillance via Pegasus [56]. Varadarajan and his colleagues in The Wire have broken multiple political scandals in India including the 2G spectrum scam done by the UPA (United Progressive Alliance) government [57] and the controversial purchase of Rafael Jets done jointly by the UPA and the NDA (National Democratic Alliance) governments[58] [56].

Varadarajan was targeted with Pegasus in April 2018 [56]. Analysis of his phone by Forbidden Stories and Amnesty International revelealed that his phone had been successfully infected [56]. Around the same time, on April 14, 2018 The Wire staff broke the coal import scam involving the Indian Government and several MNCs believed to enjoy close links with it [56]. As an investigative journalist he has been subjected to multiple defamation lawsuits [56]. In 2020 for instance, he was summoned to appear in court following an article he published about a religious ceremony that took place despite the imposition of a strict lockdown [56].

Referring to his graph we notice that his online activity has been pretty much constant until February 2021. Every month following February 2021 has seen a decrease in the number of posts when compared to the activity levels starting January 2020 according to the 12 month rolling average. Our hypothesis to explain this drop was that since The Wire was also a member of The Pegasus Project that investigated the 50,000 phone numbers leaked to The

Forbidden Stories, Varadarajan being the founder of The Wire could have been clued in to the results of the investigation much before the official Pegasus Revelations in July 2021. As also seen from the graph, starting from July 2021 Varadarajan has mentioned the keyword "Pegasus" a sizeable number of times in his online posts indicating that he wants to spread awareness about the spyware to his followers. However, since he did not respond to our request for his comments we cannot be certain that being a Pegasus victim alone contributed to the difference in his online behavior.

4.2. Relationships between journalists

The following is a summary of the interesting relationship patterns which we saw emerge upon analysis of the Pegasus victims.

Legend

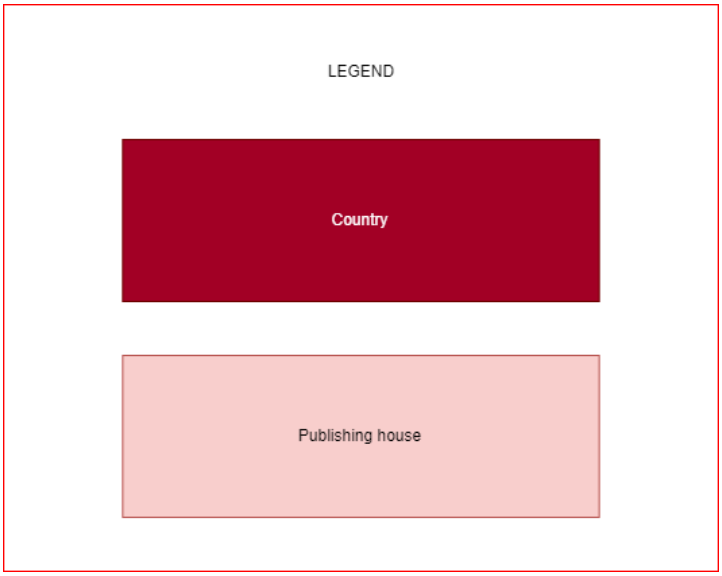

Figure 4.8.: Legend for the following relationship diagrams.

France and Morocco

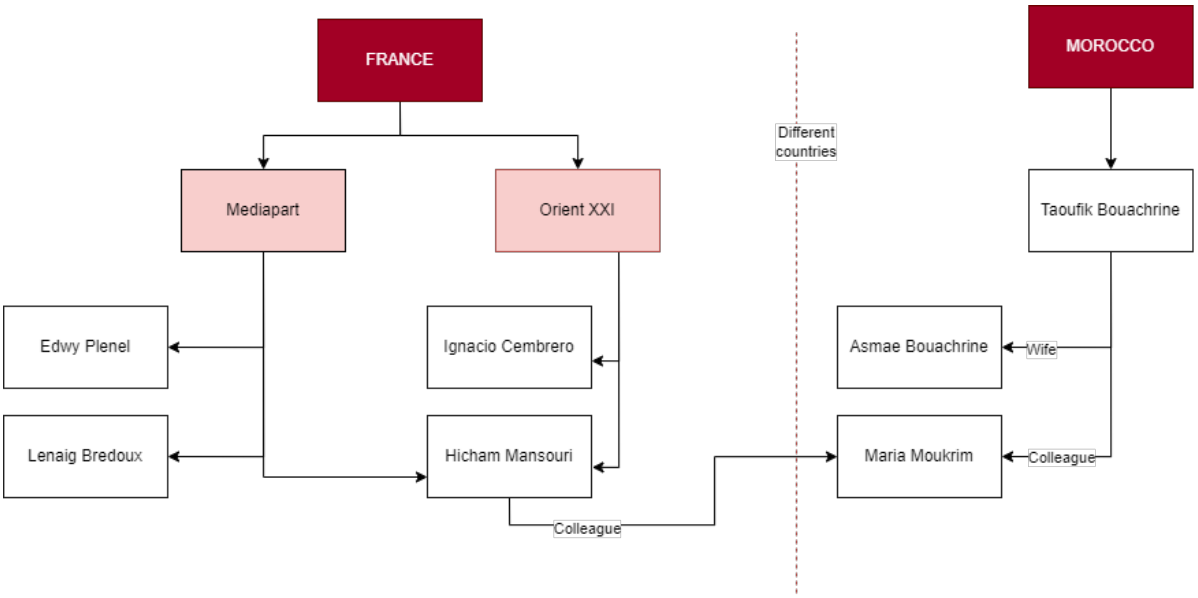

Figure 4.9.: Relations between Pegasus victims in France and Morocco.

In Morocco, Maria Moukrim, a colleague of Taoufik Bouachrine (who was featured in our shortlisted list of journalists above) was targeted with Pegasus from 2017 to 2019. She was the president of the Moroccan Association for Investigative Journalism (AMJI) until July 2014. [59]

In France, Edwy Plenel, the founder and director of the investigative media outlet Mediapart was targeted with Pegasus in 2019. On multiple public forums he has supported human rights defenders in Morocco leading to the suspicion that the Moroccan government might be behind targeting him [60]. Similarly, Plenel's colleague Lenaig Bredoux's (who also featured among our shortlisted journalists above) phone was hacked in 2019 and 2020 [61].

In another French daily, Orient XXI colleagues Hicham Mansouri and Ignacio Cembrero were targeted with Pegasus. Hicham Mansouri's phone was compromised 20 times in 2021 alone. He co-founded AMJI in 2011. Ms. Moukrim was part of the same project. AMJI was being attacked constantly according to him. The homepage of AMJI was reportedly hacked and turned into a pornography site. He also started training citizens in investigative journalism techniques. On facing resistance from the Moroccan authorities to this program, Mansouri fled to France where he currently works with the media houses Mediapart and Oritent XXI [62]. A colleague of Mansouri at Orient XXI, Ignacio Cembrero was selected for surveillance in 2019. He has repeatedly faced the wrath of the Moroccan authorities for being critical about them in his articles. As an example, in the year 2013 the Prime Minister of Morocco filed a complaint against Cembrero accusing him of trying to incite violence when he tagged a video of Al-Qaeda in a blog post [63].

Mexico

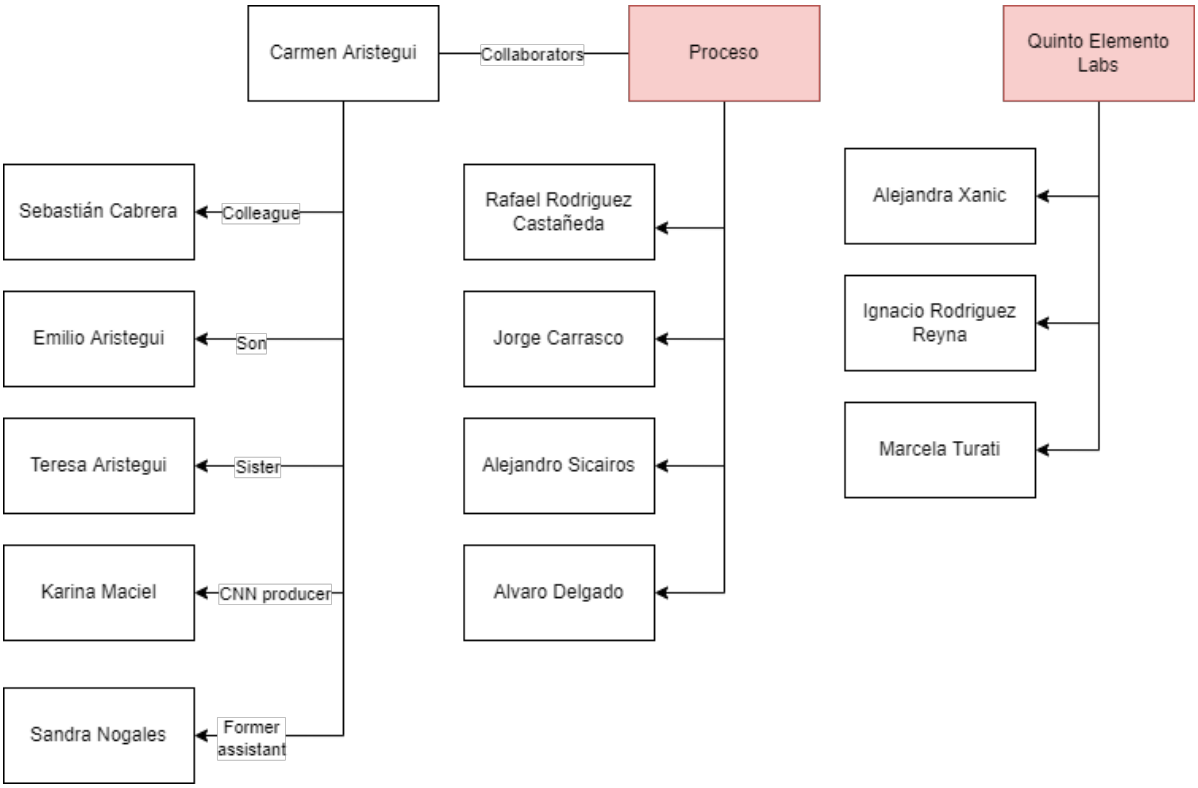

Figure 4.10.: Relations between Pegasus victims in Mexico.

Carmen Aristegui is one of the most recognized faces in investigative journalism in Mexico and currently works with CNN Spanish. She was targeted with Pegasus in 2015 and 2016. She is well known for her critical views on the Mexican government. Along with Aristegui her family members namely her son Emilio Aristegui and her sister Teresa Aristegui, were targeted with Pegasus. Her colleague Sebastian Carbera, CNN producer Karina Maciel and former assistant Sandra Nogales were victims of Pegasus, too.[64].

Employees of Proceso who were frequent collaborators with Aristegui's news outlet Aristegui Noticias also found themselves to be victims of Pegasus. Rafael Rodriguez Castañeda, the director of Proceso until January 2020 was selected for Pegasus surveillance in 2016. Proceso collaborated with Aristegui's Aristegui Noticias to reveal that Enrique Peña Nieto, the then Mexican President had fired a priest so that he could get his first marriage cancelled and get married again religiously [65]. A colleague of Castañeda, Jorge Carrasco was selected for Pegasus surveillance in June, 2016. Alejandro Sicairos who also worked at Proceso was selected for surveillance after his colleague Javier Valdez was murdered. Sicairos heads the newspaper Riodoce in Sinaloa where the cartel operates. He often published articles condemning the lack of repercussions that made threatening and murdering journalists convenient for criminals [66]. Delgado, like his other colleagues at Proceso, was also selected

for surveillance in 2016. This was around the time he had published a book which stated that the two main political rivals in Mexico had had several secret meetings to ensure that the power rotated only between their two parties [67].

Employees of the investigative news outlet Quinto Elemento Lab were also selected for Pegasus surveillance. Alejandra Xanic Von Betrab, a Pulitzer Prize winner for her articles on Walmart bribing Mexican officials, was targeted with Pegasus in 2016. Incidentally, it was in 2016 that she co-founded Quinto Elemento Lab that were behind some of the biggest scandals in Mexico. Marcela Turati who co-founded Quinto Elemento Lab with Betrab in 2016 was targeted with Pegasus in 2016. She previously investigated human rights abuses. The efforts to intimidate her were so grave that she once said, "We were told that if we kept going, we weren't going to make it out alive" [68].

India

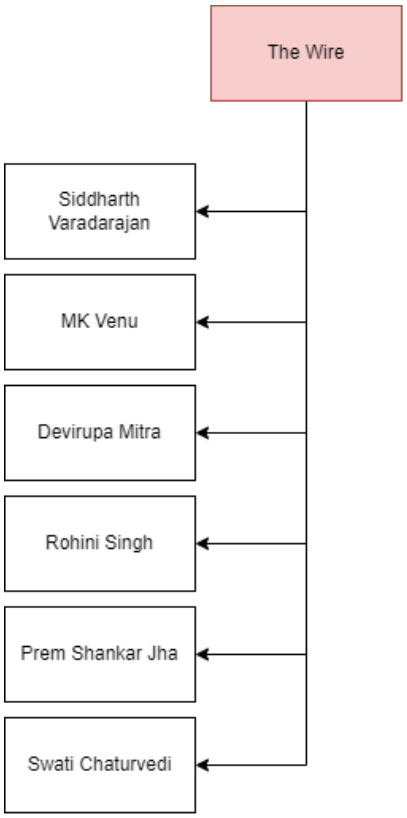

Figure 4.11.: Relations between Pegasus victims in India.

A striking pattern emerges when we look at the journalists targeted in India. More than a handful of them work for the same publishing house - The Wire. Siddharth Varadarajan (who featured on our shortlist of journalists above) was targeted with Pegasus in April, 2018. MK Venu co-founded The Wire along with Varadarajan. He was infected with Pegasus in

July, 2021. The time he was selected for surveillance matches with the time he was primarily covering India's COVID vaccine response. According to Forbidden Stories, Venu was also selected for surveillance in 2018 and 2019 [69]. Rohini Singh, a colleague of Venu was selected for surveillance in 2018 and 2019. She garnered widespread attention when she broke the story that Jay Amit Shah's , the son of Union Home Minister, net worth increased by a factor of 16,000 in the one year following Modi's election in 2014. Swati Chaturvedi was targeted in 2018 and 2019. She is the author of the book "I Am A Troll: Inside the World of the BJP's Digital Army" published in 2018 that detailed out the ruling BJP Party's network of internet trolls that propelled Modi to win a second term in Office. She also investigated the Rafale sale publishing an article about the CBI director Alok Verma getting sacked right when he was preparing to probe the Rafale contract [70].

5. Further research ideas

Due to the limited scope of this seminar project, we were of course unable implement all ideas we had. The JSON files with publications collected for this project provide a solid base for further analysis.

For example, a more in-depth quantitative analysis (e.g. correllation analysis, sign test of publishing behavior before / after revelations) would certainly be interesting. As we spent a large amount of time on collecting the data, especially parsing the publishing houses, we were unable to conduct this analysis due to a lack of time.

Similarly, the detection of posts related to Pegasus could be improved by increasing the list of keywords. Both more general keywords and some language-specific keywords could be used. For example, „NSO" for all languages and „surveillance", „spyware" etc. in the respective language.

The sentiment analysis we had initially planned, would still be interesting. The issue here is the wide variety of languages, some of them very niche, which greatly increases the effort required to complete such an analysis. Still, sentiment of texts about the Pegasus findings and, more generally, the change in sentiment around the time of the findings, could provide valuable insight into how the affected journalists perceived their targeting.

Our dataset could also be combined / merged with the dataset from Digital Violence [35], which includes information on real-world harassment and contextual events. We initially attempted this combination, but it turned out that the intersection of journalists between our dataset and the one from Digital Violence is very small. This would therefore require further data collection on the journalists from the Digital Violence Dataset.

To amend our dataset, it would also be great to include more sources of journalistic activity. Our dataset includes social media activity (Twitter and Facebook) and written publications from online newspapers, blogs etc. Other types of activity, such as on TV, publication of videos / podcasts etc. are not covered by our data collection. This could be implemented in a further project.

Another interesting approach would be to conduct a survey among the targeted individuals whose activities were analyzed in this project in order to understand their perception of the events and their reaction. Especially when it comes to self-censorship, comparing the journalist's actual change in behavior to their perceived change could yield some interesting results.

6. Discussion & Conclusion

As we have seen in the previous sections, the publication activity around the time of the Project Pegasus revelations greatly varies between affected individuals. For many journalists, we were unable to see any clear connection between the Pegasus Project revelations and their publishing activity. This does not mean that there is none, but it is simply not strong enough to be clearly visible among seasonal and other changes.

Many journalists showed decreased activity after the Pegasus project revelations, either temporarily such as Lenaïg Bredoux and Vijaita Singh or permanently (until now, further observation is necessary) such as Sushant Singh and Taoufik Bouachrine. This was the reaction we expected from most journalists, as finding out that they were spied on by a highly sophisticated, stealthy smartphone spyware must be quite disturbing and stressful. In contrast to other oppressive targeting of journalists (such as with lawsuits or being tailed in public), smartphone surveillance targets a very private area of life as phones nowadays no only contain all our data, but are also equipped with cameras and microphones that could be active at any time. We expected a significant number of journalists to be intimidated and stressed by these revelations and show reduced activity. In addition to feeling threatened and intimidated, we also expected some journalists to feel exhausted and thus take a temporary break (e.g. a vacation), posting less in that time.

This is exactly the behavior we've seen with five individuals. Lenaïg Bredoux and one other journalist who asked to stay anonymous confirmed that exhaustion and feeling overwhelmed was the reason for their decreased activity after the revelations.

Some, like Ferdinand Ayité and Carlos Ketohou, show significantly more activity after finding out that they were subjected to surveillance. As for the reason for this behavior, it seems likely that they were not demotivated or intimidated by the findings, but even more convinced of the importance of their own work. Thus they work even more determinedly and in defiance of the authorities trying to suppress free journalism. Ferdinand Ayité confirmed this assumption in his reply to our request for comment, saying that „The fact of having been a victim of this spyware is a motivation to denounce those responsible.[...]".

However, for all journalists analyzed, Pegasus is not the only factor influencing their activity. In addition to „normal" fluctuations (such as planned vacations, time invested in projects, ...) almost all of the affected journalists are also targeted with other forms of harassment and oppression. Some of them were even imprisoned during the time period we analyzed, like Carlos Ketohu, which of course impacts their ability to publish content. Even less aggressive measures such as lawsuits, smear campaigns etc. occurring in parallell to targeting with Pegasus influence the jounalists' perception of being threatened and their publishing behavior.

The Pegasus spyware is thus just one part of the toolbox used by autocrats to put pressure on independent journalists, activists and opposition figures. It is particularly dangerous due

to its stealth nature and creates a feeling that „no situation or data is safe from being spied on", but it is rarely used alone. Therefore, it is very hard to determine the effect caused by Pegasus alone on its victims. We have seen some examples of significant changes in publishing behavior around the time of discovery of the spyware, showing that the surveillance weapon can have a significant effect.

Of course, it is very difficult to deal with the oppression, begin spied on, harassed and arrested. As we've seen before, Ferdinand Ayité has a very unique way of dealing with these circumstances. Therefore, we'd like to close with his words from a Facebook post he made after being released from prison in January 2021:

> When you are a journalist [...] in addition to investigation, prison is a must in some countries. [...] It is [...] the risk of our profession. Today I come out of this 21-day hostage situation more determined than before [...] I would just like to remind you that neither persecution nor imprisonment, let alone intimidation, will undermine our determination to be on the side of the truth. [71]

Sadly, it is true that being spied on, subjected to harassment and imprisoned is an inevitable part of being an independent journalist in some countries. The fact that people like the journalists included in this report still keep doing their important, but dangerous work, lets us hope that some day, free journalism will be possible without repercussions anywhere.

A. Results for all analyzed journalists

This section provides the graphs for all the 44 journalists we analyzed.

The following journalists were left out due to not finding any suitable data sources (neither a Twitter nor Facebook account nor a suitable publishing house):

- Jaspal Heran: Publishes an online newspaper in the vernacular language Punjabi as PNG images. Parsing would require Punjabi optical character recognition (OCR)
- David Dercsenyi: No Twitter account, no Facebook profile, Publishing House doesn't allow access by author.
- Omar Brouksy: No Twitter account, no Facebook profile, can only find a single article by him on newspaper website.
- Cecilio Pineda: Murdered in 2017. No online profile found.
- Soulemaine Raissouni: No online presence.

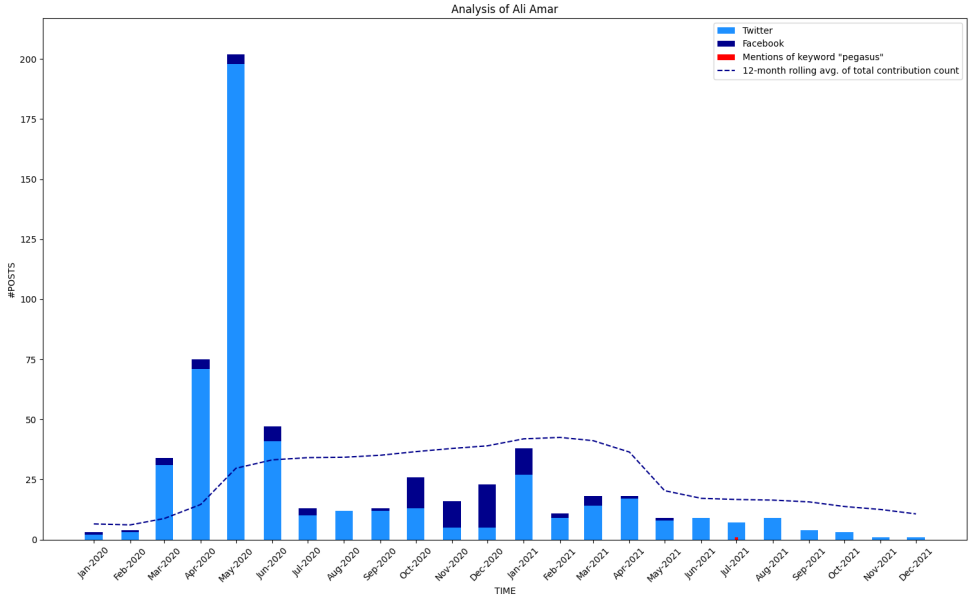

Figure A.1.: Publication statistics for Ali Amar.

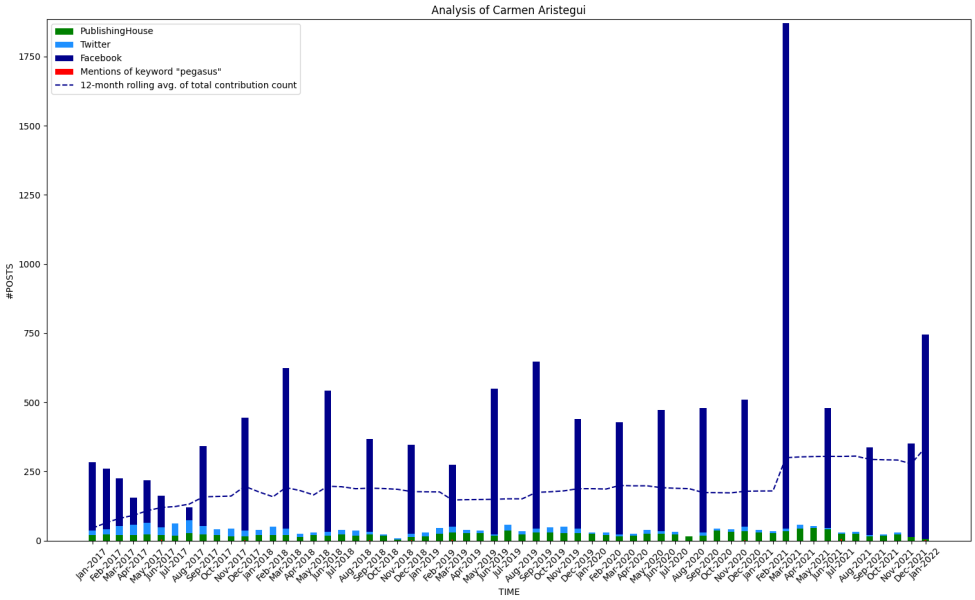

Figure A.2.: Publication statistics for Carmen Aristegui.

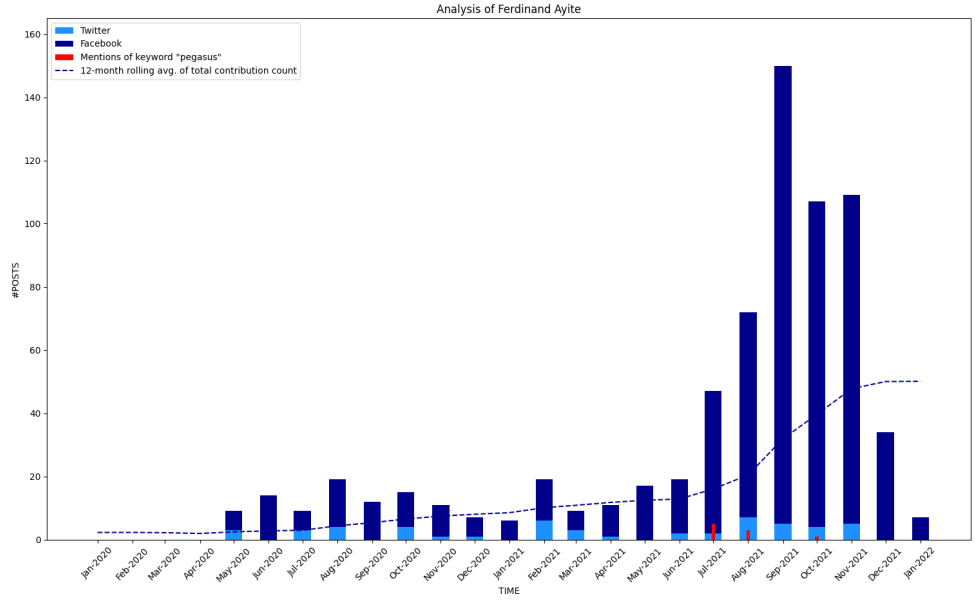

Figure A.3.: Publication statistics for Ferdinand Ayité.

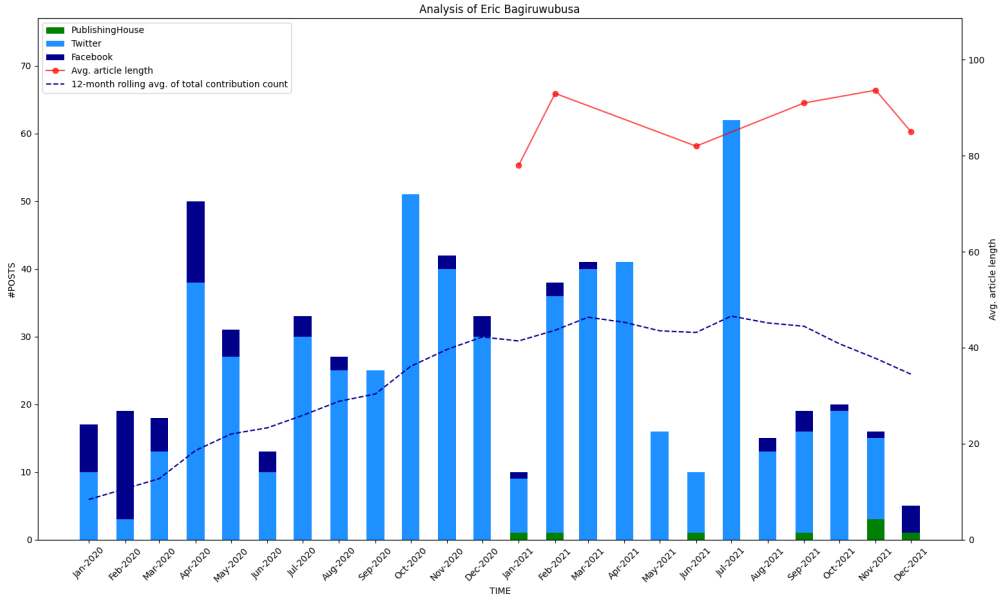

Figure A.4.: Publication statistics for Eric Bagiruwubusa.

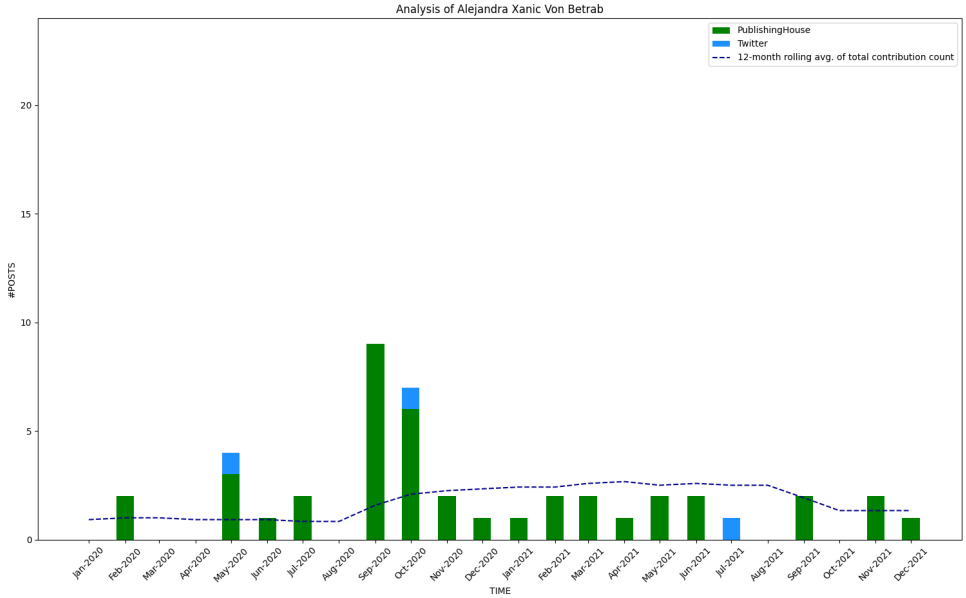

Figure A.5.: Publication statistics for Alejandra Xanic Von Betrab. Unable to parse her news website, Quinto Elemento Lab, by author. Instead parsed the entire website, which was co-founded by her and Marcela Turati. Thus the Publishing House data includes activity from other journalists as well.

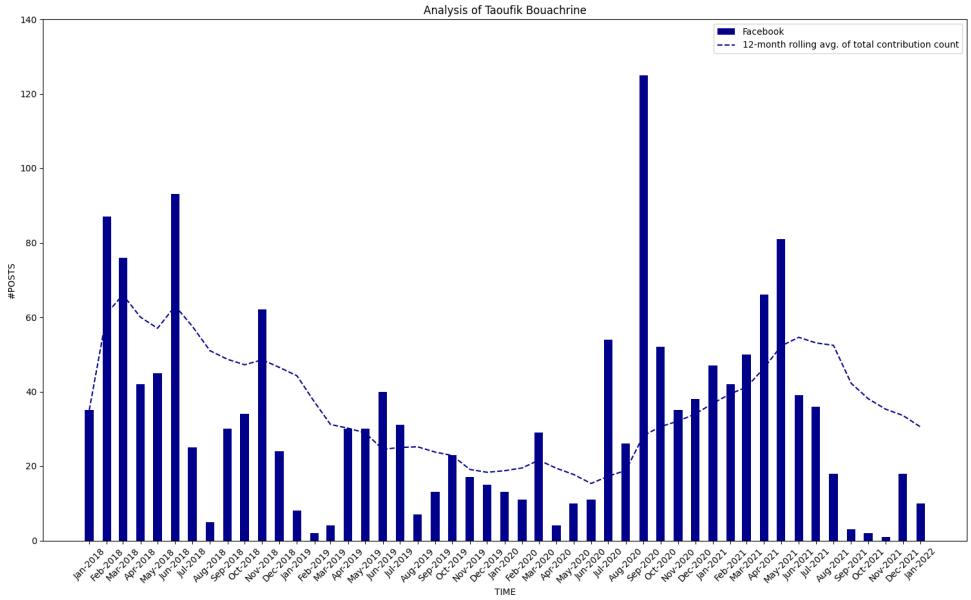

Figure A.6.: Publication statistics for Taoufik Bouachrine.

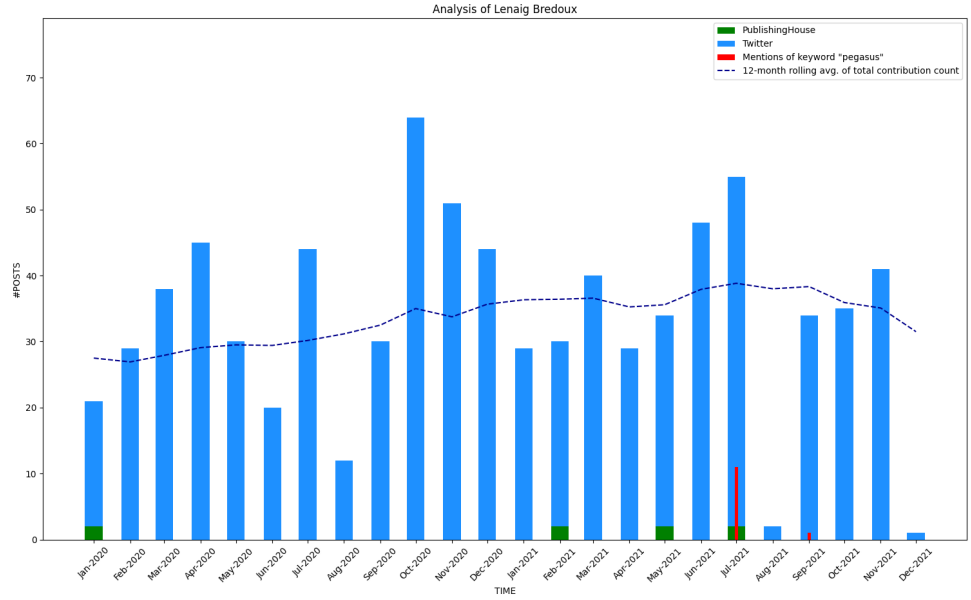

Figure A.7.: Publication statistics for Lenaïg Bredoux. Article length could not be parsed because full text is only accessible with a paid subscription

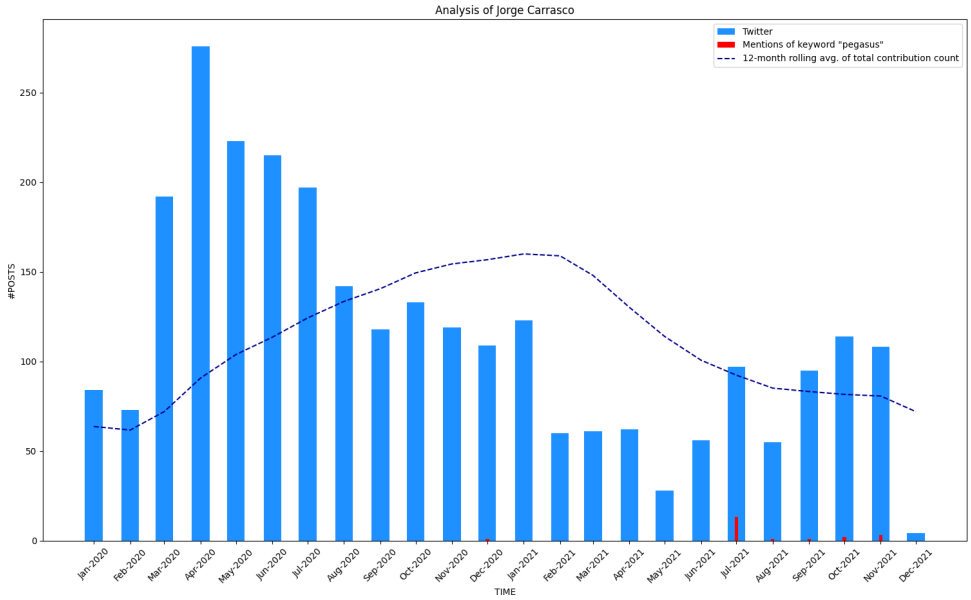

Figure A.8.: Publication statistics for Jorge Carrasco.

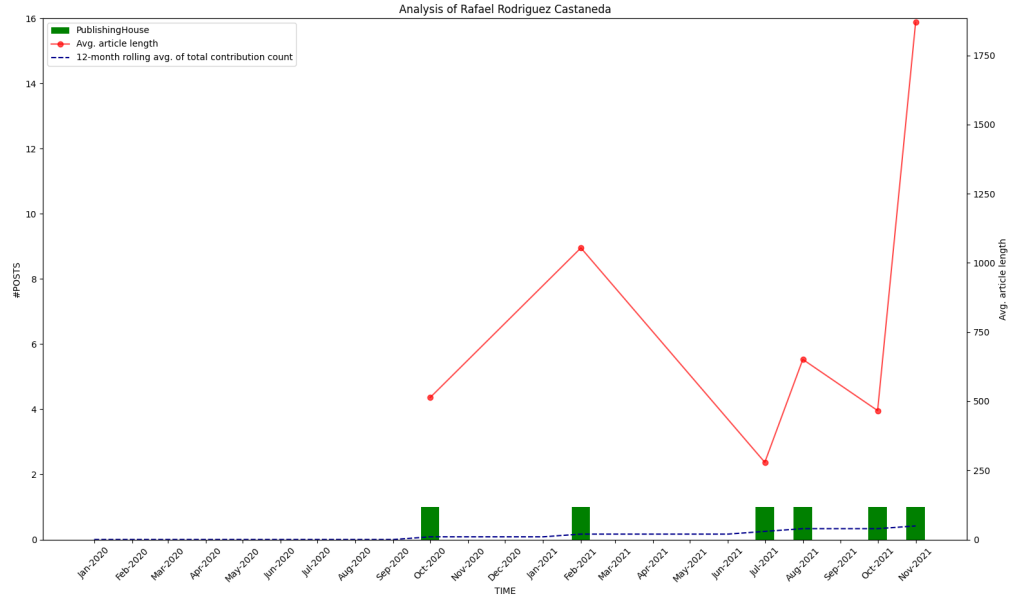

Figure A.9.: Publication statistics for Rafael Rodriguez Castañeda.

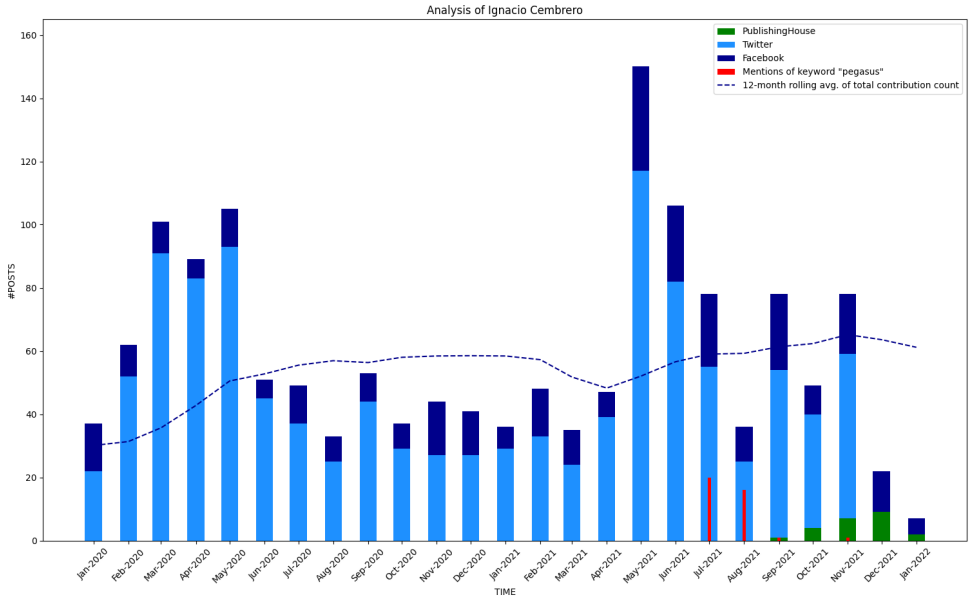

Figure A.10.: Publication statistics for Ignacio Cembrero.

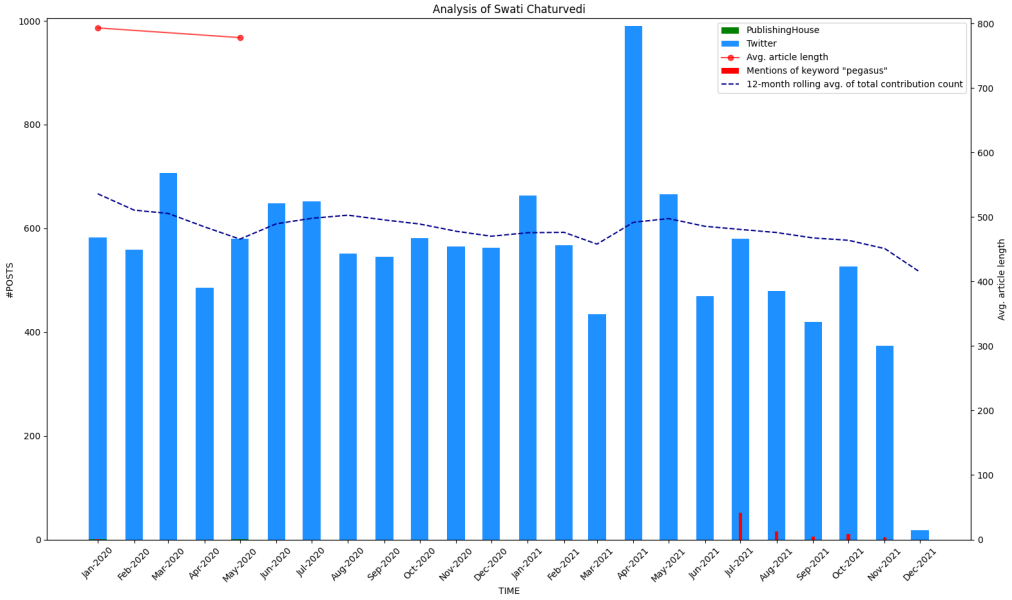

Figure A.11.: Publication statistics for Swati Chaturvedi.

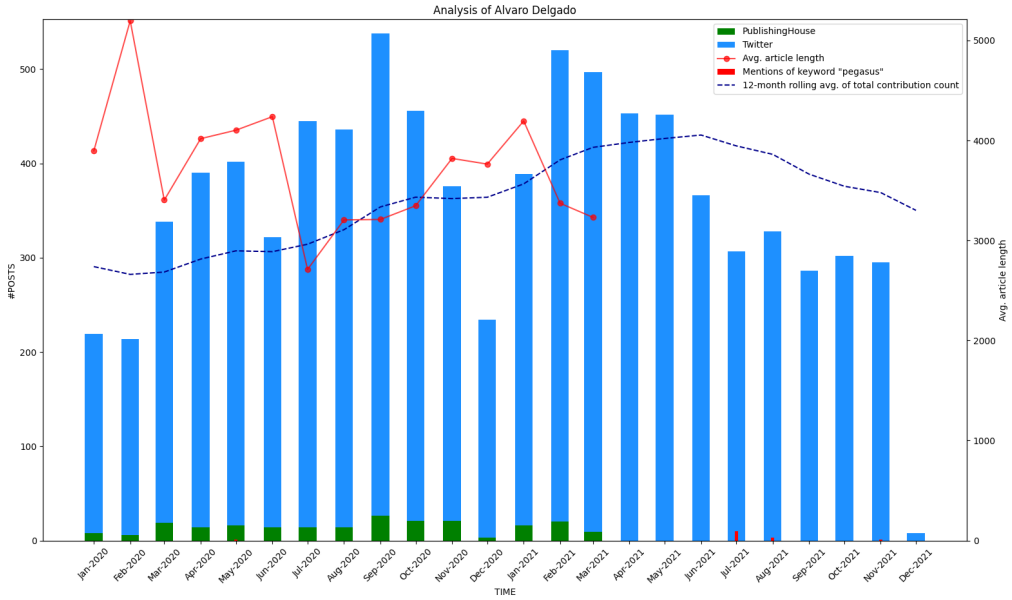

Figure A.12.: Publication statistics for Alvaro Delgado.

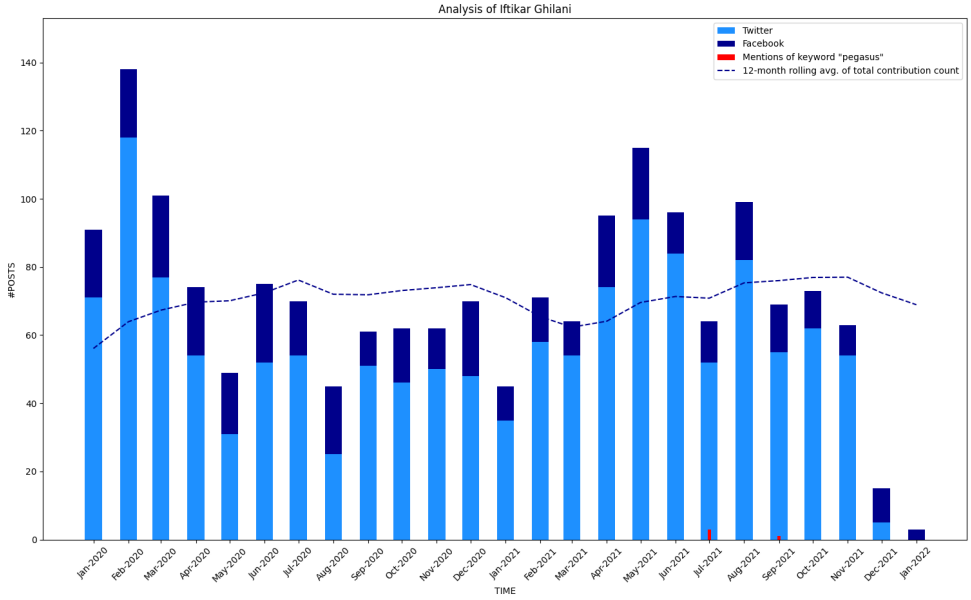

Figure A.13.: Publication statistics for Iftikar Ghilani.

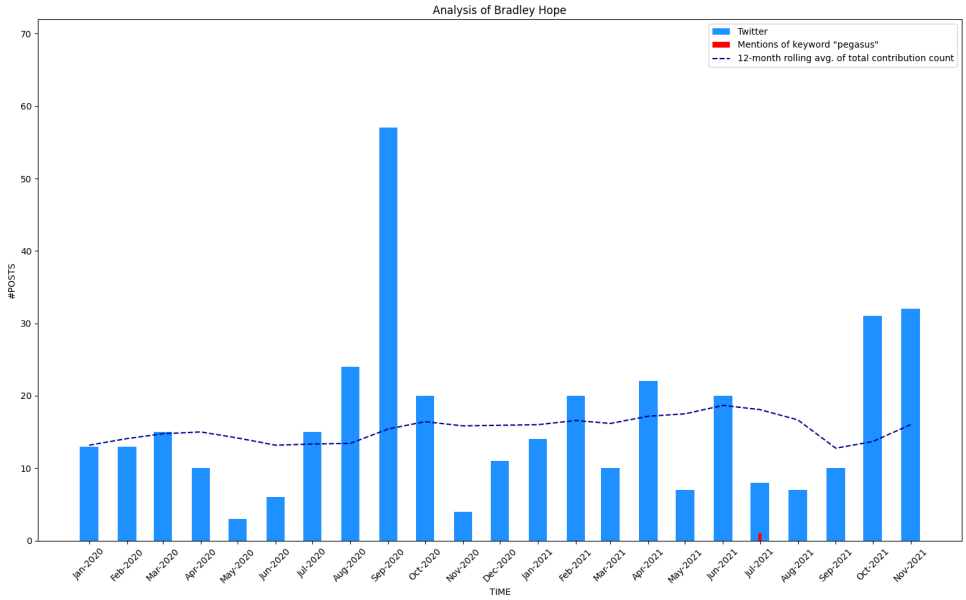

Figure A.14.: Publication statistics for Bradley Hope.

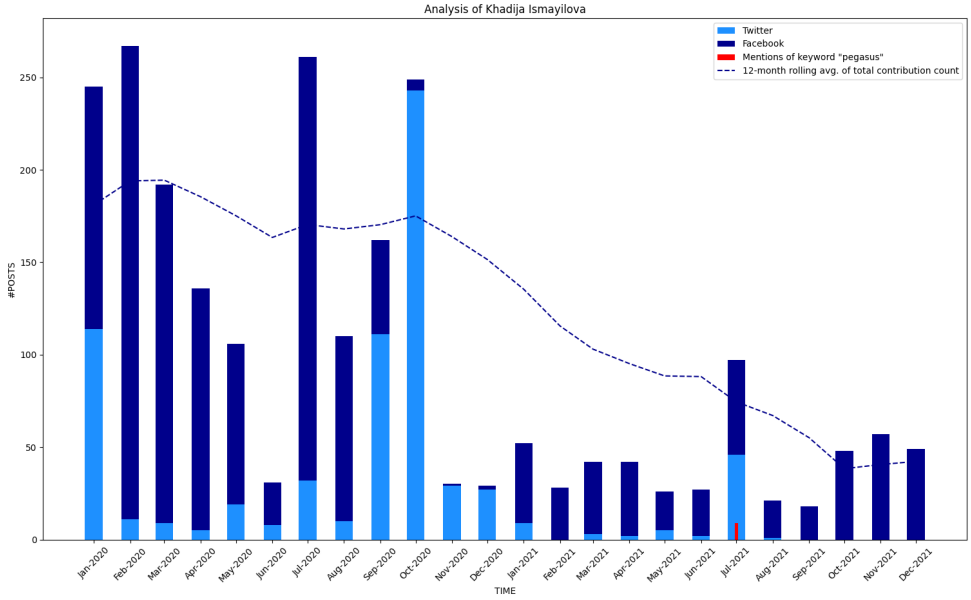

Figure A.15.: Publication statistics for Khadija Ismayilova.

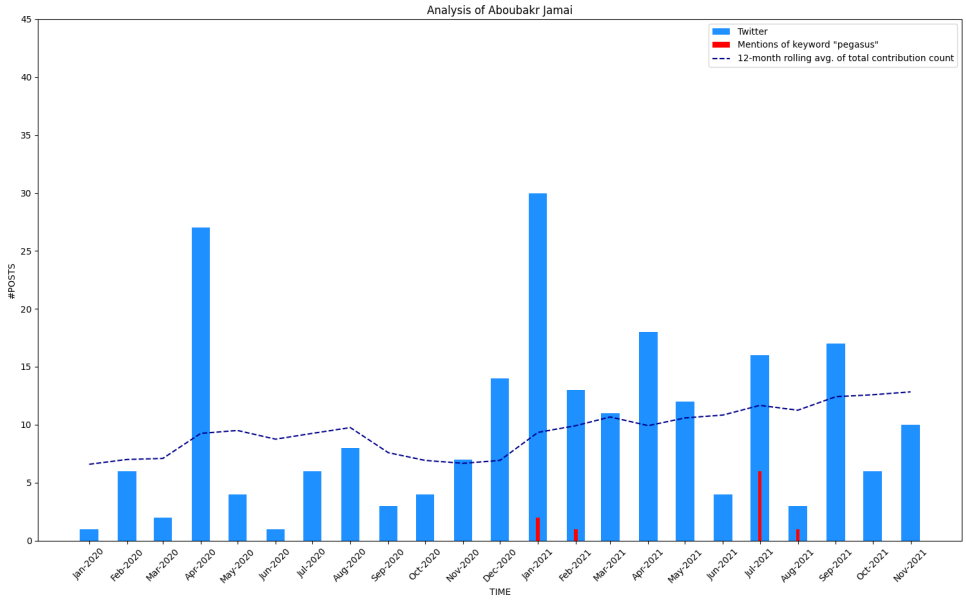

Figure A.16.: Publication statistics for Aboubakr Jamai.

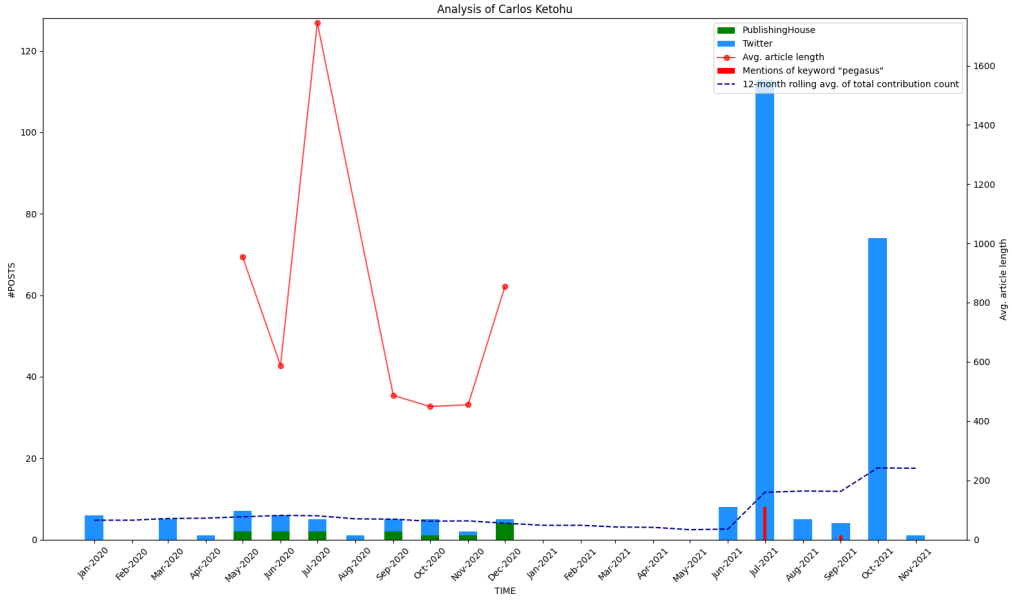

Figure A.17.: Publication statistics for Carlos Ketohou.

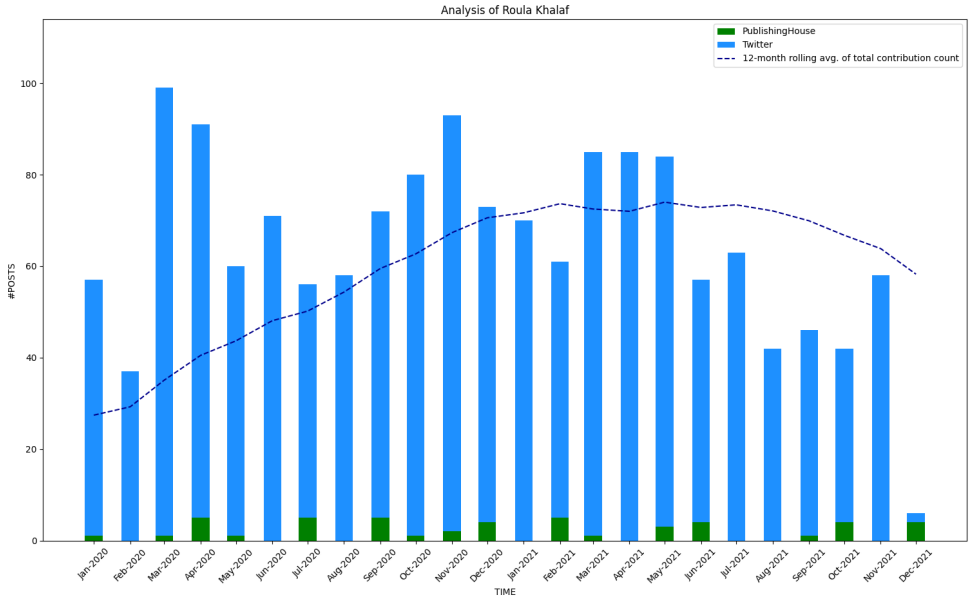

Figure A.18.: Publication statistics for Roula Khalaf. Article length could not be parsed because full text is only accessible with a paid subscription.

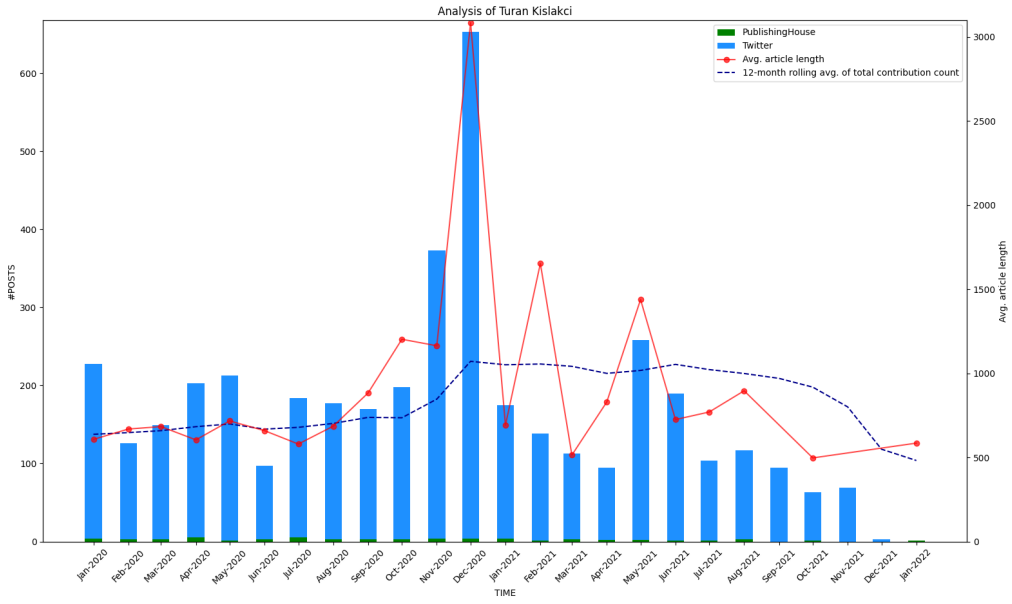

Figure A.19.: Publication statistics for Turan Kislakci.

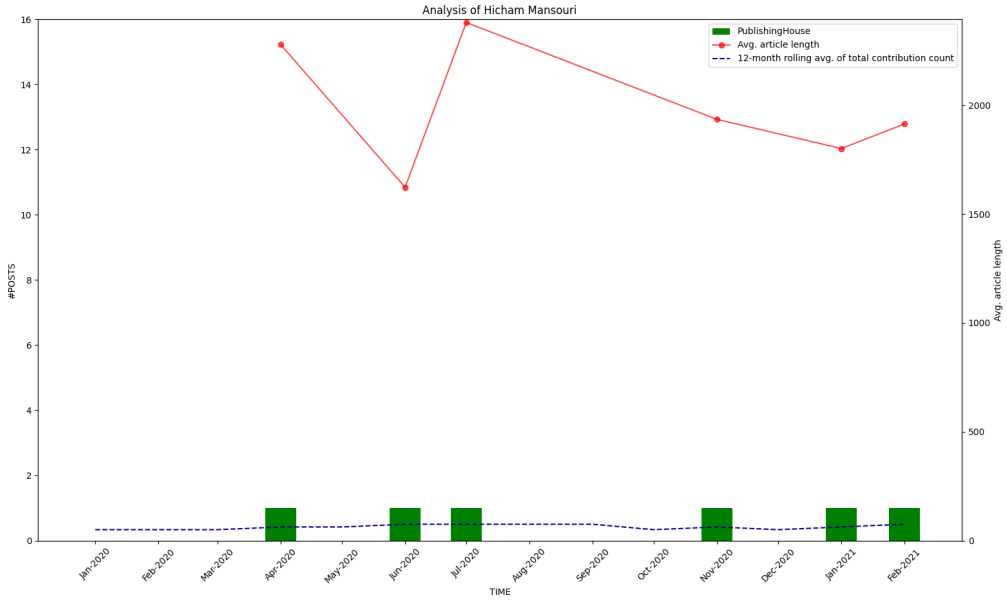

Figure A.20.: Publication statistics for Hicham Mansouri.

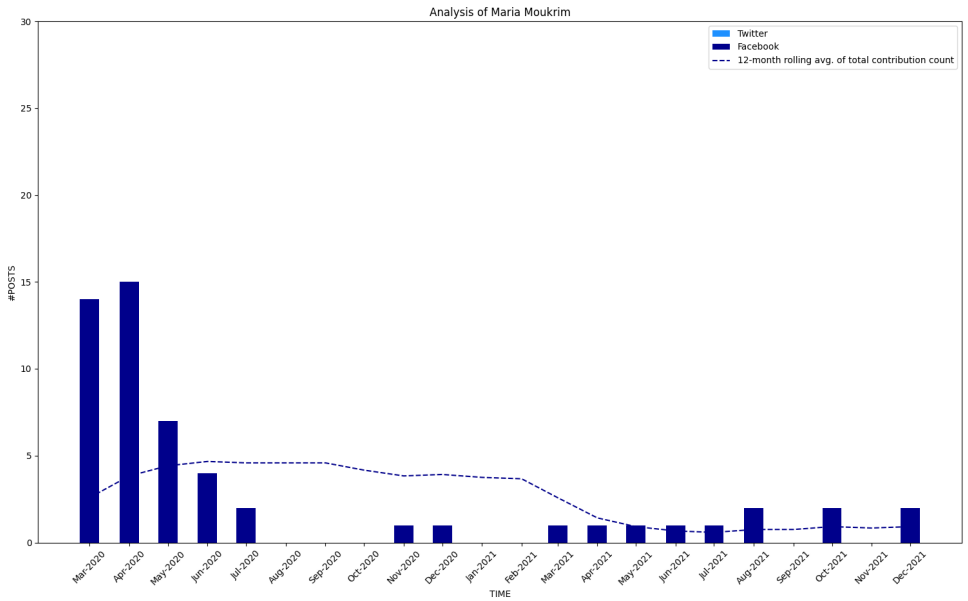

Figure A.21.: Publication statistics for Maria Moukrim. Publishing house is in Arabic, were not able to parse

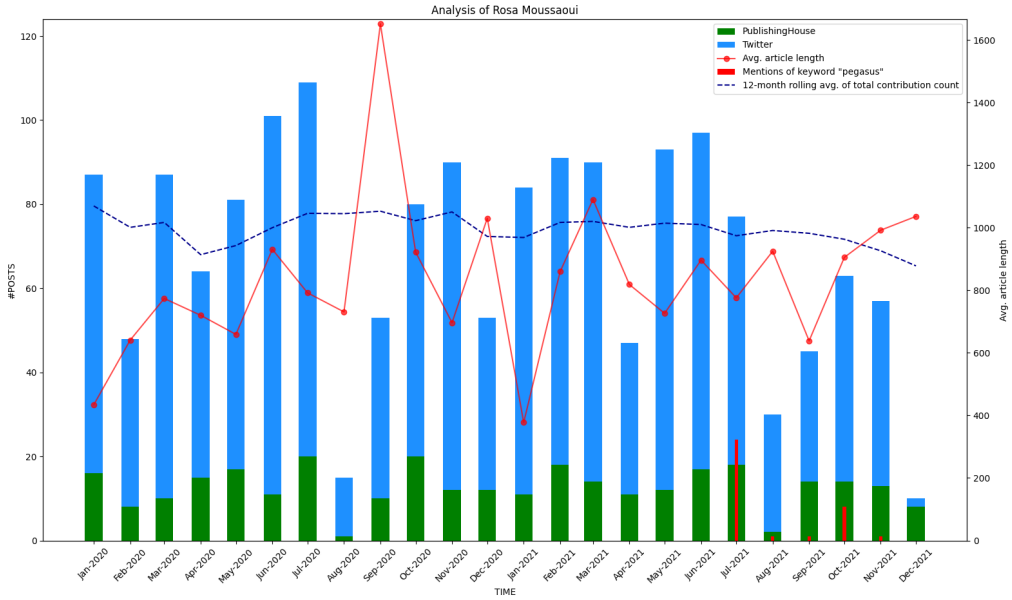

Figure A.22.: Publication statistics for Rosa Moussaoui.

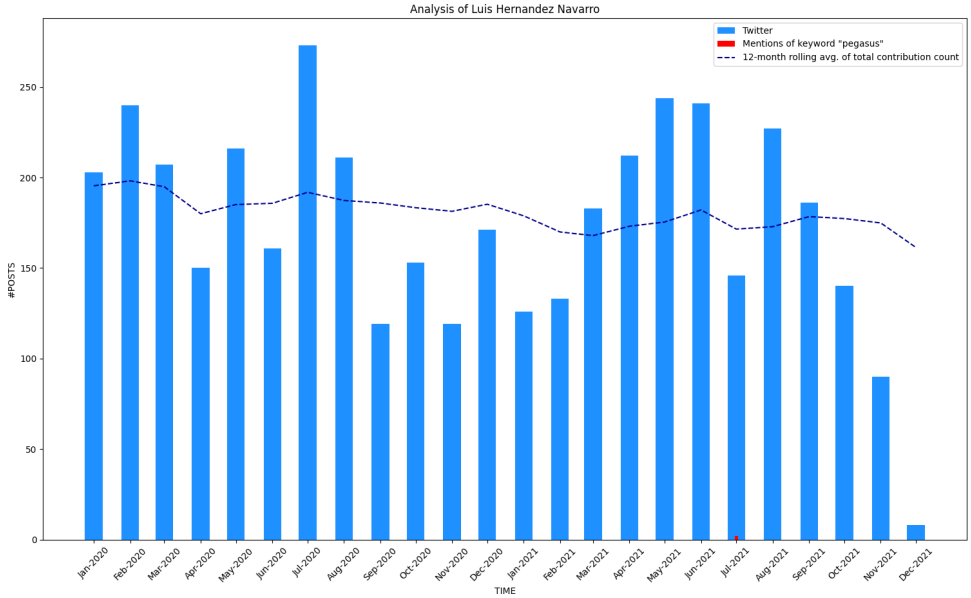

Figure A.23.: Publication statistics for Luis Hernández Navarro.

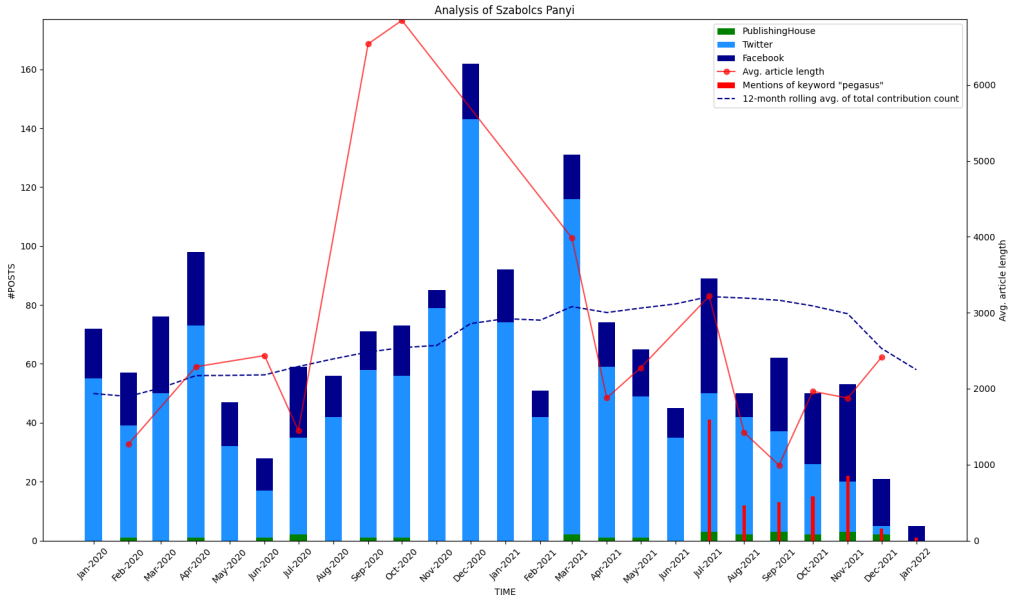

Figure A.24.: Publication statistics for Szabolcs Panyi.

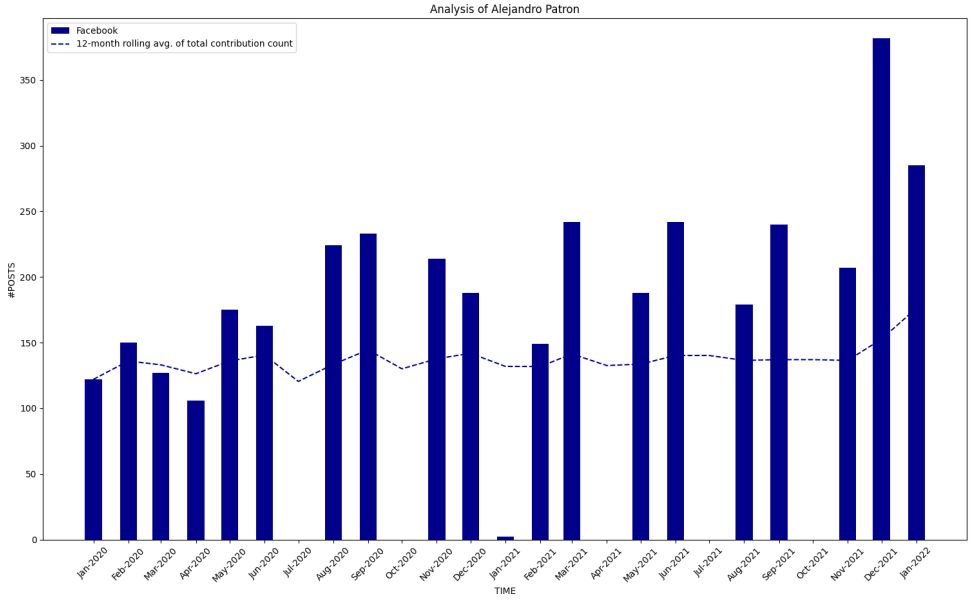

Figure A.25.: Publication statistics for Alejandro Patro.

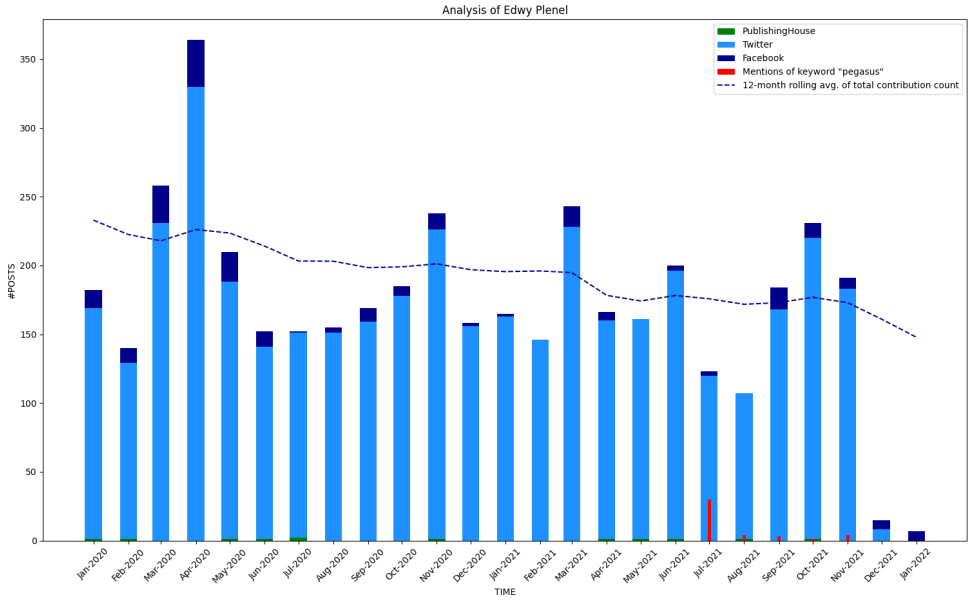

Figure A.26.: Publication statistics for Edwy Plenel.

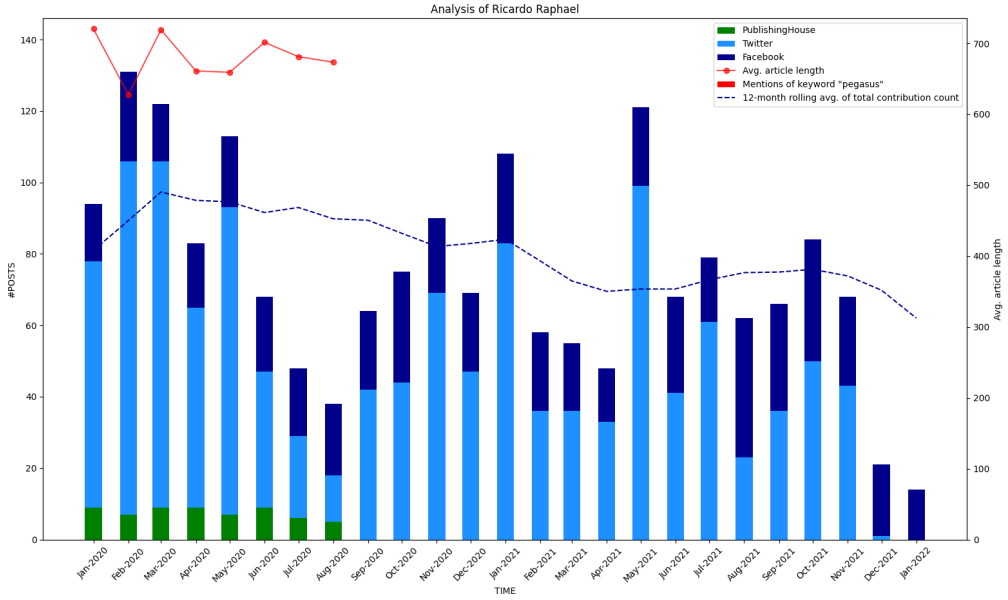

Figure A.27.: Publication statistics for Ricardo Raphael.

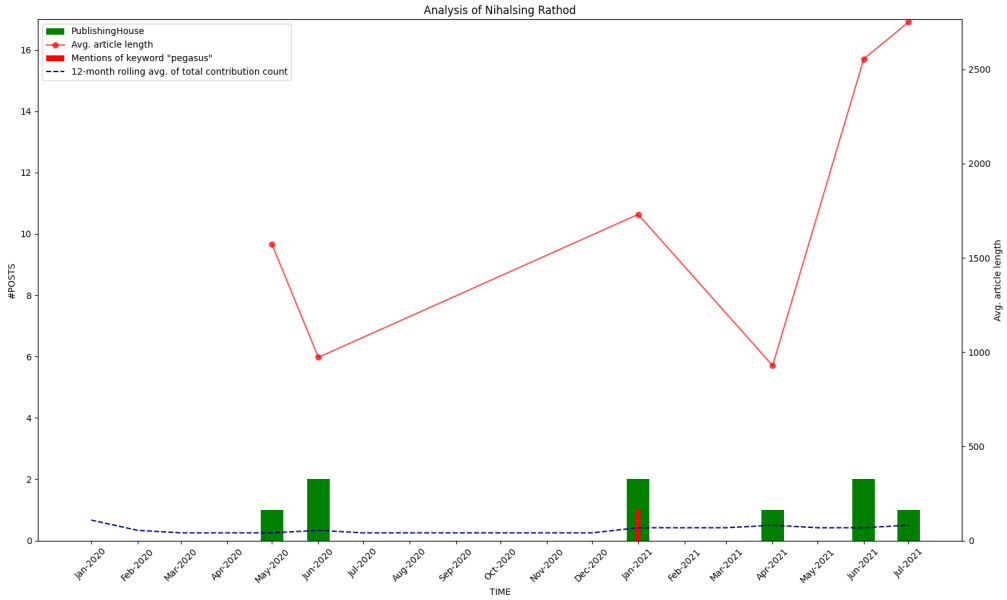

Figure A.28.: Publication statistics for Nihalsing Rathod.

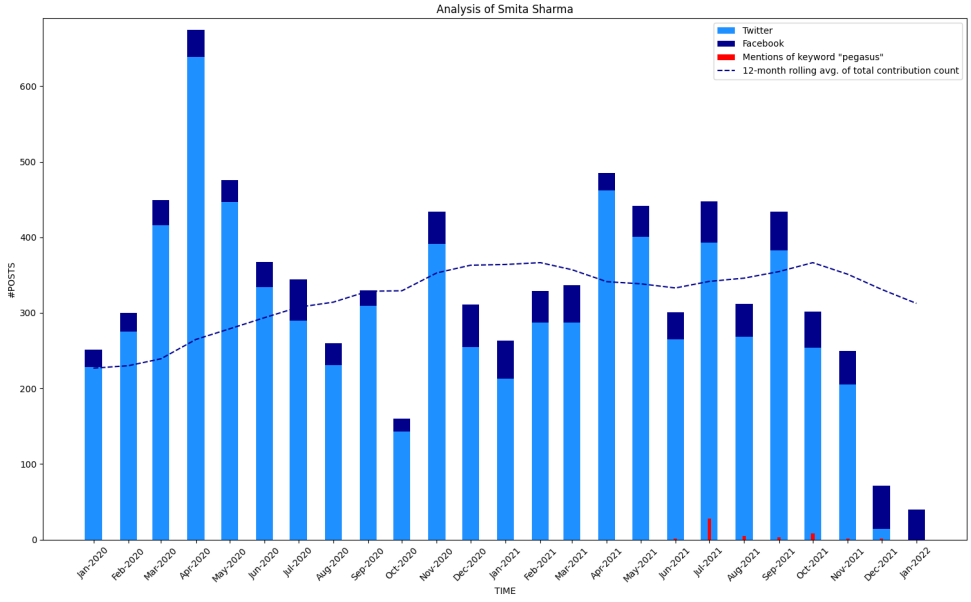

Figure A.29.: Publication statistics for Smita Sharma. Huffington Post has closed their Indian operations and is thus unavailable for parsing

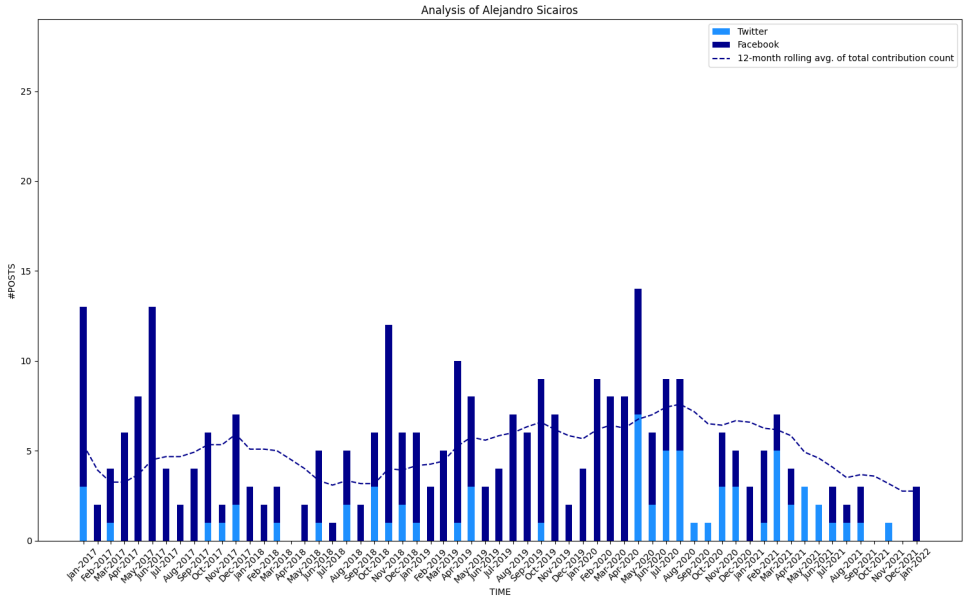

Figure A.30.: Publication statistics for Alejandro Sicairos. Noroeste Publishing House: Broken website, unable to parse.

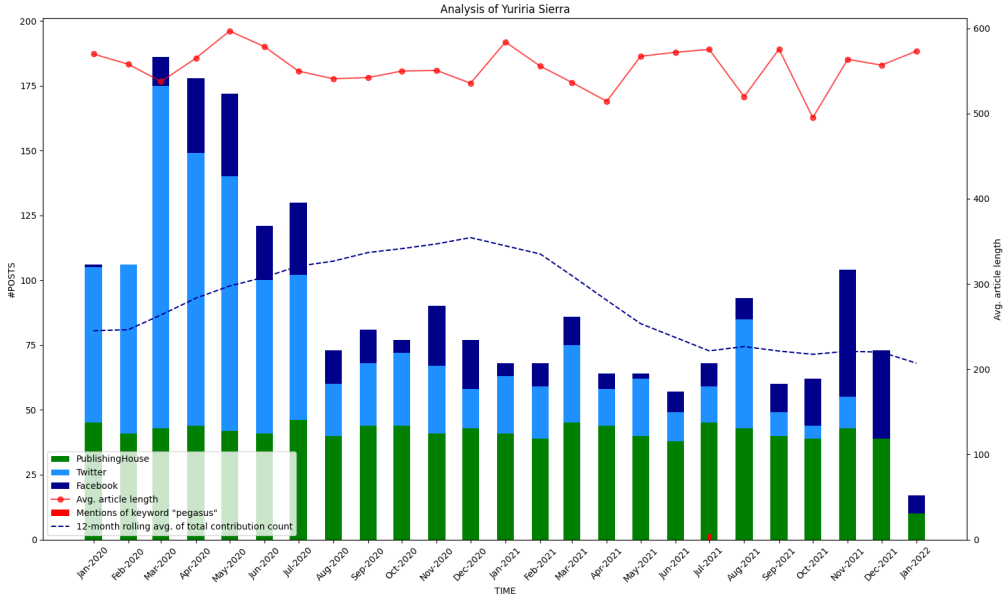

Figure A.31.: Publication statistics for Yuriria Sierra.

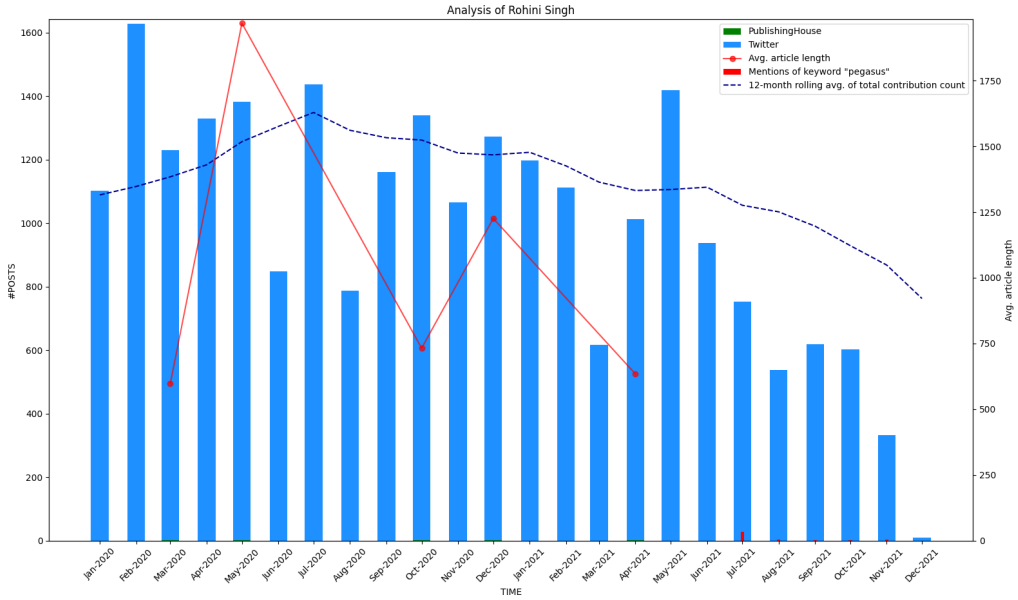

Figure A.32.: Publication statistics for Rohini Singh.

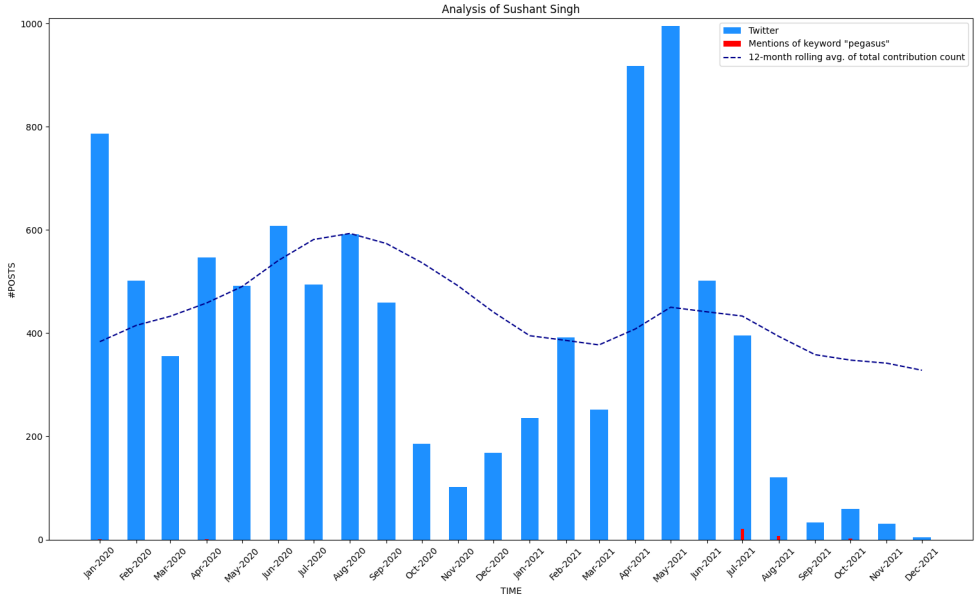

Figure A.33.: Publication statistics for Sushant Singh.

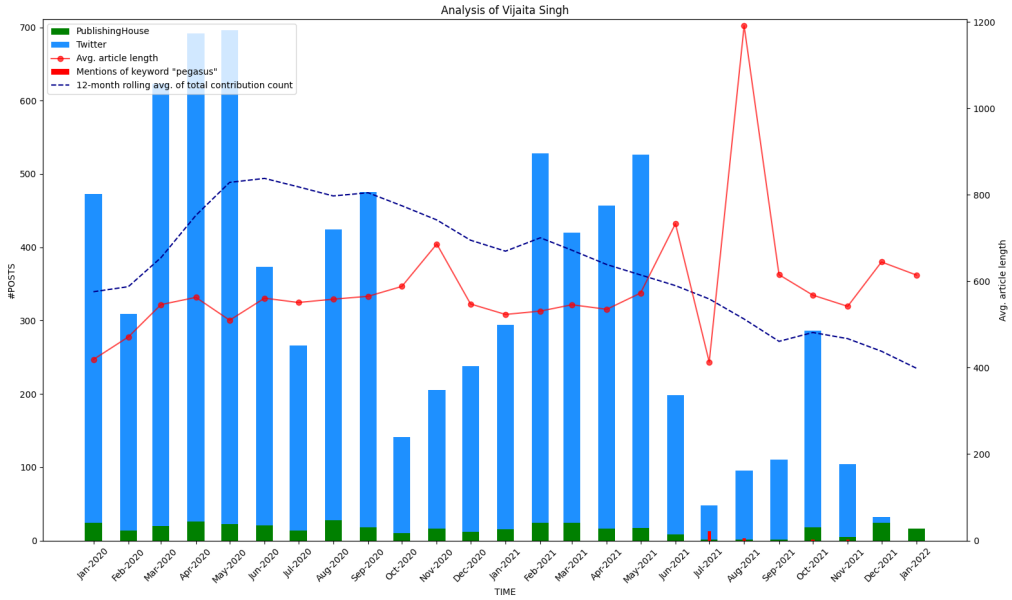

Figure A.34.: Publication statistics for Vijaita Singh.

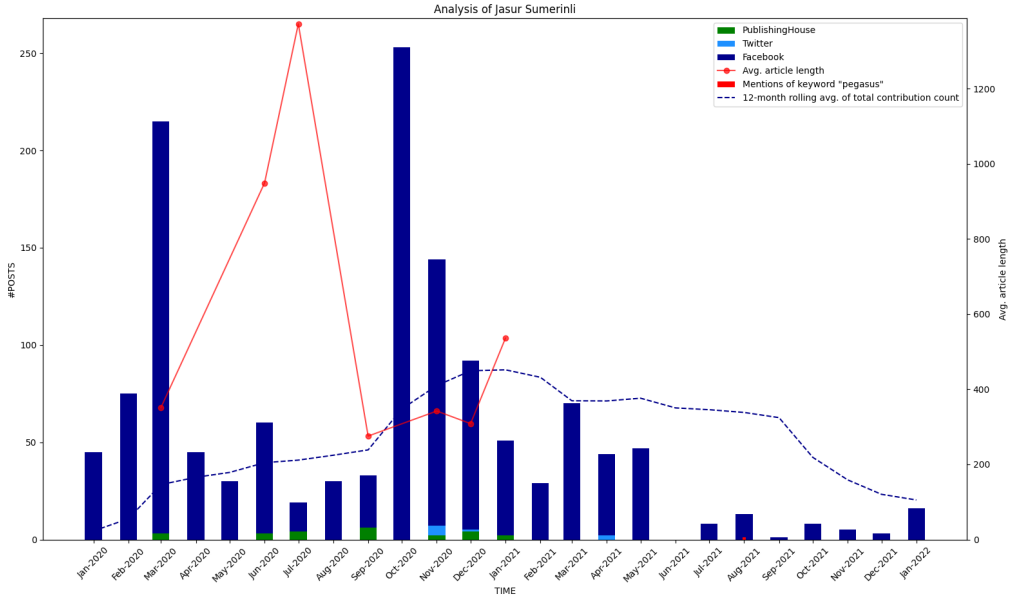

Figure A.35.: Publication statistics for Jasur Sumerinli.

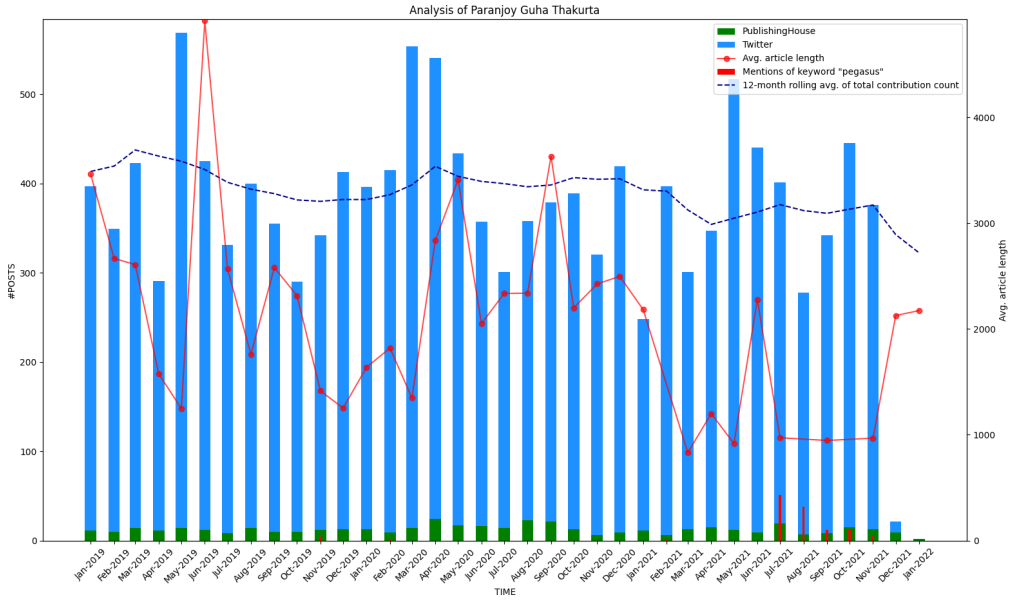

Figure A.36.: Publication statistics for Paranjoy Guha Thakurta.

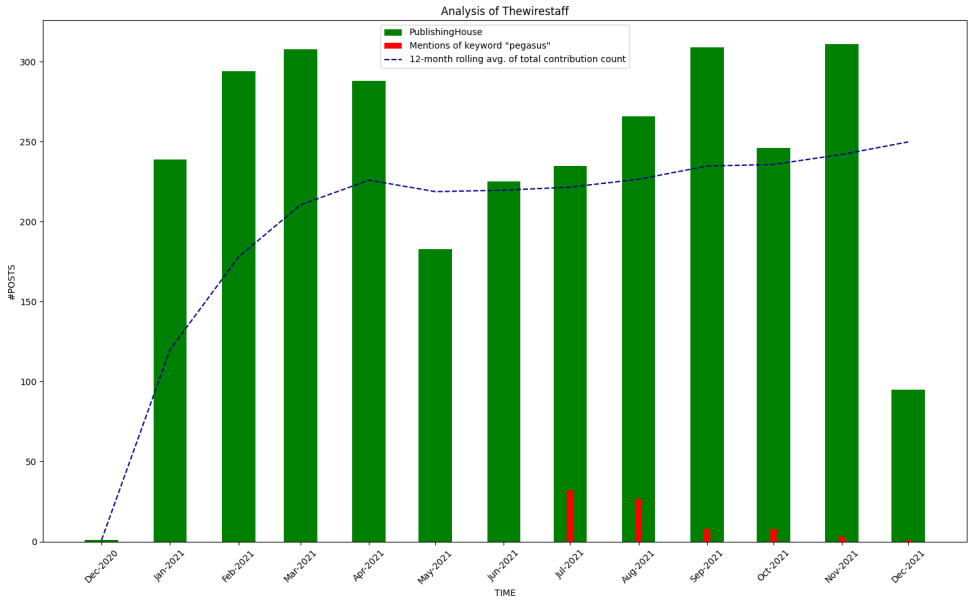

Figure A.37.: Publication statistics for Thewirestaff. The Wire is the name of the news outlet of Siddhart Varadarajan, Rohini Singh, MK Venu and Swati Chaturvedi on our list. Most articles on their website are not posted under a specific author's name, but rather under the name "The Wire Staff". Because The Wire was heavily targeted with Pegasus, we wanted to include this activity in our analysis.

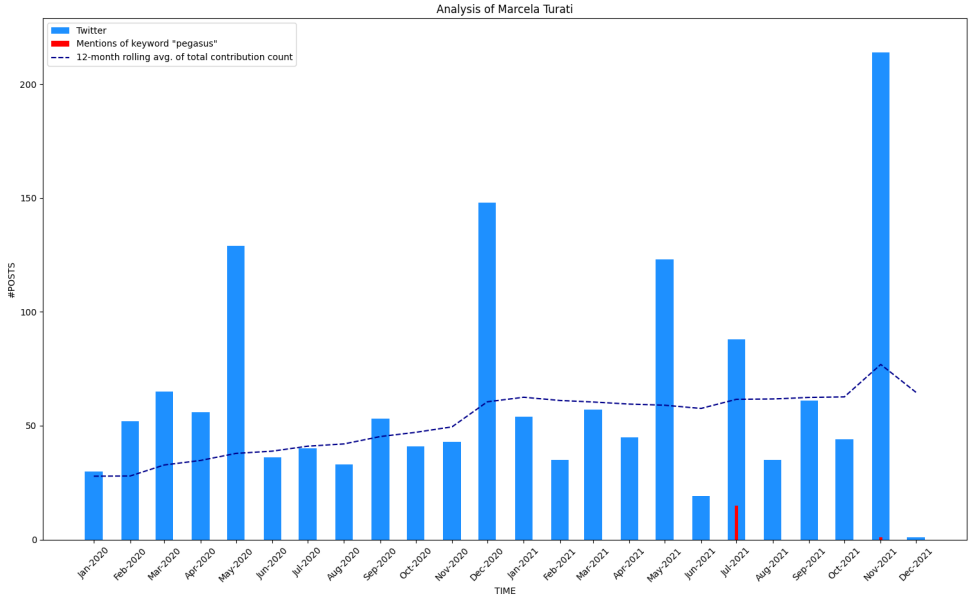

Figure A.38.: Publication statistics for Marcela Turati. Works with Alexandra Xanic von Betrab at Quinto Elemento Lab. Article counts for entire news site are included in her graph, see Figure A.5.

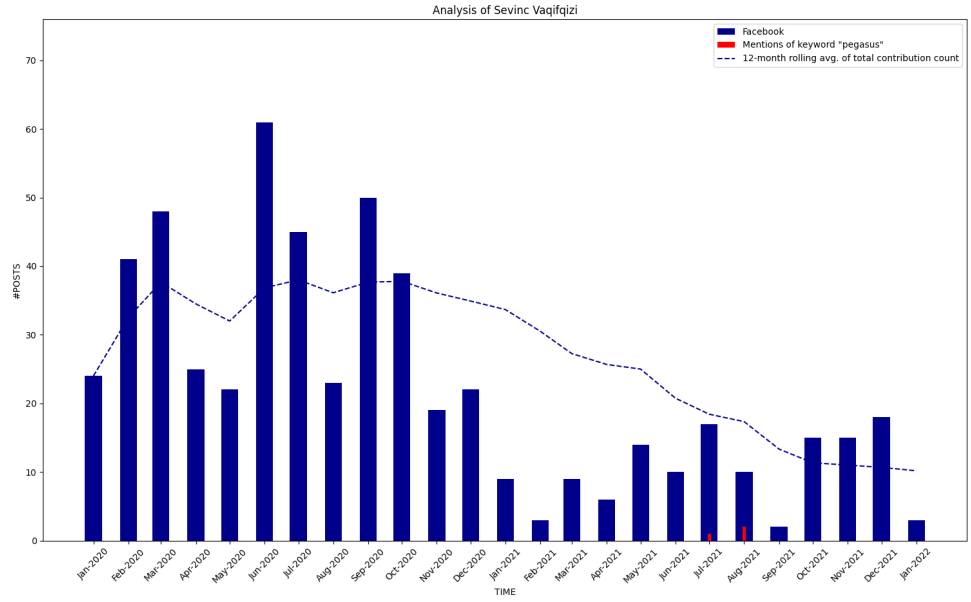

Figure A.39.: Publication statistics for Sevinc Vaqifqizi.

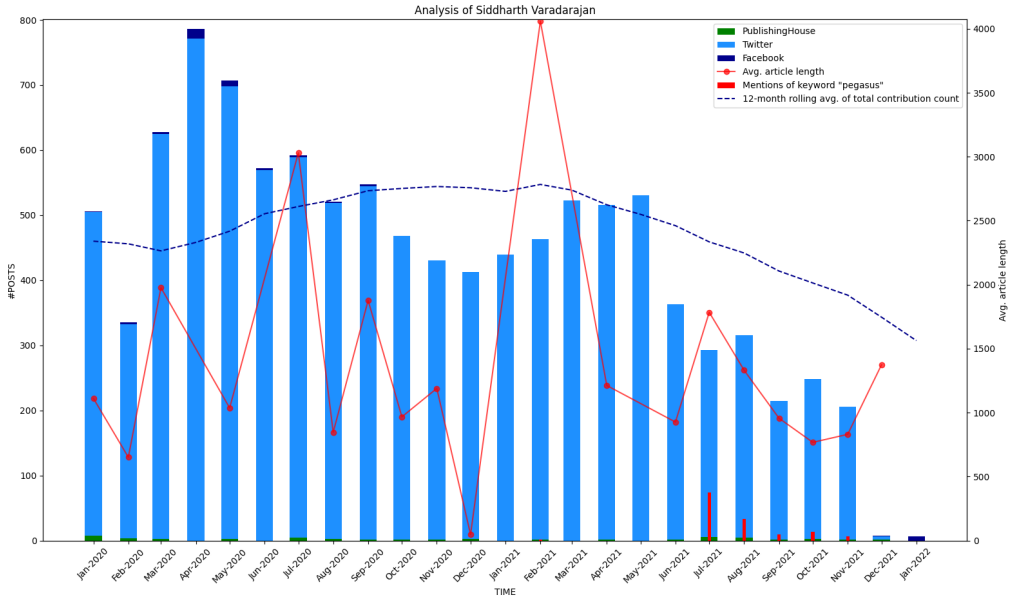

Figure A.40.: Publication statistics for Siddharth Varadarajan.

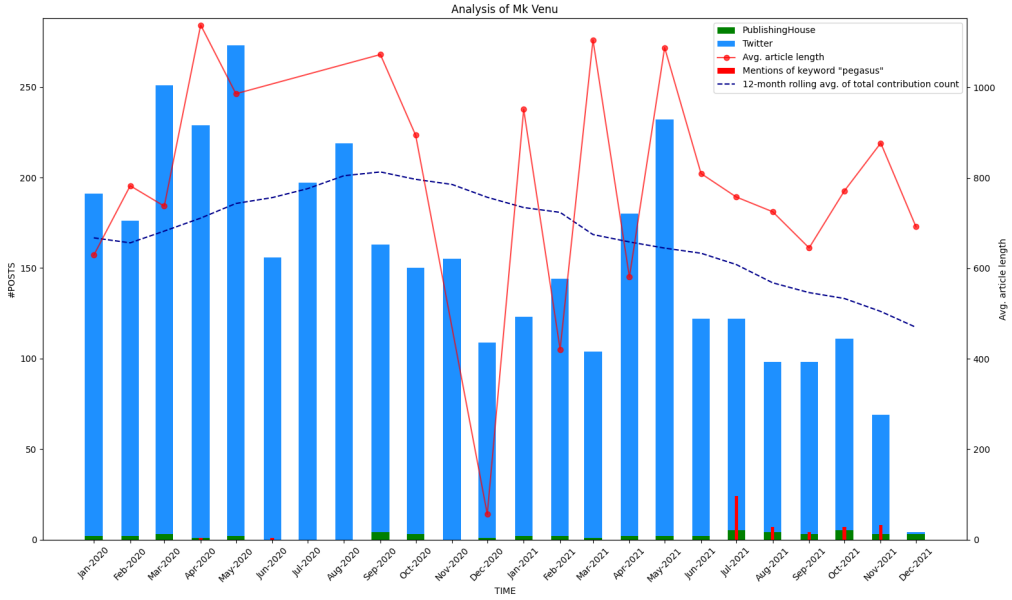

Figure A.41.: Publication statistics for MK Venu.

List of Figures

Bibliography

[1] Amensty International. *Forensic Methodology Report: Pegasus Forensic Traces per Target*. https://www.amnesty.org/en/latest/research/2021/07/forensic-methodology-report-appendix-d/, accessed Feb 10th, 2022.

[2] Dave Lee. *Who are the hackers who cracked the iPhone?* https://www.bbc.com/news/technology-37192670, accessed Feb 10th, 2022.

[3] Lookout. *Technical Analysis of Pegasus Spyware*. https://info.lookout.com/rs/051-ESQ-475/images/lookout-pegasus-technical-analysis.pdf, accessed Feb 10th, 2022.

[4] Lorenzo Franceschi-Bicchierai. *Government Hackers Caught Using Unprecedented iPhone Spy Tool*. https://www.vice.com/en/article/3da5qj/government-hackers-iphone-hacking-jailbreak-nso-group, accessed Feb 10th, 2022.

[5] Chaim Levinson. *With Israel's Encouragement, NSO Sold Spyware to UAE and Other Gulf States*. https://www.haaretz.com/middle-east-news/.premium-with-israel-s-encouragement-nso-sold-spyware-to-uae-and-other-gulf-states-1.9093465, accessed Feb 10th, 2022.

[6] Forbidden Stories. *About the pegasus project*. https://forbiddenstories.org/about-the-pegasus-project/, accessed Feb 10th, 2022.

[7] Stephanie Kirchgaessner, Paul Lewis, David Pegg,Sam Cutler,Nina Lakhani , Michael Safi. *Revealed: leak uncovers global abuse of cyber-surveillance weapon*. https://www.theguardian.com/world/2021/jul/18/revealed-leak-uncovers-global-abuse-of-cyber-surveillance-weapon-nso-group-pegasus, accessed Feb 10th, 2022.

[8] Forbidden Stories. *THE PEGASUS PROJECT: A WORLDWIDE COLLABORATION TO COUNTER A GLOBAL CRIME*. https://forbiddenstories.org/the-pegasus-project-a-worldwide-collaboration-to-counter-a-global-crime/, accessed Feb 10th, 2022.

[9] Devirupa Mitra. *Pegasus Project: 14 World Leaders in Leaked Database*. https://thewire.in/world/pegasus-project-14-world-leaders-macron-ramaphosa-michel-imran, accessed Feb 10th, 2022.

[10] Shaun Walker, Stephanie Kirchgaessner, Nina Lakhani and Michael Safi. *Pegasus project: spyware leak suggests lawyers and activists at risk across globe*. https://www.theguardian.com/news/2021/jul/19/spyware-leak-suggests-lawyers-and-activists-at-risk-across-globe, accessed Feb 10th, 2022.

[11] Forbidden Stories. *Journalists under surveillance*. https://forbiddenstories.org/pegasus-journalists-under-surveillance/, accessed Feb 10th, 2022.

[12] S. Kirchgaessner, P. Lewis, D. Pegg, S. Cutler, N. Lakhani, and M. Safi. "Revealed: leak uncovers global abuse of cyber-surveillance weapon". en-GB. In: *The Guardian* (July 2021). ISSN: 0261-3077. URL: https://www.theguardian.com/world/2021/jul/18/revealed-leak-uncovers-global-abuse-of-cyber-surveillance-weapon-nso-group-pegasus (visited on 02/19/2022).

[13] Kristina Ludwig and Hannes Munzinger. "Wie Aserbaidschan Kritiker ausspäht und demütigt". de. In: *Süddeutsche Zeitung* (). Section: Politik. URL: https://www.sueddeutsche.de/projekte/artikel/politik/pegasus-project-wie-aserbaidschan-kritiker-bekaempft-e297191/ (visited on 02/19/2022).

[14] K. Willsher. "Pegasus spyware found on journalists' phones, French intelligence confirms". en-GB. In: *The Guardian* (Aug. 2021). ISSN: 0261-3077. URL: https://www.theguardian.com/news/2021/aug/02/pegasus-spyware-found-on-journalists-phones-french-intelligence-confirms (visited on 02/19/2022).

[15] Siddhart Varadarajan. "Revealed: How The Wire and Its Partners Cracked the Pegasus Project and What It Means for India". In: *The Wire* (). URL: https://thewire.in/media/revealed-how-the-wire-partners-cracked-pegasus-project-implications-india (visited on 02/19/2022).

[16] Amnesty International Security Lab. *Forensic Methodology Report: How to Catch NSO Group's Pegasus*. en. Tech. rep. Amnesty International, 2018. URL: https://www.amnesty.org/en/documents/doc10/4487/2021/en/ (visited on 02/19/2022).

[17] B. Marczak, J. Scott-Railton, S. Anstis, and R. Deibert. *Independent Peer Review of Amnesty International's Forensic Methods for Identifying Pegasus Spyware*. Tech. rep. Citizen Lab, University of Toronto, July 2021. URL: https://citizenlab.ca/2021/07/amnesty-peer-review/ (visited on 02/19/2022).

[18] *FAQ: On the Pegasus Project's Digital Forensics*. FAQ:OnthePegasusProject'sDigitalForensics, accessed March 10th, 2022.

[19] B. Marczak and J. Scott-Railton. *The Million Dollar Dissident: NSO Group's iPhone Zero-Days used against a UAE Human Rights Defender*. Tech. rep. Citizen Lab Research Report No. 78. University of Toronto, Aug. 2016. URL: https://citizenlab.ca/2016/08/million-dollar-dissident-iphone-zero-day-nso-group-uae/ (visited on 02/19/2022).

[20] J. Scott-Railton, B. Marczak, C. Guarnieri, and M. Crete-Nishihata. *Bitter Sweet: Supporters of Mexico's Soda Tax Targeted With NSO Exploit Links*. Tech. rep. Citizen Lab Research Report No. 89. University of Toronto, Feb. 2017. URL: https://citizenlab.ca/2017/02/bittersweet-nso-mexico-spyware/ (visited on 11/09/2021).

[21] J. Scott-Railton, B. Marczak, B. A. Razzak, M. Crete-Nishihata, and R. Deibert. *Reckless Exploit: Mexican Journalists, Lawyers, and a Child Targeted with NSO Spyware*. Tech. rep. Citizen Lab Research Report No. 93. University of Toronto, June 2017. URL: https://citizenlab.ca/2017/06/reckless-exploit-mexico-nso/ (visited on 02/19/2022).

[22] J. Scott-Railton, B. Marczak, B. A. Razzak, M. Crete-Nishihata, and R. Deibert. *Reckless Redux: Senior Mexican Legislators and Politicians Targeted with NSO Spyware*. Tech. rep. Citizen Lab Research Report No. 94. University of Toronto, June 2017. URL: https://citizenlab.ca/2017/06/more-mexican-nso-targets/ (visited on 02/19/2022).

[23] J. Scott-Railton, B. Marczak, B. A. Razzak, M. Crete-Nishihata, and R. Deibert. *Reckless III: Investigation Into Mexican Mass Disappearance Targeted with NSO Spyware*. Tech. rep. Citizen Lab Research Report No. 96. University of Toronto, July 2017. URL: https://citizenlab.ca/2017/07/mexico-disappearances-nso/ (visited on 02/19/2022).

[24] J. Scott-Railton, B. Marczak, S. Anstis, B. A. Razzak, M. Crete-Nishihata, and R. Deibert. *Reckless VII: Wife of Journalist Slain in Cartel-Linked Killing Targeted with NSO Group's Spyware*. Tech. rep. Citizen Lab Research Report No. 117. University of Toronto, Mar. 2019. URL: https://citizenlab.ca/2019/03/nso-spyware-slain-journalists-wife/ (visited on 02/19/2022).

[25] J. Scott-Railton, B. Marczak, S. Anstis, B. A. Razzak, M. Crete-Nishihata, and R. Deibert. *Reckless VI: Mexican Journalists Investigating Cartels Targeted with NSO Spyware Following Assassination of Colleague*. Tech. rep. Citizen Lab Research Report No. 116. University of Toronto, Nov. 2018. URL: https://citizenlab.ca/2018/11/mexican-journalists-investigating-cartels-targeted-nso-spyware-following-assassination-colleague/ (visited on 02/19/2022).

[26] J. Scott-Railton, B. Marczak, B. A. Razzak, M. Crete-Nishihata, and R. Deibert. *Reckless IV: Lawyers for Murdered Mexican Women's Families Targeted with NSO Spyware*. Tech. rep. Citizen Lab Research Report No. 98. University of Toronto, Aug. 2017. URL: https://citizenlab.ca/2017/08/lawyers-murdered-women-nso-group/ (visited on 02/19/2022).

[27] J. Scott-Railton, B. Marczak, S. Anstis, B. A. Razzak, M. Crete-Nishihata, and R. Deibert. *Reckless VII: Wife of Journalist Slain in Cartel-Linked Killing Targeted with NSO Group's Spyware*. Tech. rep. Citizen Lab Research Report No. 117. University of Toronto, Mar. 2019. URL: https://citizenlab.ca/2019/03/nso-spyware-slain-journalists-wife/ (visited on 02/19/2022).

[28] B. Marczak, J. Scott-Railton, S. McKune, B. A. Razzak, and R. Deibert. *HIDE AND SEEK: Tracking NSO Group's Pegasus Spyware to Operations in 45 Countries*. Tech. rep. Citizen Lab Research Report No. 113. University of Toronto, Sept. 2018. URL: https://citizenlab.ca/2018/09/hide-and-seek-tracking-nso-groups-pegasus-spyware-to-operations-in-45-countries/ (visited on 02/19/2022).

[29] B. Marczak, S. Anstis, M. Crete-Nishihata, J. Scott-Railton, and R. Deibert. *Stopping the press: New York Times journalist targeted by Saudi-linked Pegasus spyware operator*. Citizen Lab Research Report No. 124 124. University of Toronto, 2020.

[30] B. Marczak, J. Scott-Railton, S. Anstis, B. A. Razzak, and R. Deibert. *Breaking the News: New York Times Journalist Ben Hubbard Hacked with Pegasus after Reporting on Previous Hacking Attempts*. Tech. rep. Citizen Lab Research Report No. 145. University of Toronto, Oct. 2021. URL: https://citizenlab.ca/2021/10/breaking-news-new-york-times-journalist-ben-hubbard-pegasus/ (visited on 02/20/2022).

[31] B. Marczak, C. Guarnieri, M. Marquis-Boire, and J. Scott-Railton. *Mapping Hacking Team's "Untraceable" Spyware*. Tech. rep. Citizen Lab Research Report No. 33. University of Toronto, Feb. 2014. URL: https://citizenlab.ca/2014/02/mapping-hacking-teams-untraceable-spyware/ (visited on 02/20/2022).

[32] B. Marczak, G. Alexander, S. McKune, J. Scott-Railton, and R. Deibert. *Champing at the Cyberbit: Ethiopian Dissidents Targeted with New Commercial Spyware*. Tech. rep. Citizen Lab Research Report No. 102. University of Toronto, Dec. 2017. URL: https://citizenlab.ca/2017/12/champing-cyberbit-ethiopian-dissidents-targeted-commercial-spyware/ (visited on 02/20/2022).

[33] S. Woodhams. *Spyware: An Unregulated and Escalating Threat to Independent Media*. Tech. rep. Washington, DC: Center for International Media Assistance, 2021. URL: https://www.skeyesmedia.org/documents/bo_filemanager/CIMA_Spyware-Report_web_150ppi.pdf.

[34] D. D. Kirkpatrick. "Israeli Software Helped Saudis Spy on Khashoggi, Lawsuit Says". en-US. In: *The New York Times* (Dec. 2018). ISSN: 0362-4331. URL: https://www.nytimes.com/2018/12/02/world/middleeast/saudi-khashoggi-spyware-israel.html (visited on 02/20/2022).

[35] Forensic Architecture, Amnesty International, and The Citizen Lab. *Digital Violence: How the NSO Group Enables State Terror*. https://digitalviolence.org/, accessed Feb 2nd, 2022.

[36] Forensic Architecture. *Digital Violence: How The NSO Group Enables State Terror (About the Project)*. https://forensic-architecture.org/investigation/digital-violence-how-the-nso-group-enables-state-terror/, accessed March 10th, 2022.

[37] Forbidden Stories. *Journalists under surveillance*. Website. https://forbiddenstories.org/pegasus-journalists-under-surveillance/, accessed Feb 2nd, 2022.

[38] *Homepage of snscrape lirary*. https://github.com/JustAnotherArchivist/snscrape, accessed Feb 10th, 2022.

[39] *Homepage of facebook-scraper lirary*. https://github.com/kevinzg/facebook-scraper, accessed Feb 10th, 2022.

[40] *Requests: HTTP for Humans*. https://docs.python-requests.org/en/latest/.

[41] *Beautiful Soup*. https://www.crummy.com/software/BeautifulSoup/.

[42] VeNoMouS. *Cloudscraper: A Python module to bypass Cloudflare's anti-bot page*. https://github.com/venomous/cloudscraper.

[43] *Pandas*. https://pandas.pydata.org/.

[44] *Matplotlib: Visualization with Python*. https://matplotlib.org/.

[45] *Punjabi Language - Wikipedia*. https://en.wikipedia.org/wiki/Punjabi_language, accessed Feb 2nd, 2022.

[46] *TextBlob: Simplified Text Processing*. https://textblob.readthedocs.io/en/dev/.

[47] *sentiment-spanish*. https://pypi.org/project/sentiment-analysis-spanish/.

[48] Forbidden Stories. *Carlos Ketohou*. https://forbiddenstories.org/journaliste/carlos-ketohou/, accessed Feb 9th, 2022.

[49] Forbidden Stories. *Lenaïg Bredoux*. https://forbiddenstories.org/journaliste/lenaig-bredoux/, accessed Feb 10th, 2022.

[50] Forbidden Stories. *Edwy Plenel*. https://forbiddenstories.org/journaliste/edwy-plenel/, accessed Feb 10th, 2022.

[51] L. Bredoux. *Tweet by @LenaBred*. https://twitter.com/LenaBred/status/1417035149372448768, accessed Feb 10th, 2022.

[52] Forbidden Stories. *Sushant Singh*. https://forbiddenstories.org/journaliste/sushant-singh/, accessed Feb 12th, 2022.

[53] *Forbidden stories page for Ferdinand Ayite*. https://forbiddenstories.org/journaliste/ferdinand-ayite/, accessed Feb 10th, 2022.

[54] *Forbidden stories page for Taoufik Boucharine*. https://forbiddenstories.org/journaliste/taoufik-bouachrine/, accessed Feb 10th, 2022.

[55] *Forbidden stories page for Vijaita Singh*. https://forbiddenstories.org/journaliste/vijaita-singh/, accessed Feb 10th, 2022.

[56] *Forbidden stories page for Siddharth Varadarajan*. https://forbiddenstories.org/journaliste/siddharth-varadarajan/, accessed Feb 10th, 2022.

[57] *Wiki page for United Progressive Alliance (UPA)*. https://en.wikipedia.org/wiki/United_Progressive_Alliance, accessed Feb 10th, 2022.

[58] *Wiki page for National Democratic Alliance (NDA)*. https://en.wikipedia.org/wiki/National_Democratic_Alliance, accessed Feb 10th, 2022.

[59] *Forbidden stories page for Maria Moukrim*. https://forbiddenstories.org/journaliste/maria-moukrim/, accessed Feb 10th, 2022.

[60] *Forbidden stories page for Edwy Plenel*. https://forbiddenstories.org/journaliste/edwy-plenel/, accessed Feb 10th, 2022.

[61] *Forbidden stories page for Lenaig Bredoux*. https://forbiddenstories.org/journaliste/lenaig-bredoux/, accessed Feb 10th, 2022.

[62] *Forbidden stories page for Hicham Mansouri*. https://forbiddenstories.org/journaliste/hicham-mansouri/, accessed Feb 10th, 2022.

[63] *Forbidden stories page for Ignacio Cembrero*. https://forbiddenstories.org/journaliste/ignacio-cembrero/, accessed Feb 10th, 2022.

[64] *Forbidden stories page for Carmen Aristegui*. https://forbiddenstories.org/journaliste/carmen-aristegui/, accessed Feb 10th, 2022.

[65] *Forbidden stories page for Rafael Rodriguez Castaneda*. https://forbiddenstories.org/journaliste/rafael-rodriguez-castaneda/, accessed Feb 10th, 2022.

[66] *Forbidden stories page for Alejandro Sicairos*. https://forbiddenstories.org/journaliste/alejandro-sicairos/, accessed Feb 10th, 2022.

[67] *Forbidden stories page for Alvaro Delgado*. https://forbiddenstories.org/journaliste/alvaro-delgado/, accessed Feb 10th, 2022.

[68] *Forbidden stories page for Marcela Turati*. https://forbiddenstories.org/journaliste/marcela-turati/, accessed Feb 10th, 2022.

[69] *Forbidden stories page for MK Venu*. https://forbiddenstories.org/journaliste/mk-venu/, accessed Feb 10th, 2022.

[70] *Forbidden stories page for Swati Chaturvedi*. https://forbiddenstories.org/journaliste/swati-chaturvedi/, accessed Feb 10th, 2022.

[71] F. A. (Facebook). https://www.facebook.com/permalink.php?story_fbid=1351096718684578&id=132851133842482&__cft__[0]=AZUdooQJohfw5K4427GbXUH0qwsVsgm9fahub-oPE83nWpZ5Bs6_qPNRabgKq5nx5O3aw_iUFa91rSoTUcjscoPNsxJg0fLg2QtFDSmun4qo8O0-_uBrh9se3ggIf3iQmvZv4xZljUO&__tn__=%2C0%2CP-R, accessed March 10th, 2022.

[72] Forbidden Stories. *About Us - Forbidden Stories*. https://forbiddenstories.org/about-us/, accessed Feb 2nd, 2022.

[73] Forbidden Stories. *About The Pegasus Project*. https://forbiddenstories.org/about-the-pegasus-project/, accessed Feb 2nd, 2022.

[74] S. Kirchgaessner, P. Lewis, D. Pegg, S. Cutler, N. Lakhani, and M. Safi. *Revealed: leak uncovers global abuse of cyber-surveillance weapon*. https://www.theguardian.com/world/2021/jul/18/revealed-leak-uncovers-global-abuse-of-cyber-surveillance-weapon-nso-group-pegasus, accessed Feb 2nd, 2022. 2021.

[75] David Pegg, Sam Cutler. *What is Pegasus spyware and how does it hack phones?* https://www.theguardian.com/news/2021/jul/18/what-is-pegasus-spyware-and-how-does-it-hack-phones, accessed Feb 10th, 2022.

[76] *Forbidden stories page for Jorge Carrasco*. https://forbiddenstories.org/journaliste/jorge-carrasco/, accessed Feb 10th, 2022.

[77] *Forbidden stories page for Alejandra Xanic Von Betrab*. https://forbiddenstories.org/journaliste/alejandra-xanic-von-betrab/, accessed Feb 10th, 2022.

[78] *Forbidden stories page for Rohini Singh*. https://forbiddenstories.org/journaliste/rohini-singh/, accessed Feb 10th, 2022.

[79] Nicole Perlroth. *Invasive Spyware's Odd Targets: Mexican Advocates of Soda Tax: Foreign Desk*. Late Edition (East Coast). New York, N.Y: New York Times Company, 2017. ISBN: 0362-4331; 0362-4331.

[80] *Pegasus: The new global weapon for silencing journalists | Forbidden Stories*. en-US. URL: https://forbiddenstories.org/pegasus-the-new-global-weapon-for-silencing-journalists/ (visited on 11/09/2021).

[81] S. Woodhams. "Spyware: An Unregulated and Escalating Threat to Independent Media". en. In: *m e d i a d e v* (), p. 23.

[82] J. Penney. *Understanding Chilling Effects*. en. Tech. rep. Toronto and Cambridge: Ctizen Lab at the University of Toronto and Berkman Klein Center for Internet & Society at Harvard University, May 2021. URL: https://papers.ssrn.com/abstract=3855619 (visited on 11/14/2021).

[83] C. Khoo, K. Robertson, and R. Deibert. "Installing fear: A Canadian legal and policy analysis of using, developing, and selling smartphone spyware and stalkerware applications". In: (2019).

[84] A. Mare. "A qualitative analysis of how Investigative Journalists, Civic Activists, Lawyers and Academics are adapting to and resisting communications surveillance in South Africa". In: *Media Policy and Democracy Project* (2016). URL: https://www.mediaanddemocracy.com/uploads/1/6/5/7/16577624/duncan_2_comm_surveillance.pdf (visited on 11/09/2021).

[85] M. Pîrvu. "THE DEGRADATION OF HUMAN RIGHTS AND FREE PRESS THROUGH THE PEGASUS SOFTWARE IN THE ERA OF SURVEILLANCE, AS A THREAT TO INTERNATIONAL SECURITY. A DEBATE OF CIVIL LIBERTIES AND CENSORSHIP". en. In: *STRATEGIES XXI: The Complex and Dynamic Nature of the Security Environment* (2021), pp. 263–272. ISSN: 2668-7828. DOI: 10.53477/2668-6511-22-29. URL: https://revista.unap.ro/index.php/XXI_CSSAS/article/view/1375 (visited on 02/19/2022).

[86] D. Harkin and A. Molnar. "Operating-System Design and Its Implications for Victims of Family Violence: The Comparative Threat of Smart Phone Spyware for Android Versus iPhone Users". en. In: *Violence Against Women* 27.6-7 (May 2021), pp. 851–875. ISSN: 1077-8012, 1552-8448. DOI: 10.1177/1077801220923731. URL: http://journals.sagepub.com/doi/10.1177/1077801220923731 (visited on 02/19/2022).

[87] J. Scott-Railton, B. Marczak, B. A. Razzak, M. Crete-Nishihata, and R. Deibert. *Director of Mexican Anti-Corruption Group Targeted with NSO Spyware*. Tech. rep. Citizen Lab Research Report No. 99. University of Toronto, Aug. 2017. URL: https://citizenlab.ca/2017/08/nso-spyware-mexico-corruption/ (visited on 02/19/2022).

[88] B. Marczak, J. Scott-Railton, N. Al-Jizawi, S. Anstis, and R. Deibert. *The Great iPwn: Journalists Hacked with Suspected NSO Group iMessage 'Zero-Click' Exploit*. Tech. rep. Citizen Lab Research Report No. 135. University of Toronto, Dec. 2020. URL: https://citizenlab.ca/2020/12/the-great-ipwn-journalists-hacked-with-suspected-nso-group-imessage-zero-click-exploit/ (visited on 02/19/2022).

[89] B. Marczak, J. Scott-Railton, B. A. Razzak, N. Al-Jizawi, S. Anstis, K. Berdan, and R. Deibert. *Pegasus vs. Predator: Dissident's Doubly-Infected iPhone Reveals Cytrox Mercenary Spyware*. Tech. rep. Citizen Lab Research Report No. 147. University of Toronto, Dec. 2021. URL: https://citizenlab.ca/2021/12/pegasus-vs-predator-dissidents-doubly-infected-iphone-reveals-cytrox-mercenary-spyware/ (visited on 02/20/2022).