An Analysis of the 2024 Web Monetization Landscape

Sophia Lichterfeld[1] and Kyle Hogan[2]

[1]The Winsor School, Boston, MA
[2]Massachusetts Institute of Technology, Cambridge, MA

Abstract

W3C's Web Monetization (WM) API offers users the ability to compensate content creators online by continuously streaming micropayments to the website owner while viewing a page. While WM could be a feasible alternative to advertisements or subscriptions, it has not yet been widely adopted by websites.

Rates of WM adoption were tracked from 2019 to 2021 but have not been evaluated for the past several years. To implement WM, website owners must add a meta tag or link with a payment pointer directing the money to their online wallet into their page's HTML head. Using the presence of the meta tag or link as an indicator of WM adoption, we built a web scraper to determine the current WM adoption rate in 2024. To expand our adoption rate results, we analyzed a dataset curated by HTTP Archive through Google's BigQuery database. We further assessed the breakdown of wallet providers, the distribution of website hosts, and the comparison of these metrics across time points and subsets of the dataset. We hope our findings will fill this data gap and better inform approaches to increasing widespread WM adoption.

Keywords: *Web Monetization; Micropayments; Solidity; Digital advertising; Decentralized Finance (DeFi); Payment Pointer; World Wide Web Consortium (W3C); HTTP Archive; Cloudflare*

The authors declare no conflicts of interest.

Acknowledgements

I would like to sincerely thank the MIT PRIMES program for providing me with the opportunity to pursue this research.

Contents

1 Introduction

Advertisement schemes dominate the current online financial landscape; indeed, the 2024 digital advertisement market is estimated to reach over 740 billion US dollars, growing from nearly 680 billion US dollars in 2023 [16]. However, ad-based systems are often irritating to users, acting as an obstacle to accessing or enjoying the content they seek. Moreover, as user preference data is collected to tailor ads, ads pose a risk to user privacy. These factors build a strong argument to explore alternate solutions for content creators to generate revenue online.

A substitute for ads that has been widely employed is a subscription model, often used by video streaming services or news sources. Nonetheless, subscription models have been challenging for small businesses to set up and popularize as they require more complicated management on both the owner's and the user's side.

In this paper, we discuss the adoption trends of a disparate option for online content compensation, namely the World Wide Web Consortium's (W3C) Web Monetization (WM) API. We evaluate the advantages it provides to both users and website owners in terms of accessing content, financial gain, and privacy. We simultaneously discuss the obstacles WM has encountered in its adoption, including the general unfamiliarity with online payment schemes that are required for WM implementation. Finally, we analyze and characterize the changes in WM adoption from 2019 to 2024, specifically focusing on drastic differences between 2021, when the last documented study of WM adoption rates was conducted, and our current WM environment.

2 Background

W3C's WM API is an alternative to ad or subscription-based models for compensating online content creators [20]. Users pay for their viewership by continuously streaming micropayments, transactions of fractions of a cent, directly to the owner's online wallet while actively on a page. WM hopes to establish a decentralized, accessible, and standardized model for users to give content creators adequate financial rewards for their efforts [21].

2.1 W3C's WM API: Requirements

To adopt web monetization, users and website owners must establish several key components. These elements, including the Interledger Protocol, WM Agent, and payment pointer, are essential to WM's functioning as they are necessary to create and process payments between users and the website owner. As prerequisites for WM usage, they can also be considered indicators of WM adoption.

2.1.1 Decentralized and Interoperable Transfers with the Interledger Protocol

Both parties must create online wallets, referred to as the WM Provider on the user's end and the WM Receiver on the owner's end, that are interoperable through the Interledger Protocol, an "open" online system that permits the transfer of money through distinct financial systems. While transactions between separate banking institutions can be complicated due to the lack of reestablished trust in these financial interactions, the Interledger Protocol greatly simplifies transfers of this type by temporarily withholding the money from either side until both have declared and confirmed that their payment is complete. The user's payment information for their Interledger wallet is securely stored such that they are not requested to sign in anew to their account each time WM is used. Through Interledger, transactions can be made regardless of geographical location or currency, paving the pathway toward a more decentralized online financial environment [9].

2.1.2 Payment Pointers as Indicators of Adoption

In addition, the website owner must create a payment pointer, a concise string that shares the syntax of a URL and can be resolved to direct payments to an online wallet address. Payment pointers are secure and do not reveal any private information. Website owners must add a payment pointer into the HTML head of their website in a specified format, either as part of a monetization meta tag or monetization link, as shown in Figure 1. This monetization line must currently be manually added by website owners to their site's HTML head [7]. As payment pointers, and, thereby, the monetization meta tag or link, are a necessary requirement for WM adoption, the presence of either of these in the website's HTML head can be taken as an indicator of WM adoption for the site.

Figure 1: WM link format specification

2.1.3 WM Agent as a Mediator

To employ WM, users must also add a browser extension to act as their WM Agent. When the user opens a new site, the WM Agent automatically checks for WM adoption on the site, searching specifically for the monetization meta tag or link and permitting a three-second grace interval to identify WM components before loading ads if those components are missing. To begin WM, the WM Agent follows the payment pointer in the monetization meta tag or link and queries for the unique Interledger wallet address, or payment destination endpoint, indicated by the payment pointer. It then connects the WM Provider with the WM Receiver as parsed through the payment pointer. Next, the WM Agent generates a Universal Unique Identifier (UUID) that acts as the session ID for the user's current visit and payment session to the site. Then, the WM Agent calculates the WM streaming rate according to the amount the website owner would earn with ad revenue; this streaming rate can be adjusted by the user as well. Throughout the user's viewership of the site, micropayments continuously flow at this streaming rate from the WM Receiver to the WM Provider until they become inactive on or exit the site and the payment session is closed [21].

2.2 W3C's WM API: Benefits

As compared to advertisements or subscriptions, WM offers several key benefits, both on the user's and on the owner's side. For users, WM replaces ad revenue, so they no longer encounter ads that distract from or delay the intended content of the page. Moreover, ad-based systems often pose a risk of infringing upon user privacy by collecting data about their preferences and interests to curate more suitable and individualized ads; this threat can be mitigated by transitioning to WM. Similarly, users will not encounter paywalls for subscription-based websites and forgo the need to manage numerous complicated subscription systems [21].

On the owner's end, WM ensures that content creators will earn at least an equal amount of revenue as they would with ads; the default streaming rate (the amount in micropayments sent per second) of WM is equal to the rate of return the website owner would receive with ads, and should a user choose to decrease their WM streaming rate, the number of ads shown to the user automatically adjusts inversely proportionally to compensate for the reduction of WM revenue. Should a user choose to increase beyond the default streaming rate that equalizes WM profits with ad revenue for the website owner, they could get access to premium content; owners can strategically leverage this effect to incentivize higher rates with the promise of additional features. Finally, WM provides a more practical and efficient pathway for small websites, which often face challenges in both setting up and popularizing their subscription models because users hesitate to maintain more than a handful of subscriptions at a time. Therefore, WM can be considered a mutually advantageous solution for online content compensation [21].

Nonetheless, despite these apparent gains, WM has experienced several critical barriers in its widespread adoption. To facilitate the transfer of micropayments, WM operates with cryptocurrencies such as Bitcoin (BTC) or Ethereum (ETH) [1] and requires both parties to create online wallets that are interoperable through the Interledger Protocol [9]. WM's reliance on online banking and cryptocurrencies may be an obstacle for potential users who are unfamiliar with these types of banking systems.

Furthermore, WM falls to a cycle of adoption hesitance where users are unsure how often they will be able to apply WM and, hence, choose not to undergo the process of installing the necessary elements for WM usage, and website owners, conversely, are afraid that the effort they expend to implement WM will not be worthwhile because there are not enough users who will appreciate it. As a result, both sides are stuck in a circular pattern of reluctance to fully adopt WM.

3 Approach and Implementation

We attempted several approaches to accurately estimate the rate of WM adoption on the web. Initially, we intended to build our own web scraper using the Scrapy framework in Python [14]. We also leveraged HTTP Archive's Google BigQuery dataset for

data science analysis in Python and SQL [3].

3.1 Building our own Web Scraper

Web scraping is the automated extraction of data or information from websites. Web scraping bots, often known as spiders, visit sites and access or search for specific data in the site's HTML [17].

Algorithm 1 Scrapy WM Web Scraper

Input: *URLList* ← list of URLs for scraper to search one by one
Notation: *item* ← one scraper entry after scraping the site, recorded in a new line in the resulting CSV file
Notation: *tag* ← the site's monetization meta tag or link, if found

```
for URL in URLList do
    if URL.request == failure then
        item.status = failure.value
        item.url = URL    ▷ scraper encountered an error when trying to access the site; record URL along with type of error
        item.dateExtracted = datetime.datetime.now                                  ▷ recording date scraping attempted
        item.monetized = None                                                       ▷ recording unknown monetization status
        item.pointer = None                                                         ▷ recording unknown payment pointer on site
        item.provider = None                                                        ▷ recording unknown wallet provider
    else
        tag = response.xpath("//meta[@name='monetization']")    ▷ searching for monetization meta tag/link in HTML head
        item.status = response.status
        item.url = URL                          ▷ scraper received a response from the site; record URL with exit status code
        item.dateExtracted = datetime.datetime.now                                  ▷ recording date of information access
        if tag != None then
            item.monetized = 1                                                      ▷ recording that a monetization meta tag/link was found
            item.pointer = tag                                                      ▷ recording the entire tag found
            item.provider = tag.provider                                            ▷ recording wallet provider as extracted from full tag
        else
            item.monetized = 0                                                      ▷ recording that a monetization meta tag/link was not found
            item.pointer = -1                                                       ▷ recording that payment pointer was absent
            item.provider = -1                                                      ▷ recording that wallet provider was absent
        end if
    end if
     yield item
end for
```

To create our WM spider, we used Scrapy, a Python system for web scraping spider development. With each crawl, we gave our spider a list of start URLs that it processed one at a time. On every site, the spider attempted to identify the monetization meta tag or link within the HTML head (based on the request path we had given it of "//meta[@name='monetization']"). Finally, we had our scraper save all data collected from the site, including the date and time of scraping, the exit or error status, whether or not the site was monetized, the payment pointer and wallet provider if the site was monetized, along with the URL of the site, in a CSV file.

3.2 Politeness and Ethical Considerations

One of our foremost priorities when running our spider was respecting the guidelines of polite web scraping. These principles include not retrieving personal or copyrighted data, not overloading the server or website, identifying one's scraper and its purpose, and, most notably, following the website's robots.txt file [2].

The robots.txt file exists for every website and can be viewed by adding "/robots.txt" to the end of a URL. It contains the Robots Exclusion Protocol, which specifies which portions of the site should not be accessed by robots by marking them with "Disallow." Moreover, the robots.txt file can suggest a preferable delay time between crawl request processing to prevent a high

throughput of robot requests that slow down the site for actual, ad-revenue-generating viewers. Adhering to the robots.txt file is considered polite and ethical; thus, we ensured to program our bot to obey this file on all sites [22].

3.3 Scraper Operation

After building our scraper, we first tested it on a random sample of 10,000 sites from Cloudflare's top one million most popular sites from April 1, 2024 [4]. Yet over one third of the websites our scraper checked returned errors and could not be scraped. Of these errors, nearly half were DNS Lookup Errors, while another large portion were Robots.txt Forbidden errors 2. As we were maintaining polite web scraping protocols, we would have been unable to address the latter type of error, although certain robots from reputable organizations that are recognized as "good bots" may be able to overcome and be exempt from those restrictions. However, we wanted to further inspect the subset of websites that were yielding DNS Lookup Errors and attempt to resolve them.

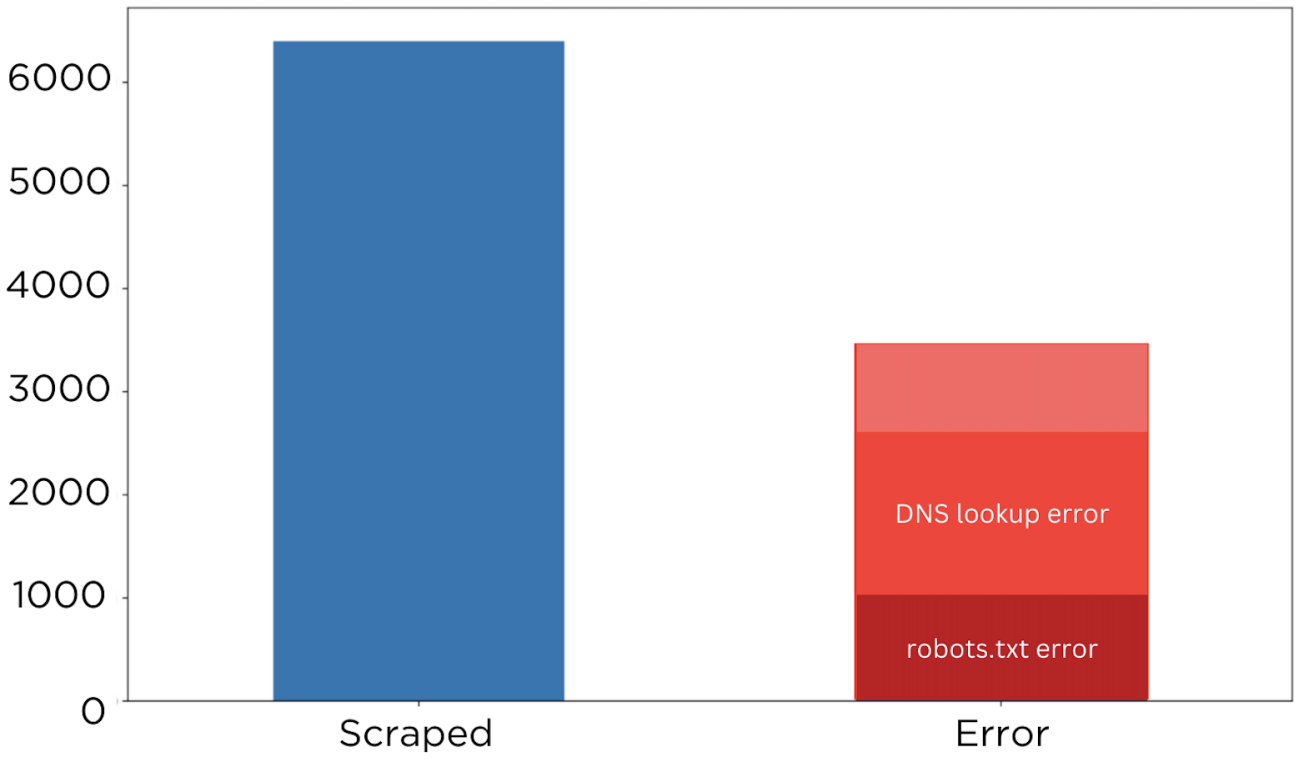


Figure 2: Breakdown of Scraper Exit Status for Random 10,000 Sites

To classify these sites, we employed Cloudflare's top 100 site list, which provides a site categorization based on the kind of services the site is responsible for. Among the top 100 sites, the scraper encountered 38 DNS Lookup Errors. When we examined the category of the sites that were unsuccessfully scraped due to a DNS Lookup Error, we found that a large majority of them were content servers, indicating that these sites are not user-facing and would not contribute to WM adoption assessments (Figure 3). Furthermore, sites with DNS Lookup Errors were also not accessible when we cross-checked the results by attempting to manually open the sites in our browsers. Nevertheless, due to the large proportion of error responses we received overall from our scraper, we instead chose to explore other potential methods to measure the rate of WM adoption more accurately and comprehensively.

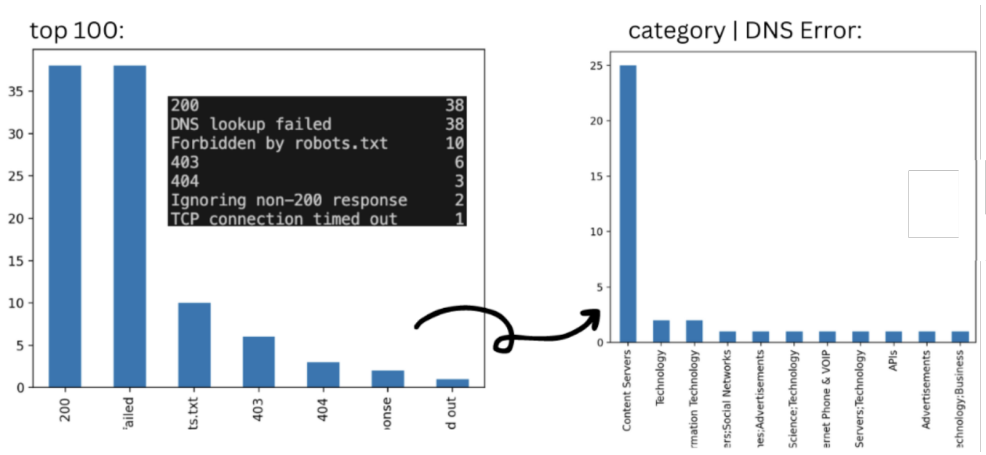


Figure 3: Breakdown of Scraper Exit Status for Top 100 Sites and Site Category for Sites with DNS Lookup Error

3.4 HTTP Archive's State of the Web Project

The most recent analysis of WM adoption was performed by the HTTP Archive [19] in July 2021 as part of their annual State of the Web project [18]. However, WM adoption analysis was not included in the State of the Web report for 2022, and the project as a whole was put on hold for 2023.

It is important to note that the analyses included in the State of the Web report are predominantly volunteers; thus, we suspect those members who had previously been responsible for and had conducted the WM portion of the survey no longer chose to contribute, resulting in this data gap over the past three years.

Nonetheless, when we checked HTTP Archive's BigQuery dataset, we found that they had still continuously been recording whether each site had either a monetization meta tag or monetization link. These data were recorded on the first of every month. We were able to use similar SQL code that the HTTP Archive had published on their GitHub for WM adoption calculations in the past on their more recent data collections [10]; thereby, we ensured that our methodology was consistent with the one used for their previous calculations. Specifically, we both assessed the overall WM adoption percentage in each time point we considered as well as compiled a list of web monetized sites for our two bounding dates, July 2021, when the last HTTP Archive analysis was performed, and April 2024, the most recent dataset published.

4 Results

We investigated several questions around WM using the HTTP Archive BigQuery dataset. This dataset chooses the most popular user websites based on the Chrome User Experience Report's assessment on the first day of each month [6]. Over the time frame analyzed, the dataset grew from 3,946,347 desktop sites in April 2019 to 12,951,215 sites in April 2024, plateauing at above 12.2 million for every time point after October 2022.

4.1 WM API Overall Adoption Trend

The first analysis we performed examined the change in WM adoption rate at six-month time intervals from 04/2019, when 0 sites in the HTTP Archive dataset were recorded to have WM, to 04/2024, the most recent time point. Since WM can be adopted with either a meta tag or a link in the website's HTML head, we recorded the percent prevalence of each of these adoption methods separately, as well as the overall adoption rate of sites having either. The red dashed line in Figure 4 indicates the date at which HTTP Archive's WM adoption percentage analyses paused. Up until that point in July 2021, our trend stayed consistent with their reported findings of approximately 0.0031% in October 2019, 0.0091% adoption in October 2020, and 0.023% adoption in July 2021.

Beyond that date, we observed that WM adoption rates grew gradually and then spiked to their documented peak of 0.047% in October 2022, where 4850 out of 10,358,534 monetized sites (with either a meta tag or a link) were identified before surprisingly falling sharply to mid-2020 levels of 0.0063%, or merely 815 of 12,922,767 sites, by April 2023 (Figure 4).

Algorithm 2 BigQuery WM Adoption Percentage Analysis

Input: *websites* ← HTTP Archive dataset list of websites scraped, where *website* is a single entry in this dataset
Notation: *distinctURLs* ← total number of websites in HTTP Archive dataset
meta ← number of sites with monetization meta tag
link ← number of sites with monetization link
either ← number of sites with either monetization meta tag or link

```
for website in websites do
    if date == '2024-04-01' then                                  ▷ considering single time period
        distinctURLs++                                            ▷ add to URLs searched count total
        if website includes feature 'HTMLMetaElementMonetization' then
            meta++                                                ▷ monetization meta tag was found on this site; add to meta count total
        end if
        if website includes feature 'HTMLLinkElementMonetization' then
            link++                                                ▷ monetization link was found on this site; add to link count total
        end if
        if website includes feature 'HTMLMetaElementMonetization' OR 'HTMLLinkElementMonetization' then
            either++                                              ▷ monetization adoption was found on this site; add to either count total
        end if
    end if
end for
metaPercent = meta / distinctURLs                                 ▷ calculating adoption percentages
linkPercent = link / distinctURLs
eitherPercent = either / distinctURLs
```

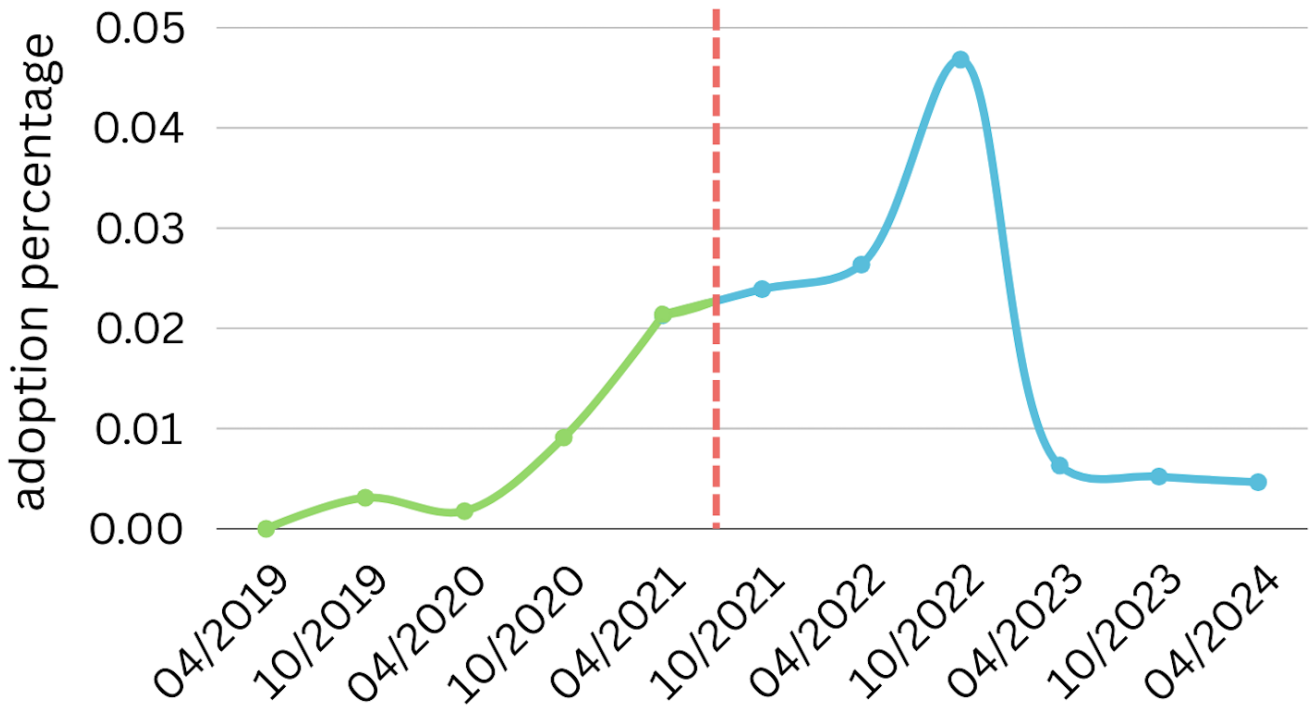


Figure 4: WM Adoption Percentage for Every Six-Month Time Interval from April 2019 to April 2024

This precipitous drop was highly unexpected to us; thus, we chose to inspect this short time frame more closely. As the HTTP Archive dataset is generated on the first of every month, we calculated WM adoption rates for each of the months between October 2022 and May 2023. During this time period, we found that there had been two waves of WM removal, first in October 2022 and again in March 2023 (Figure 5). Between October and November of 2022, WM adoption rates were cut by more than half, decreasing from 0.047% to 0.021%. While there was a slight rise to 0.022% by February 2023, between March and April 2023, WM adoption once again fell, this time from 0.020% to 0.0063%.

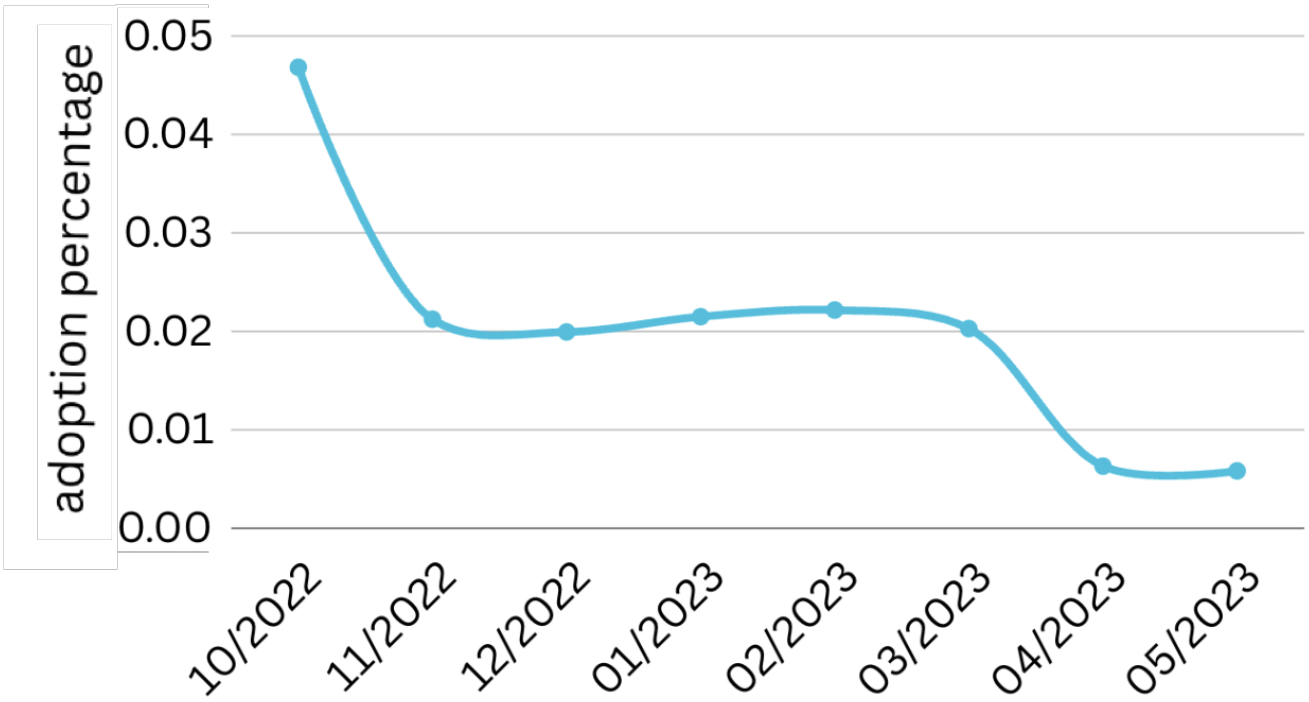


Figure 5: WM Adoption Percentage on the First of Every Month from October 2022 to May 2023

4.2 Result Cross-Checking with Scraper

As these findings were astonishing to us, we first wanted to confirm that the trend we saw was not merely an artifact of the data but rather a reflection of a real pattern and an indication of external forces at play. For this purpose, we generated the complete list of sites in the 2021 BigQuery dataset that were monetized and the complete list of sites in the 2024 BigQuery dataset that were monetized. We then used our web scraper to test the current monetization status of these sites to determine whether it agrees with the data from the HTTP Archive.

As anticipated and in agreement with our previous result, only about one-quarter of the sites that had been monetized in 2021 according to the BigQuery dataset still retained either a monetization tag or link, while the vast majority were no longer adopting WM; this finding suggests that most sites actively chose to remove their monetization status. On the other hand, when running our scraper on the list of sites BigQuery had determined were monetized in 2024, nearly all of them were marked as monetized when cross-checked by our scraper as well. In conclusion, the results obtained from our scraper agree with the web monetization data from the HTTP Archive and can be taken to signify that the overall adoption trend observed accurately characterizes the reality that WM has been removed from a large portion of sites (Figure 6).

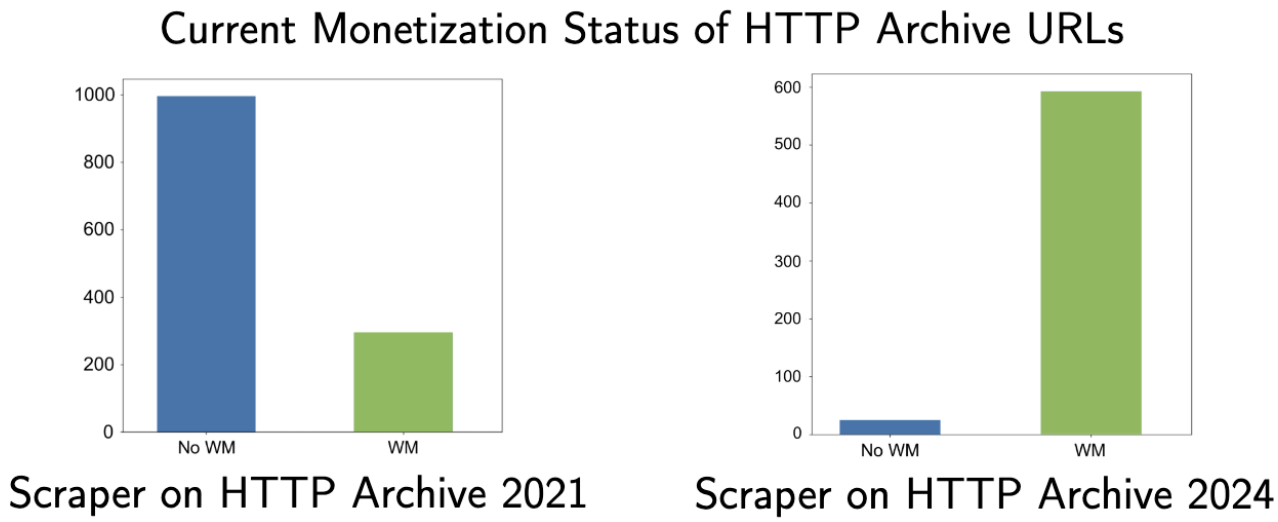
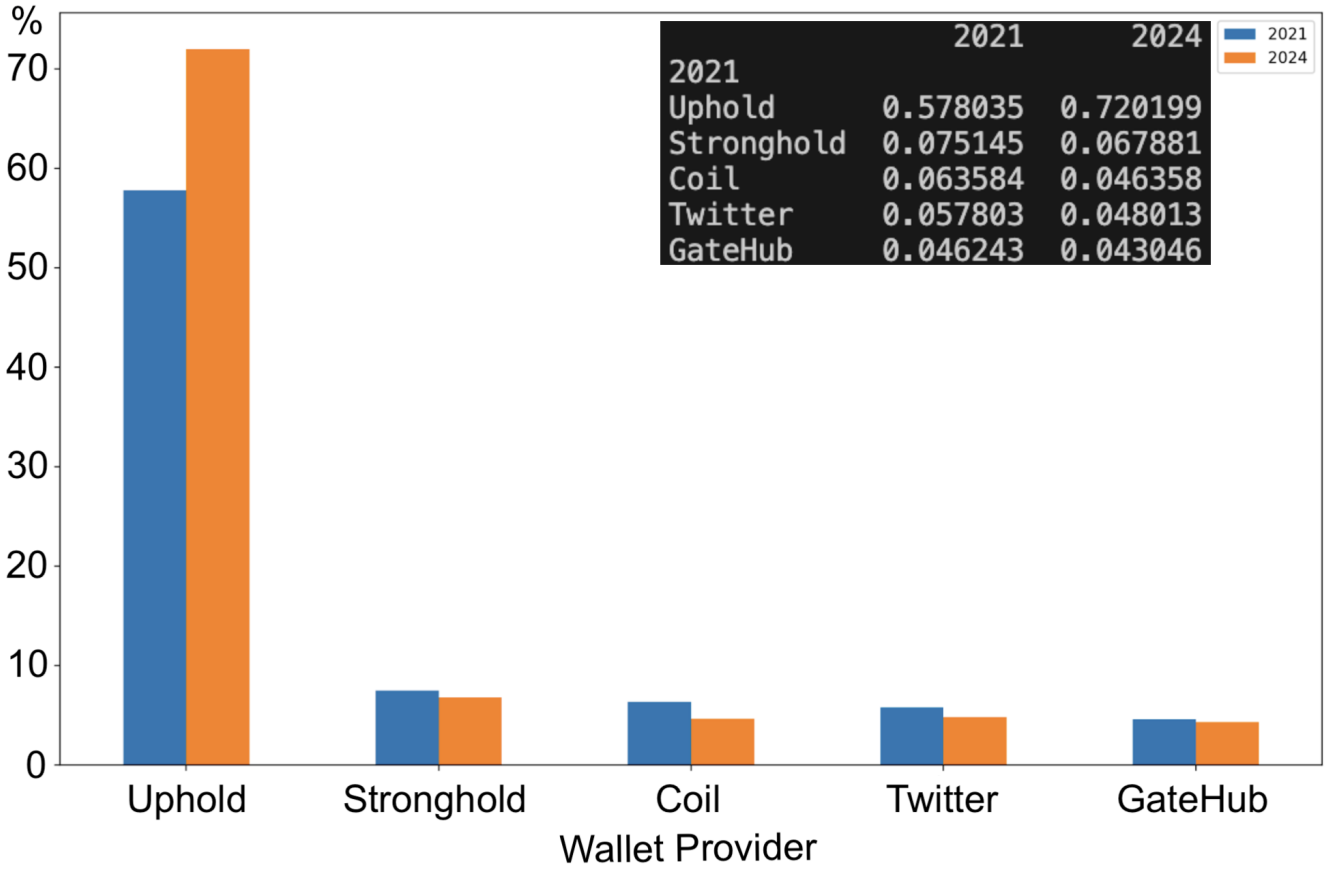


Figure 6: Monetization Status as Determined by WM Scraper on HTTP Archive Monetized Sites from 2021 and 2024

4.3 Wallet Provider Breakdown



	2021	2024
2021		
Uphold	0.578035	0.720199
Stronghold	0.075145	0.067881
Coil	0.063584	0.046358
Twitter	0.057803	0.048013
GateHub	0.046243	0.043046

Figure 7: Breakdown of Wallet Providers for Monetized Sites in 2021 and 2024

Having established that WM adoption experienced an abrupt decrease between October 2022 and April 2023, we next wondered whether the observed trend could be due to the dropping out of a major wallet provider, inactivating many online wallets crucial to WM adoption and functionality. As our own web scraper recorded the wallet provider of each monetized site scraped, we were able to generate a histogram of wallet providers after running our scraper on the HTTP Archive list of monetized sites for 2021 and the HTTP Archive list of monetized sites for 2024 that we had previously generated.

In this analysis, we found that the top five largest wallet providers, Uphold, Stronghold, Coil, Twitter, and GateHub, had persisted through this time frame with only small variations in their user-share breakdowns (Figure 7). These findings portray that the significant drop in overall WM adoption cannot be attributed to changes in wallet provider accessibility or popularity, and the question remained of what extrinsic factor could be affecting WM adoption rates.

4.4 Website Hosting

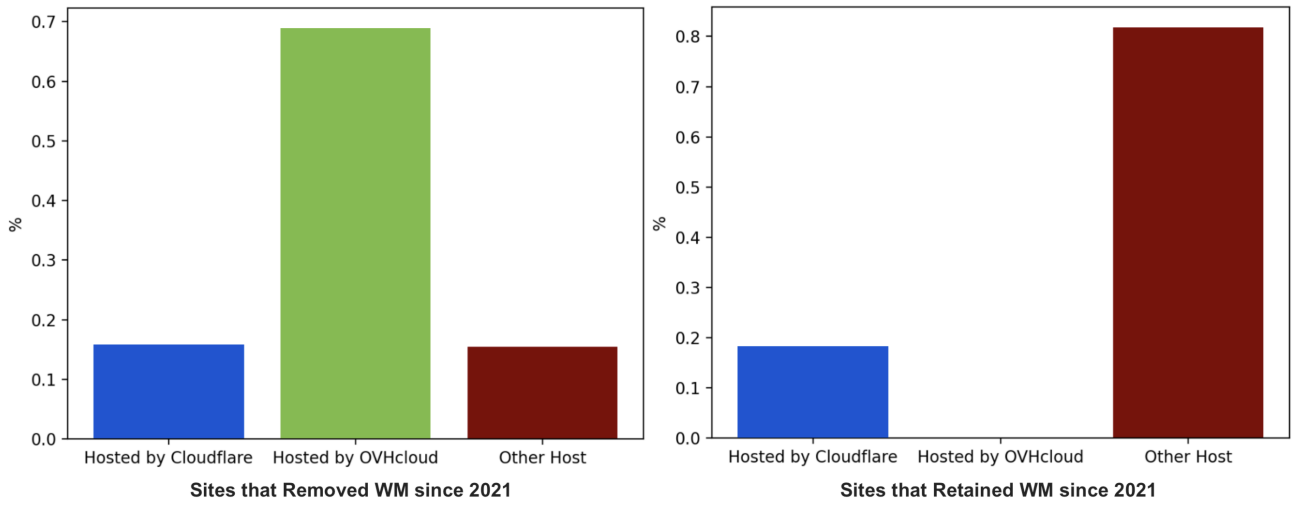


Figure 8: Breakdown of Website Hosts for Sites that Removed vs. Retained WM since 2021

Since our findings do not suggest wallet provider type as the catalyzing factor for the observed rapid decline in WM adoption, we sought an alternate explanation. As the decrease was so sudden, we suspected that it was unlikely websites had coordinated a removal all at once; instead, we considered the website hosting provider as a potential component contributing to this drop.

To investigate that possibility, we compared the percentage of websites hosted by two major hosting providers, Cloudflare [5] and OVHcloud [12], between sites that had removed WM and sites that had retained WM from 2021. From the HTTP Archive list of web monetized sites in 2021, we sorted out and categorized the sites into two distinct lists based on their current WM status as determined by our scraper. Then, we separately ran pydig queries on each site in these two lists to obtain the IP addresses of the sites' hosting providers [13]. We checked these against the IP ranges listed for Cloudflare [8] and OVHcloud [11]. For sites that had removed WM, 15.74% were hosted by Cloudflare, and 68.80% were hosted by OVHcloud. The percentage of sites that still had WM hosted by Cloudflare stayed relatively similar at 18.25%; however, by contrast, there were no sites hosted by OVHcloud that still had WM (Figure 8).

Further examining the OVHcloud hosted sites, we noticed that these sites predominantly shared the domain name "newgrounds.com" in their URLs with different subdomains that appeared to be usernames for each account on the site (alice.newgrounds.com and bob.newgrounds.com would both be included in this WM tally, for instance).

Ergo, "newgrounds.com" had been multi-counted thousands of times; since the original adoption percentages from the HTTP Archive had employed this type of counting methodology, we had followed the same one in our continued trend analysis and inadvertently been multi-counting "newgrounds.com" as well. Whether this counting technique is favorable is up for debate: it is possible that, after the entire domain "newgrounds.com" allowed for WM to be implemented on its sites, each user was required to proactively opt-in by setting up a wallet account and several key components themselves for each subdomain; however, it appears evident that the domain owner still maintained a degree of control over this feature since the WM tag was eliminated at seemingly one time point. This finding signifies that it may have made more sense to only count the entire domain once.

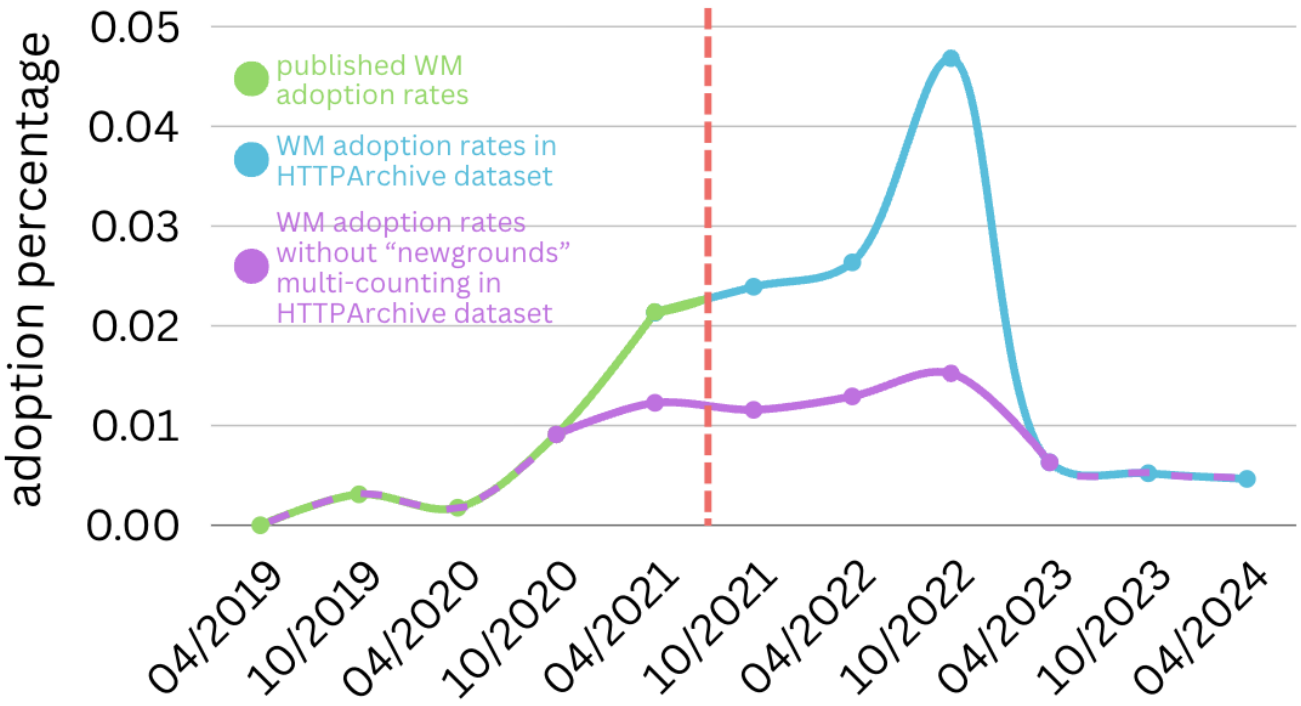


Figure 9: WM Adoption Percentage Trend Adjusted for Multi-Counting of "newgrounds" Domain

With this observation in mind, we returned to the HTTP Archive dataset and filtered out sites under the "newgrounds" domain, correcting these URLs to only be counted as one site. Based on those numbers, we computed new WM adoption rates, as plotted by the purple line in Figure 9. The first instances of "newgrounds.com" began appearing in October 2020, and the last ones had been removed by April 2023. Although there is still a slight drop between October 2022 and April 2023, it is far less extreme than the one previously noted, decreasing from 0.015% to 0.0063% adoption. However, based on these data, this six-month interval still saw the WM adoption percentage cut by more than half, raising further questions about determinants that may have influenced this trend. As all sites hosted by OVHcloud that we found in this dataset were of the domain "newgrounds.com," it remains unclear whether this enduring decline is due to the removal of WM from other OVHcloud-hosted sites (if the website provider OVHcloud no longer supported WM) or the result of other domains being multi-counted as "newgrounds.com" was.

5 Conclusions

Based on the results from both our own WM Scraper and the HTTP Archive's Dataset, we found that WM adoption rates had sharply decreased from 0.047% in October of 2022 to merely 0.0063% by April 2023. When we focused in on this time period, we observed that this decline had come in two waves, one between October and November of 2022 and another between March and April of 2023. We confirmed this trend by running the WM Scraper we built on the list of web-monetized sites from 2021; in that analysis, we found that WM only remained on approximately one quarter of the sites that had implemented it in 2021 while it had been removed from the vast majority of sites.

To explore the potential causes of this downturn, we characterized the breakdown of wallet providers, comparing 2021 to 2024 sites. We found that, between these two time points, the proportion of sites using each of the top five wallet providers stayed relatively stable, with Uphold, the largest wallet provider, gaining slightly while the other four top wallet providers decreased minimally. Thus, we concluded that the trend we observed could not be attributed to the withdrawal of a wallet provider. Moreover, we evaluated whether the cause of this trend could be the termination of WM support by a large website host; specifically, we concentrated on Cloudflare and OVHcloud. Our data showed that, when comparing sites that had removed WM since 2021 and sites that had maintained it, the proportion hosted by Cloudflare was relatively similar, yet the proportion hosted by OVHcloud differed highly significantly. Indeed, while over two-thirds of sites that had removed WM since 2021

were hosted by OVHcloud, not a single site that retained WM was. While these results could suggest that the website hosting provider OVHcloud no longer supported WM, resulting in this widespread drop, in this analysis process, we found that the domain "newgrounds.com" accounted for the entirety of OVHcloud-hosted sites and had been multi-counted for various user subdomains. When we eliminated this multi-counting from our adoption percentage analysis, we saw a much more gradual and smoother trend; nonetheless, we note that the drop between October 2022 and April 2023, although diminished, has not been altogether eliminated, posing itself as a continued area for inquiry and investigation.

6 Discussions and Future Work

These findings raise several intriguing questions for further exploration. Primarily, we hope to investigate the root cause of the drop in WM adoption that has yet to be wholly accounted for; we suspect that external or third-party decisions were at play in this sudden decrease. The two factors that we have evaluated thus far, walled provider availability and website hosting, have proven partially conclusive, though even when adjusting for the differences revealed, a portion of the drop is still present. Hence, we hope to continue searching for potential explanations for this unexpected trend.

Furthermore, we hope to consider possible methods by which we could counteract the pattern we have observed and other alternatives to ads or subscription models for online content compensation. In 2023, our research focused on developing a scheme to financially incentivize the adoption of WM on the owner's side [15]. Simultaneously, we are interested in finding new accessible forms for website owners to generate revenue that would promote a decentralized financial environment and protect user privacy, just as WM hoped to do.

References

[1] A. Case, "The Current State of Micropayments and Web Monetization," *Medium* (July 22, 2021). https://caseorganic.medium.com/part-ii-the-current-state-of-micropayments-and-web-monetization-50ee2e58d332. Accessed June 4, 2024.

[2] A. Datsenko, "The Etiquette of Web Scraping and How to Use Web Scraping Legally," *Webbiquity* (June 16, 2022). https://webbiquity.com/marketing-technology/the-etiquette-of-web-scraping-and-how-to-use-web-scraping-legally/?doing_wp_cron=1700498673.2903699874877929687500. Accessed June 5, 2024.

[3] "BigQuery," *Google Cloud*. cloud.google.com/bigquery. Accessed June 5, 2024.

[4] "Domain Rankings," *Cloudflare Radar* (June 3, 2024). https://radar.cloudflare.com/domains. Accessed June 5, 2024.

[5] "Connect, protect and build everywhere," *Cloudflare* (2024). https://www.cloudflare.com/. Accessed June 20, 2024.

[6] "Frequently Asked Question," *HTTPArchive*. https://httparchive.org/faq#how-does-the-http-archive-decide-which-urls-to-test. Accessed June 4, 2024.

[7] "HTML <link> rel=monetization," *Web Monetization*. https://webmonetization.org/docs/references/html-link-rel-monetization/#:~:text=Payment%20pointers%20and%20meta%20elements,-The%20first%20version&text=Now%2C%20Web%20Monetization%20only%20supports,replace%20the%20%24%20with%20https%3A%2F%2F%20.. Accessed June 5, 2024.

[8] "IP Ranges," *Cloudflare* (Sep 28, 2023). https://www.cloudflare.com/ips/. Accessed June 20, 2024.

[9] "Interledger Protocol (ILP): Web Monetization," *Interledger* (2023). https://interledger.org/rfcs/0028-web-monetization/. Accessed Jun. 30, 2023.

[10] K. Farrugia, "monetization.sql," *GitHub* (Nov 22, 2021). https://github.com/HTTPArchive/almanac.httparchive.org/blob/9ba79b1c8c4169cc1efddc19f33927884ea8b91d/sql/2021/markup/monetization.sql. Accessed June 4, 2024.

[11] "Our IP address services," *OVHcloud* (2024). https://us.ovhcloud.com/hosted-private-cloud/vmware/ip/. Accessed June 20, 2024.

[12] "OVHcloud," *OVHcloud* (2024). https://us.ovhcloud.com/. Accessed June 20, 2024.

[13] "pydig 0.4.0," *PyPI* (2024). https://pypi.org/project/pydig/. Accessed June 20, 2024.

[14] "Scrapy," *Scrapy*. https://scrapy.org/. Accessed June 4, 2024.

[15] S. Lichterfeld, G. Rastogi, and K. Hogan, "Leveraging the Escrow-Holding Abilities of Ethereum Smart Contracts to Incentivize Account Creation for the Widespread Adoption of Web Monetization Schemes," *MIT Mathematics* (Jan 16, 2024). https://math.mit.edu/research/highschool/primes/papers.html. Accessed June 18, 2024.

[16] Statista Market Insights, "Digital Advertising - Worldwide," *Statista* (Nov 2023). https://www.statista.com/outlook/dmo/digital-advertising/worldwide#ad-spending. Accessed June 18, 2024.

[17] S. Wu, "Web Scraping Basics," *Medium* (July 15, 2020). https://towardsdatascience.com/web-scraping-basics-82f8b5acd45c. Accessed June 5, 2024.

[18] "The 2021 Web Almanac," *Web Almanac by HTTP Archive* (2024). https://almanac.httparchive.org/en/2021/. Accessed June 20, 2024.

[19] "The HTTP Archive," *HTTP Archive* (2024). https://httparchive.org/. Accessed June 20, 2024.

[20] Web Incubator Community Group, "Web Monetization," *GitHub* (2024). https://github.com/WICG/webmonetization. Accessed June 5, 2024.

[21] "Web Monetization Explainer | Web Monetization," *WebMonetization.org* (2024). https://webmonetization.org/docs/explainer/. Accessed June 5, 2024.

[22] "What is robots.txt? | How a robots.txt file works," *Cloudflare*. https://www.cloudflare.com/learning/bots/what-is-robots-txt/. Accessed June 5, 2024.