On the Infrastructure Providers that Support Misinformation Websites

Catherine Han
Stanford University
cathan@stanford.edu

Deepak Kumar
Stanford University
kumarde@stanford.edu

Zakir Durumeric
Stanford University
zakird@stanford.edu

Abstract

In this paper, we analyze the service providers that power 440 misinformation sites, including hosting platforms, domain registrars, DDoS protection companies, advertising networks, donation processors, and e-mail providers. We find that several providers are disproportionately responsible for hosting misinformation websites compared to mainstream websites. Most prominently, Cloudflare offers DDoS protection to 34.3% of misinformation sites while servicing only 17.9% of mainstream websites in our corpus. While many mainstream providers continue to service misinformation sites, we show that when misinformation and other abusive websites are removed by hosting providers, DDoS protection services, and registrars, these sites nearly always resurface after finding alternative providers. More encouragingly, we show that misinformation sites also disproportionately rely on popular ad networks and donation processors, but that anecdotally, sites struggle to remain online when mainstream monetization channels are severed. We conclude with insights for infrastructure providers and researchers to consider for stopping the spread of online misinformation.

Introduction

Technical infrastructure providers like Amazon, Cloudflare, and Google both support and regulate websites that spread misinformation. To counter the spread of misinformation, influential platforms have extended their service agreements to prohibit misinformation and, in extreme cases, terminated service to violating sites (Cox 2021; Wong 2019; Infostormer 2019). For example, in 2017, the *Daily Stormer* lost its DDoS protection from Cloudflare and was subsequently cut off from registrar providers GoDaddy and Google, resulting in a website hiatus (Burch 2017). Similarly, in 2021, the "far-right alternative to Twitter," Parler, was knocked offline for a month after Apple and Google removed Parler from their app stores and Amazon cut off Parler's hosting services. Yet, despite their increasing role in policing content, there has been little attention paid to identifying the infrastructure providers that support misinformation websites more broadly.

In this paper, we investigate the technical infrastructure that powers misinformation websites, including domain registrars, web hosting and email providers, online advertising partners, and Distributed Denial of Service (DDoS) protection providers. We specifically seek to: (1) identify the service providers that misinformation sites disproportionately rely on and (2) analyze whether the act of deplatforming misinformation websites affects their long-term availability. To answer these questions, we crawl and analyze the network dependencies of 440 misinformation websites from the OpenSources dataset (OpenSources 2017). We compare these misinformation sites to a baseline of mainstream sites. We crawl each website and collect its DOM, cookies, and network requests, which we then augment with hosting and registrar data. To understand how misinformation websites monetize, we map third-party web dependencies to known advertising providers and payment processors.

We show that misinformation sites disproportionately rely on several hosting providers, most prominently Cloudflare, which serves content for 34.3% of misinformation sites compared to 17.9% of mainstream sites. In manually investigating each misinformation website, we find that sites prefer Cloudflare because of its lax acceptable use policies and its free DDoS protection services that help protect against vigilante attacks. Misinformation websites also disproportionately rely on other mainstream providers including GoDaddy, Unified Layer, and Automattic because of their WordPress offerings that allow users to quickly setup and scale sites without much technical expertise. By manually analyzing past deplatforming events, we find that when major sites are deplatformed by mainstream providers, they nearly always find new homes on alternative providers who actively ignore site content, similar to how bullet-proof hosting providers are utilized by Internet attackers.

Next, we investigate monetization platforms like online advertisement providers and payment processors that enable revenue collection for misinformation sites. Most misinformation sites rely on mainstream ad networks like Google's DoubleClick, but also disproportionately depend on otherwise niche providers including RevContent and Outbrain. Misinformation sites also disproportionately rely on donations through PayPal and Patreon, as well as direct cryptocurrency donations. Despite little evidence showing that deplatforming by hosting providers is effective at keeping such sites offline, we note that anecdotally, websites cease to produce more misinformation content after they are deplatformed from both ad providers and payment processors. In

other cases, sites have lamented the decrease in site revenue after being deplatformed from mainstream ad providers, instead soliciting users for direct donations.

We conclude with a discussion of different strategies for preventing the spread of misinformation based on our results. We argue that while deplatforming sites from hosting infrastructure is likely not an effective solution for curbing the spread of misinformation, targeting site monetization may be a promising approach for combating misinformation. We hope that by shedding light on what has anecdotally been most effective, we encourage providers—particularly those who have announced that they are committed to fighting online abuse and misinformation—to further explore monetization as a critical channel for curbing the spread of misinformation at scale.

Related Work

Our work is inspired by research that highlights the growing complexities of the web. Prior work has studied how websites have grown in complexity (Butkiewicz, Madhyastha, and Sekar 2011; Nikiforakis et al. 2012; Englehardt and Narayanan 2016; Kumar et al. 2017) and are increasingly relying on centralized network entities and third-party content (Kumar et al. 2017). Beyond this, several studies have leveraged the nuances of technical infrastructure to better understand and combat traditional computer abuse, including spam and scams (Hao et al. 2009; 2016) and phishing (Ho et al. 2017; 2019).

In the context of misinformation, much work has focused on the classification of websites, primarily through content analysis (Rashkin et al. 2017; Kumar, West, and Leskovec 2016) or social graph features (Nguyen et al. 2020; Jin et al. 2014; Popat et al. 2017; Shu, Wang, and Liu 2019). Studies have leveraged infrastructure properties (e.g., HTTPS configuration or domain expiration) to classify misinformation sites (Hounsel et al. 2020), but these studies do not consider web resources broadly as features.

Most closely aligned with our work are several recent studies of the web infrastructure components of misinformation sites. Zeng et al. investigated the ads and ad platforms that power mainstream and misinformation sites, finding that although some advertisers are more prevalent on misinformation sites, both categories share similar fractions of problematic advertising content (Zeng, Kohno, and Roesner 2020). Similarly, Agarwal et al. explored the web trackers on hyper-partisan, biased websites. They found that right-leaning, hyper-partisan sites track users more aggressively and rely on many third-party networks (e.g., Doubleclick, Taboola, AdNexus) to function (Agarwal et al. 2020).

Methodology

Our study investigates the technical infrastructure that supports misinformation websites, including web hosting, domain registration, DDoS protection, online ads, and payment processing. In this section, we describe the set of misinformation websites we analyze and how we collect data about the providers that support each website.

Misinformation Websites. We analyze two sets of misinformation websites. In this context, we deem misinformation to be non-satirical websites that have potentially misleading content (e.g., "fake news"), determined by the OpenSources project (OpenSources 2017). OpenSources publishes lists of known, vetted, labeled misleading websites by analyzing sites across several axes: (1) domain name, (2) "About Us" page, (3) article source, (4) writing style, (5) page and image aesthetic, and (6) social media network; the set has been used extensively in prior research (Zeng, Kohno, and Roesner 2020; Hounsel et al. 2020; Budak 2019; Sharma et al. 2019). While the OpenSources master list contains 826 websites, this list was published in 2017, and because of this, many of these sites are unavailable today. Thus, we removed 191 unreachable sites and 123 parked domains (e.g., those that pointed back to a domain registrar). As our objective is to exclude satirical sites, two independent researchers then manually coded the remaining websites to identify 72 satirical websites, based on the "About pages" of each website in question. Our final misinformation corpus contains 440 websites. The misinformation corpus spans several flavors of unreliability, according to the labels provided by OpenSources. For example, some sites consistently present extreme bias (e.g., breitbart.com), peddle conspiracy theories or bigoted propaganda (e.g., infowars.com, barenakedislam.com), or promote junk science (e.g., naturalnews.com).

Additionally, we analyzed 128 QAnon websites.Separate two-sample proportion tests comparing the site resources and hosting providers of our QAnon corpus to our misinformation corpus indicate that the two sets of sites were drawn from different distributions. As such, we analyze the two data sets separately. Any results describe the OpenSources websites unless noted otherwise.

Mainstream Website Sample. We analyze 10K random sites from the Alexa Top Million (Alexa Internet, Inc. 2020) as a control set. We analyze a random sample of websites rather than mainstream news sites because mainstream news sites differ significantly in structure compared to sites similar in scale to those in our misinformation corpus. For instance, mainstream news sites are often extremely well-managed and optimized compared to smaller websites that are similar in scale to those in our misinformation corpus. No mainstream news sites use content management systems like WordPress that are common amongst smaller websites. We note that 154 (35%) of our misinformation sites appear in the Alexa Top Million; we exclude these from our mainstream corpus.

Data Collection. We crawled each website using Crawlium, a crawler based on headless Chrome (Arshad 2020). We visit each website using a fresh browser instance with no cookies. To identify potentially hidden resources or infrastructure on each website, we additionally spider to four first-party links on the same domain. For each page, we allot 10 seconds to navigate to the URL and wait 10 seconds for dynamic content to load. We then collect (1) the page DOM, (2) cookies, and (3) logs of network events. We crawled 52K pages from December 2020 to January 2021.

AS	% Misinfo	% Mainstream	p-value	Effect Size
Cloudflare	34.3%	17.9%	6.2e-16	0.34
GoDaddy	6.6%	2.6%	6.9e-07	0.08
Unified Layer	5%	2.1%	3.8e-06	0.06
Liquid Web	2.7%	0.9%	1.6e-05	0.04
Sucuri	1.8%	0.4%	1.9e-05	0.03

Table 1: **ASes Disproportionately Hosting Misinformation**—The results of a two-sample proportion test of ASes sorted by effect size. We find that Cloudflare disproportionately represented on misinformation sites.

AS Owner	Sites	% Mis.	AS Owner	Sites	% Mis.
Cloudflare	151	34.3%	Liquid Web	12	2.7%
GoDaddy.com	29	6.6%	OVH SAS	12	2.7%
Google	29	6.6%	DigitalOcean	12	2.7%
Amazon.com	22	5%	Fastly	10	2.3%
Unified Layer	22	5%	Automatticc	9	2%

Table 2: **Top Misinformation ASes**—We show the top ten ASes responsible for hosting misinformation sites and the portion of misinformation sites for which each is responsible. We find that Cloudflare has the largest market share.

Resource Analysis. For each website, we construct an inclusion tree, a structure derived from a webpage's DOM that represents the sequence of resource loads that fetch site content. We annotate each resource with the origin AS from which it is loaded. We use the AS of the root page to determine each site's web hosting and DDoS protection provider. To understand ad providers, we analyzed images larger than a 1×1 pixel loaded by advertising providers listed by WhoTracks.me (WhoTracks.Me 2021). One limitation of this approach is that while we restricted our ad detection method to observe images served from ad domains excluding tracking pixels, other images like Facebook's "Like" button are still counted toward Facebook's presence as an ad provider on a site. Finally, we performed a WHOIS lookup on each domain to determine domain registratrar and an MX lookup to identify e-mail provider. We will release our code and dataset at publication time.

Ethical Considerations. We visit each site in our study five times. While this is negligible load for widely-known websites, there are ethical considerations at play as there are with any active scanning. We followed the best practices defined by Durumeric et al. and refer to their work for more detailed discussion of the ethics of active network research (Durumeric, Wustrow, and Halderman 2013). We do not block ads loading because they are an element of our study, but we never click or interact with ads. We argue that we do not significantly impact the misinformation ecosystem along two axes: (1) we do not meaningfully contribute to site traffic in a way that may negatively affect the site itself, and (2) we only negligibly contribute to the ad revenue of misinformation publishers.

Hosting and DDoS Protection

We first consider the primary hosting provider for each website. We note that because many sites are protected by DDoS providers like Cloudflare, in some cases, we can only determine the DDoS provider and not the backend hosting provider. In those instances, we classify the site as being hosted by the DDoS provider since they are the company serving web content to users.

A handful of providers host a disproportionate number of misinformation websites, most notably Cloudflare. Cloudflare provides free CDN and DDoS protection services to sites and is a popular provider across the board, serving nearly 18% of sites in our mainstream sample. However, it also serves the largest fraction of misinformation sites (151 domains, 34.3%) (Table 1). Misinformation sites also disproportionately rely on GoDaddy, Unified Layer, Liquid Web, and Sucuri compared to mainstream sites. To measure this, we conducted a two-sample proportion test, measuring whether the proportion of websites in our misinformation and our mainstream corpus hosted by each hosting provider differed between the two sets. Because we were simultaneously measuring multiple comparisons, we corrected our p-values with Bonferroni corrections $ps < 6.02 \times 10^{-5}$). Given our large sample size, most p-values are statistically significant, so we compute effect size using Cohen's h to better understand the strength of the relationship between these hosting providers and misinformation sites. Our analysis shows that Cloudflare has the largest effect size (0.38, 34.4% of misinformation vs. 17.9% of mainstream sites).

Though not responsible for a disproportionate number of sites, several reputable hosting providers, including Google and Amazon, provide critical infrastructure to dozens of misinformation websites, partially contributing to the global misinformation predicament (Table 2). We also find websites protected by well-known CDNs Akamai (e.g., unclesamsmisguidedchildren.com, an extremist site known for its consistent publication of conspiracy theories) and Fastly (e.g., cnsnews.com, an website known for unreliable claims). Across all providers, we find misinformation served from 90 distinct ASes.

QAnon Providers

We observe a reliance on similar providers for QAnon sites. Cloudflare is the most popular provider (13% of sites), followed by Automattic (11%) and VanwaTech (9%). Automattic is likely popular due providing a zero-touch WordPress platform, which is commonly used to spread QAnon content. A number of smaller players absent from our misinformation corpus are commonly used to host QAnon sites, notably VanwaTech, which hosts 9% of QAnon sites and rose to prominence when it began hosting 8kun (formerly 8chan), an online community associated with extreme content and mass shootings (Wong 2019; Krebs 2020).

Despite its prominence in hosting conspiracy content, VanwaTech also hosts a number of innocuous popular domains. Through passive DNS data, we observe 266 domains that are hosted by VanwaTech, of which 25 sites appear on the Alexa Top Million. A similar provider, DDoS Guard (who previously hosted Parler, 8kun, and sites belong to Hamas) hosts 33K domains total and 561 sites in the Alexa Top 1M. Although IP allocation providers have

Cloudflare Sites	Attack	Cloudflare Migration	
		Date	Post-Attack
barenakedislam.com	2/4/15	8/31/17	✓
drudgereport.com	12/30/16	1/4/17	✓
frontpagemag.com	3/23/15	3/24/15	✓
godlikeproductions.com	4/13/16	8/9/17	✓
naturalnews.com	8/8/17	8/8/17	✓
off-guardian.org	9/26/19	5/6/19	✗
returnofkings.com	9/2/15	10/23/14	✗
russia-insider.com	4/11/18	4/13/18	✓
thegatewaypundit.com	4/15/18	6/12/15	✓
weaselzuppers.us	1/5/15	1/1/14	✗
infostormer.com	12/7/19	8/15/17	✗

Table 3: **DDoS Attacks Against Cloudflare Misinfo. Sites**—Misinformation sites with known DDoS attack history and when they were first observed using Cloudflare hosting in our dataset.

recently taken action against DDoS Guard (Krebs 2021) for potentially operating in a region with no presence, we note that a collective ban on such providers simply because they are amenable to hosting misinformation content may inadvertently harm thousands of non-misinformation websites.

Acceptable Use Policies

In spite of growing concerns regarding misinformation, most hosting providers do not explicitly prohibit hate speech or misinformation. Providers such as GoDaddy, Amazon, Unified Layer, WordPress, and Fastly, do explicitly disavow sites that incite violence, but their ToS/AUP do not extend to hate speech or misinformation (GoDaddy 2020; Amazon Web Services, Inc. 2016; Automattic 2021; Fastly, Inc. 2021; Liquid Web 2021). Two hosting providers, OVH and Digital Ocean, specifically prohibit harassing or abusive content, including racially or ethnically offensive content (OVHcloud 2020; DigitalOcean 2020). In contrast, a handful of companies have taken a counter, "content-neutral" approach, notably Cloudflare, whose ToS simply state that it "cannot remove material from the Internet that is hosted by others" (Cloudflare 2020).

The OpenSources dataset tags each website with additional labels, one of which is whether the website contains hate speech. In our corpus, 30 websites are labeled as hate speech. We find that hateful websites do in fact appear on providers that prohibit the practice. One such website is hosted on OVH (vdare.com), and another by Digital Ocean (actforamerica.org). The most prominent provider among sites specifically serving hateful content is Cloudflare (9 sites, 30%), followed by GoDaddy and Sucuri (3, 10% each).

Cloudflare DDoS Protection

It is difficult to ascertain exactly why misinformation websites prefer Cloudflare over other hosting providers. Cloudflare has only relatively recently emerged as the primary provider for misinformation and abusive websites (Figure 1). Leveraging historical passive DNS data from Farsight (Farsight Security 2021), we find that GoDaddy was the most prevalent provider between 2010–2015, hosting up to 48 (11.2%) misinformation sites as recently as 2015. It was not until October 2015 that Cloudflare overtook GoDaddy.

One explanation for Cloudflare's rise is simply that Cloudflare grew in popularity across the Internet. However, we observe that rate of growth for mainstream sites adoption Cloudflare hosting is approximately half that of misinformation websites (Figure 1). In the end, it is likely due to a confluence of reasons. First, it is likely that misinformation websites turn to Cloudflare due to their lax policies. We observe anecdotal evidence from misinformation sites about their reliance on Cloudflare. For example, AmmoLand, a popular guns rights blog, revered Cloudflare not just for its DDoS protection, but also for its self-described "content-neutral" stance (AmmoLand 2020):

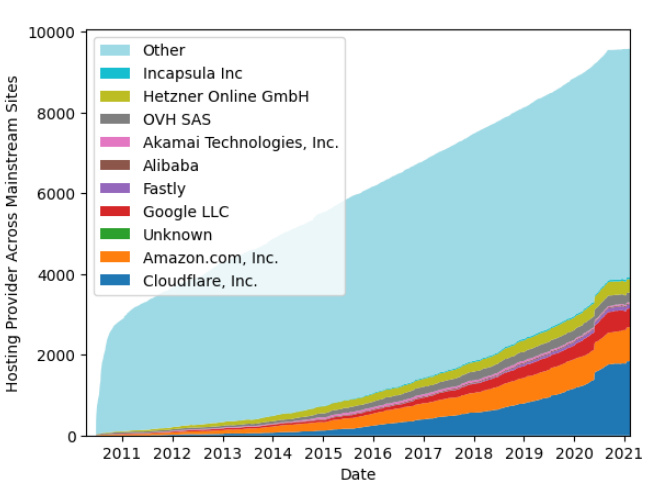
> *"Cloudflare is also pro-free speech and anti-censorship. Prince is a rarity in Silicon Valley. The SPLC and various left-wing organization have called out Cloudflare to stop providing services to websites that host content that they see as objectionable Cloudflare has responded in a way that I wish more companies would return to this type of pressure from SPLC type groups. They ignored the demands. Prince believes it is imperative for our country that his company remains content-neutral."*

Similarly, an author of Infostormer, who previously wrote for the Daily Stormer, empathized with Cloudflare's CEO concerning difficult decisions of "ban-hammering" sites from their service (Infostormer 2019):

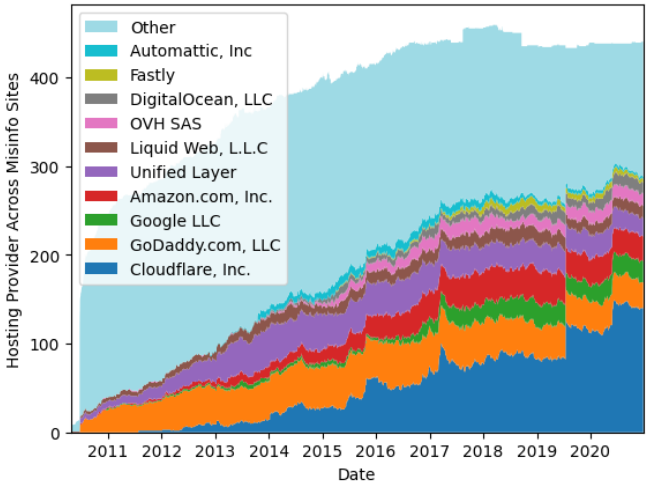
> *"Obviously I don't like what Prince did [banning the Daily Stormer]. I've been highly critical of him over the past few months since he ordered the ban. He justified banning the site because he thought we were "assholes" and happened to be in a bad mood. As a writer for the Daily Stormer I found this comment to be quite offensive. It was also an abandonment of principles. Up until that time, Cloudflare maintained a neutral stance on content. This was the correct position to have. With that said, I can understand that he was put in a tough position. He had to do what he thought was best for the company at that time."*

Second, many sites turn to Cloudflare for its free DDoS protection services because sites are regularly coming under attack from "vigilante justice" groups (BBC 2017; Wong 2019). We manually investigate the 151 misinformation websites hosted by Cloudflare and observe that 23 sites have publicly documented experiencing DDoS attacks. Of those, 11 have specific attack dates. Leveraging Farsight's DNS data, we find that 7 (64%) domains transitioned to Cloudflare after an attack, with 4 (37%) transitioning within days of being attacked (Table 3).

In one example, Natural News (naturalnews.com), a prominent anti-vaccination and conspiracy theory site, came under attack on August 8, 2017. At the time, the site relied on Codero and EasyDNS. Then, on the day that news of the DDoS attack on the website was published, Natural News began its transition to Cloudflare. Similarly, FrontPage Magazine (frontpagemag.com), a site known for its



(a) **Mainstream Hosting over Time**



(b) **Misinformation Hosting over Time**

Figure 1: **Longitudinal Hosting Providers**—Since the beginning of Farsight's historical DNS data, Cloudflare has seen the most growth of any other hosting provider among both misinformation and mainstream sites. However, misinformation websites have grown far more reliant on Cloudflare, which accounts for 34% of misinformation websites compared to just 17% of mainstream websites.

Registrar	Misinfo	Mainstream	p-value	Effect Size
GoDaddy	49%	24%	2.7e-25	0.52
NameCheap	4.5%	0.8%	1e-11	0.25
eNom	5.9%	1.7%	1.5e-8	0.23
Epik	1.7%	0.08%	2.3e-13	0.2
CloudFlare	2.5%	0.6%	8.3e-6	0.17
DNC	1.4%	0.2%	2e-6	0.16
FastDomain	2.8%	1%	0.001	0.14

Table 4: **Registrars Disproportionately Supporting Misinformation**—The results of a two-sample proportion test of domain registrars sorted by effect size. GoDaddy is the most prevalent, likely in part due to its free WordPress integration.

far-right, Islamophobic content, experienced a DDoS attack on March 23, 2015. While the site briefly used Cloudflare in May 2013, it quickly switched to using Rackspace Cloud Service's name servers (stabletransit.com). We do not observe changed DNS data until the day after the attack, March 24, 2015, when the site switched to Cloudflare. Both sites have remained on Cloudflare since coming under attack.

Hosting and Site Generation Bundles

Many misinformation websites rely on free content management tools. For example, WordPress powers 68% of misinformation websites—nearly twice the percentage of mainstream websites. The reliance on free website generation tools likely explains the prevalence of specific providers. For example, GoDaddy heavily advertises free WordPress integration. About 65% of GoDaddy, 72% of Unified Layer, and 90% of Liquid Web websites use WordPress, highlighting the role that ease of use can play in choosing a hosting provider and supporting misinformation more broadly.

Domain Registrars

The misinformation sites in our study rely on 47 domain registrars (Table 4). Sites disproportionately rely on GoDaddy (49% misinformation vs. 24% Alexa sample), NameCheap (4.5% vs. 0.8%), eNom (5.9% vs. 1.7%), Epik (1.7% vs. 0.08%), and Cloudflare (2.5% vs. 0.6%). We visited the abuse reporting pages of each registrar, and find that while all registrars have abuse reporting mechanisms, only one explicitly prohibits misinformation: Tucows.

In spite of a lack of policy, registrars have made ad-hoc decisions to deplatform hateful, violent, or misleading content in the past. Most notably, the Daily Stormer was deplatformed by a series of registrars, including GoDaddy, Google, and Namecheap, which hindered its ability remain online (Robertson and Liptak 2017). Asia Registry, an Australian registrar, booted Gab (an alt-right alternative to Twitter) off of their service in 2017, citing Australian discrimination law (Breland 2017). After losing service, Gab switched to Epik, which serves as the registrar for 1.7% of the domains in our misinformation corpus and is disproportionately relied on by misinformation websites. Epik is primarily known for hosting far-right extremist content, famous for previously offering protection services through its company Bitmitigate to 8chan and Parler (Brodkin 2019; Greenspan 2021). Our results highlight lax registrar policies, but also show that many lesser-known registrars (e.g., Epik) are willing to support abusive websites in the name of a "free and open Internet."

Monetization Strategies

Beyond hosting infrastructure, misinformation sites also rely on online advertising and direct donations to stay online. In this section, we analyze the role monetization plays in supporting misinformation websites.

Ad Tracker	% Sites	Ad Tracker	% Sites
Facebook	23.4	Outbrain	5.0
DoubleClick	21.6	Taboola	3.9
RevContent	14.3	ShareThis	2.0
Google Syndication	9.1	Connatix	2.0
Google	6.1	Amazon Advertising	1.8

Table 5: **Top Advertisers on Misinformation**—The distribution of the top 10 advertising trackers found on misinformation sites. Google constitutes three of the 10 advertising domains.

Ad Provider	Misinfo	Mainstream	p-value	Effect Size
RevContent	14.3%	0.1%	7.8e-18	0.72
DoubleClick	21.6%	4.9%	1.5e-17	0.52
Facebook	23.4%	9.3%	2.0e-12	0.39
Google Syndication	9.1%	9.1%	2.2e-07	0.32
Outbrain	5.0%	0.5%	2.2e-07	0.30

Table 6: **Advertisers Disproportionately Supporting Misinformation**—The top 5 advertising trackers that are found disproportionately often in misinformation sites over mainstream sites, ordered by effect size. All p-values were Bonferroni corrected ($n = 606$) and were statistically significant ($ps < 8.25 \times 10^{-6}$).

Advertising

Online ads continue to be a primary monetization strategy for misinformation websites, despite journalists calling on ad companies to discontinue serving ads on misinformation sites (Silverman, Singer-Vine, and Thuy Vo 2017; Silverman 2017). We find that twice as many misinformation sites use online advertisements as mainstream sites (62.7% vs. 34.9%). We conducted two-sample proportion tests on the prevalence of all ad providers on misinformation and mainstream websites with ads. All comparisons were corrected for multiple testing using the Bonferroni corrections. The results indicate that several ad providers are disproportionately used by misinformation sites, most notably RevContent and Google's DoubleClick (Table 6). While not disproportionately responsible for the monetization of misinformation sites, we also find that major ad companies like Facebook (23.4%) also appear on misinformation sites (Table 5).

RevContent. RevContent had the highest effect size (0.72), indicating that it is most disproportionately used by misinformation websites (14.3% of misinformation websites but only 0.1% of mainstream websites). RevContent was previously admonished by mainstream media outlets for serving ads on fake news sites, and even went as far as launching a *Truth in Media Initiative*, which allows users to report misinformation websites. Despite this, the company continued to place ads on known misinformation sites, and the company later defended their inaction, indicating that while fake news intended to deceive was not allowed on the site, satirical content content was not prohibited (Silverman 2017). We note that we removed satirical sites from our misinformation corpus; 15.9% of sites using RevContent in our study are labeled as junk science sites and 34.9% as conspiracy theory sites.

DoubleClick. Googles's DoubleClick had the second highest effect size (0.52). In response to rising concern over misinformation amidst the 2016 U.S. election, Google released a statement that it would "restrict ad serving on pages that misrepresent, misstate, or conceal information about the publisher, the publisher's content, or the primary purpose of the web property" (Love and Cooke 2016). According to our dataset's site labels, however, of the 95 sites serviced by DoubleClick, 27 (28.4%) are conspiracy sites; 17 (28.4%) are fake news; 9 (9.5%) are junk science.

There is anecdotal evidence that removing ad revenue is effective at curbing the spread of misinformation. In one such instance, Google AdSense deplatformed American Free Press in 2017 for serving anti-Semitic content. The site remains blocked by Google Ads. Today, most ads found on American Free Press are embedded directly into the page as first-party content. We also detect the header bidding library, Prebid.js, on American Free Press, allowing the site to directly offer bid slots to brands. American Free Press has experienced a different fate from that of ZeroHedge, which was deplatformed by Google in June 2019. ZeroHedge's ad monetization was reinstated by Google just one month later after its takedown of problematic comments (Graham 2020). All News Pipeline, a conspiracy theory site, laments the decline of ad revenue for itself and other "independent" media sites (Duclos 2018):

> *"With digital media revenue spiraling downward, especially hitting those in Independent Media, it has become apparent that traditional advertising simply isn't going to fully cover the costs and expenses for many smaller independent websites."*

This hints that ad providers may be able to effectively reduce misinformation driven revenue and site operation. However, we note that sites rely on an average of seven ad providers, underscoring the need for coordinated efforts amongst providers. Unfortunately, this does not appear to be happening in practice. Despite RevContent and Google previously claiming to be curbing misinformation on their platforms (Silverman, Singer-Vine, and Thuy Vo 2017; Silverman 2017; Love and Cooke 2016), our results indicate that their efforts are not effective, and that these organizations still financially support the spread of misinformation. Broadly, we find minimal evidence of ad providers blocking misinformation sites.

Donations

Misinformation websites often also rely on donations to sustain their operations. Donation strategies range from using third-party intermediaries like PayPal to solicit donations to directly accepting cryptocurrencies like Bitcoin. In our corpus, 43 (9.78%) misinformation sites rely on resources from PayPal compared to only 67 (0.01%) mainstream websites. A two-sample proportion test indicates that this difference in proportions is statistically significant ($p < 0.005$, $h = 0.47$): misinformation sites disproportionately rely on PayPal compared to mainstream sites.

To understand why PayPal is disproportionately represented on misinformation sites, we manually investigated

the 43 misinformation sites that loaded resources from PayPal domains. For each of these sites, we examined web pages and banners soliciting donations. We find that 93% (40) of the sites that rely on PayPal use it for donation services, but two links were inactive. The remainder (7%) utilized PayPal for subscriptions or storefronts. The misinformation sites in our investigation also solicit Bitcoin donations (14%), Patreon donations (9.3%), and Salsa Labs donations (4.7%).

PayPal has previously blocked payment on sites hosting hateful and non-inclusive content. One site author that was deplatformed by PayPal is Roosh Valizadeh (Roosh V) known for his support of men's rights and the alt-right. One of his sites, returnofkings.com, is present in our set of misinformation sites. Return of Kings (ROK) announced a hiatus in 2018, identifying deplatforming of monetization strategies (e.g., PayPal and ads) as a successful tactic in removing misinformation online:

> *"The first factor for this hiatus is that site revenues are too low. We've been banned from Paypal and countless ad partners, which forced me to lay off the site editor last year and also lower payments to regular contributors. This started a negative spiral of declining content quality, site traffic, and revenues. Even the beloved comments section, which many see as the highlight of ROK, was badly hit when Disqus banned us. Currently, ROK receives half the traffic of its peak and less than one-fifth of the income"* (Valizadeh 2018).

The Daily Stormer also faced challenges from restricted revenue streams, but remains operational. As a result, it has been forced to rely solely on donations:

> *"We are not allowed to use any form of advertisement. We cannot use PayPal. We cannot even use credit card processors. We had a P.O. box, and even that was shut down. The only way we can receive money is through crypto currency"* (DailyStormer 2021).

Our data indicates a variety of different revenue streams supporting the production of misleading content online, but hints that coordinated deplatforming by both ad providers and payment processors may be an effective way of disincentivizing the continued upkeep of online misinformation.

QAnon Monetization

Monetization strategies on QAnon sites are similar to the OpenSources misinformation sites we analyzed. DoubleClick is the most common ad provider (59%) on QAnon sites. However, we did not find ad providers that were disproportionately represented in QAnon sites when compared to misinformation sites. Many QAnon domains rely on external donations and other monetization strategies to support their online presence: 42% of QAnon sites rely on external donations, primarily through cryptocurrency donations (23% of websites), PayPal (19% of websites) and Patreon (6.3% of websites). QAnon websites also rely on a long tail of other third-party money transfer applications, including Venmo, Square, and Fundly, highlighting the role a wide breadth of financial applications have in supporting QAnon.

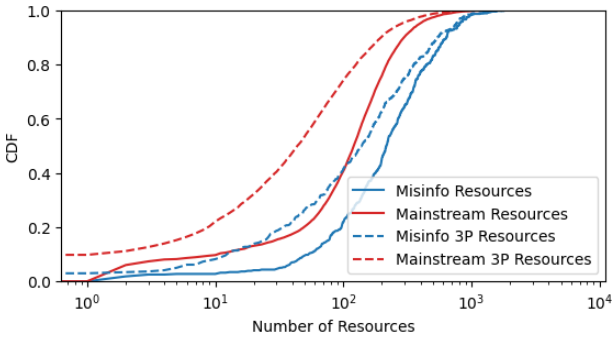


Figure 2: **CDF for Number of Resources**—Misinformation sites load more resources than mainstream sites. A higher proportion of resource loads on misinformation sites are third-party, indicating a heavier reliance on third-party services than mainstream sites.

AS	ASN	Misinfo		Mainstream	
		#	%	#	%
Google	13949	416	94.5%	7973	79.7%
Amazon.com	14618	324	73.6%	3771	37.7%
Cloudflare	13335	308	70%	3804	38%
Fastly	54113	283	64.3%	2396	24%
Akamai	393234	239	54.3%	2293	22.9%
Facebook	32934	228	51.8%	3262	32.6%
AppNexus	36805	199	45.2%	1129	11.3%
Highwinds	11588	198	45%	2138	21.4%
MCI	12199	182	41.4%	1096	11%
Automattic	2635	175	39.8%	858	8.6%

Table 7: **Top Misinformation Third-Party Resource ASes**—The top 10 autonomous systems responsible for third-party resource loads across misinformation sites.

Other Technical Dependencies

Although hosting platforms and monetization sources are the primary dependencies for misinformation sites, sites often rely on a myriad of other technical dependencies like third-party web resources and e-mail providers. In this section, we highlight these other dependencies.

Misinformation websites, which can range from complex news pages to small blogs, load a median of 215 resources, of which 77 (36%) are first-party and 138 (64%) are third-party. Compared to mainstream sites, which rely on a median of 36% third-party resources, misinformation sites more heavily rely on third-party entities (Figure 2). Misinformation sites load the same top third-party resources as mainstream sites (e.g., popular analytics, tracking, and advertising resources). In some cases, misinformation sites do have statistically different proportions: for example, 61% of misinformation websites rely on DoubleClick whereas only 35% of mainstream websites do. However, most differences are marginal. The third-party resources that misinformation sites rely on come from a variety of providers; however, a small handful of providers. Unsurprisingly, misinformation sites depend on resources from major players including Google (95%), Amazon (74%), Cloudflare (70%), Fastly (64.3%), and Akamai (54.3%) (Table 7). Beyond previously

discussed services (e.g., Google ads), large providers also support website in other manners. For example, Google also provides fonts (83% of misinformation websites) and custom search integration (69% of misinformation websites).

Many misinformation sites are also configured to accept inbound email. We find no statistically significant differences in e-mail providers. Similar to mainstream sites, misinformation sites most commonly depend on Microsoft Outlook and Gmail, which serve 32% and 17% of misinformation websites, respectively. However, similarly to analytics, we see evidence of inconsistent policies within companies. Despite being deplatformed by Google News, westernjounalism.com still uses Gmail. We encourage organizations that make deplatforming decisions to consider all products that may be used to support misinformation operations.

Discussion

Our work suggests that while deplatforming misinformation websites has broad impact on their availability, certain strategies for deplatforming may be more effective than others. We discuss potential strategies in this section:

Hosting and Domain Registration

Though many mainstream providers have policies condemning hate and violence, and policies against misinformation sites are beginning to appear, there are many alternative providers available. Similar to how bulletproof hosting providers allows customers to host illegal content, send spam, and launch DDoS attacks (Konte, Perdisci, and Feamster 2015), niche registrars and hosting companies are willing to serve misinformation and abusive content.

Broadly, we find that deplatforming misinformation from hosting providers does not prevent them from remaining online. Websites like the Daily Stormer and Parler found hosting online with VanwaTech and Beelastic, respectively. However, we note that alternative providers tend to be more expensive, and because lesser known misinformation sites have not been deplatformed in practice, it remains unclear whether small sites would be able to afford alternatives.

DDoS Protection

While DDoS protection is not a required component of infrastructure for many mainstream websites, there is a long history of particularly offensive or hateful misinformation sites coming under attack, and empirically DDoS protection is a particularly useful service for these sites. Misinformation websites disproportionately rely on Cloudflare, a provider that offers *free* DDoS protection and has neglected to address abusive content in all but the most egregious cases. We observe steady growth in Cloudflare's popularity across misinformation sites since 2010; and today, Cloudflare is the primary provider for misinformation sites.

This may be due to the absence of free alternative DDoS protection. Misinformation websites have only a few alternatives, many of which are expensive: Bitmitigate, which serviced the Daily Stormer after it was removed from Cloudflare, costs $159 dollars a month for enterprise-level protection, and DDoS-Guard, which is leveraged by several far-right websites in our dataset, costs up to $1,000 a month. While both Cloudflare and DDoS-Guard offer a free tier of protection, DDoS-Guard's free tier only offers protection for attacks with up to 1.5 Tbps compared to Cloudflare's 67 Tbps capacity (DDoS-Guard 2021; Cloudflare 2021). It remains unclear whether smaller DDoS protection providers—especially at analogous free tiers of service—can withstand significant attacks to the same degree.

Monetization

Blocking monetization channels for misinformation sites appears to have the greatest potential for curbing the spread of misinformation. Anecdotally, misinformation websites report significant decreases in revenue, and in some cases stop publishing new content entirely after losing access to advertising and/or donation platforms. For example, the website *Return of Kings* was first deplatformed by PayPal, and eventually shut off by almost all advertising partners. While ads from MGID are still displayed on the site, the decrease in revenue forced the site to announce a hiatus, and no new content has been posted since October 2018. Aligned with prior research in online abuse that suggests that increasing costs reduces harm (Ramachandran and Feamster 2006), we suggest that monetization platforms—both ad providers and payment processors—consider the role they play in supporting online misinformation and how they may be in the best position to curb its spread.

Ethics of Deplatforming

Our paper focuses on understanding the providers that directly or indirectly support misinformation websites and whether deplatforming helps curb the spread of misinformation. It remains an open question whether companies *should* deplatform misinformation sites, and if they do, how they should choose which sites to deplatform. While a few providers have policies that prohibit misinformation, many do not, which may inadvertently enable misinformation websites to thrive on their platforms. We encourage providers to actively consider writing concrete policies around abusive content and misinformation. We also note that several of the largest ad providers have publicly announced their intent to fight online abusive content and misinformation, but based on our data, have failed to take meaningful action against such sites. We encourage providers to reconsider how they are enforcing their policies.

Conclusion

In this paper, we analyzed the infrastructure that powers misinformation websites. We showed that several providers are disproportionately responsible for hosting online misinformation, most prominently Cloudflare, which hosts a third of the misinformation sites in our study. Providers rarely have clauses in their terms of service to prevent misinformation on their platforms, and even when hosting providers do have policies, enforcement is rare and seemingly ineffective. When misinformation websites are deplatformed by hosting providers and registrars, they find other willing providers to serve their content. We do, however, find that misinformation sites disproportionately rely on monetization platforms

like ad networks and donation platforms, and that anecdotally, sites appear to struggle when their monetization channels are removed. We hope our results will inform infrastructure platforms and researchers of more effective strategies to reduce the spread of online misinformation.

References

Agarwal, P.; Joglekar, S.; Papadopoulos, P.; Sastry, N.; and Kourtellis, N. 2020. Stop tracking me Bro! differential tracking of user demographics on hyper-partisan websites. In *World Wide Web Conference*.

Alexa Internet, Inc. 2020. Top 1,000,000 Sites. http://s3.amazonaws.com/alexa-static/top-1m.csv.zip. 2020-10-08.

Amazon Web Services, Inc. 2016. Aws acceptable use policy. https://aws.amazon.com/aup/. Accessed: 2021-01-15.

AmmoLand. 2020. Gun owner privacy. https://www.ammoland.com/tags/gun-owner-privacy/feed/.

Arshad, S. 2020. Crawlium. https://github.com/sajjadium/Crawlium. Accessed: 2020-10-08.

Automattic. 2021. Terms of service. https://wordpress.com/tos/. Accessed: 2021-01-15.

BBC. 2017. Daily stormer: Cloudflare drops neo-nazi site.

Breland, A. 2017. Alt-right twitter rival may lose its web domain. https://thehill.com/policy/technology/351169-alt-right-twitter-rival-might-lose-its-domain.

Brodkin, J. 2019. Dumped by cloudflare, 8chan gets back online—then gets kicked off again. *Ars Technica*.

Budak, C. 2019. What happened? the spread of fake news publisher content during the 2016 u.s. presidential election. In *The World Wide Web Conference*, WWW '19, 139–150. New York, NY, USA: Association for Computing Machinery.

Burch, S. 2017. https://www.thewrap.com/russian-internet-boot-daily-stormer/.

Butkiewicz, M.; Madhyastha, H. V.; and Sekar, V. 2011. Understanding website complexity: measurements, metrics, and implications. In *ACM Internet measurement conference*.

Cloudflare. 2020. Website and online services terms of use. https://www.cloudflare.com/website-terms/.

Cloudflare. 2021. https://www.cloudflare.com/plans/.

Cox, K. 2021. Parler goes dark, sues amazon to demand immediate reinstatement. *Ars Technica*.

DailyStormer. 2021. Support the daily stormer! https://dailystormer.su/contributions/. Accessed: 2021-01-15.

DDoS-Guard. 2021. https://ddos-guard.net/en/store/web.

DigitalOcean. 2020. Acceptable use policy. https://www.digitalocean.com/legal/acceptable-use-policy/.

Duclos, S. 2018. Independent media labeled 'dangerous' & disqus comment notifications listed in gmail as suspicious or spam as big tech finds news ways to attack alternative media websites. https://allnewspipeline.com/Tech_Attacks_On_IM_Ramps_Up.php. Accessed: 2021-01-15.

Durumeric, Z.; Wustrow, E.; and Halderman, J. A. 2013. ZMap: Fast Internet-wide scanning and its security applications. In *22nd USENIX Security Symposium*.

Englehardt, S., and Narayanan, A. 2016. Online tracking: A 1-million-site measurement and analysis. In *ACM SIGSAC conference on computer and communications security*.

Farsight Security. 2021. Introducing dnsdb 2.0. https://www.farsightsecurity.com/solutions/dnsdb. 2021-01-11.

Fastly, Inc. 2021. Acceptable use policy; reporting a violation; and dmca safe harbor. https://www.fastly.com/acceptable-use/. Accessed: 2021-01-15.

GoDaddy. 2020. Godaddy legal agreements and policies. https://www.godaddy.com/legal/agreements.

Graham, M. 2020. Google says zero hedge can run google ads again after removing 'derogatory' comments. https://www.cnbc.com/2020/07/14/google-reinstates-zero-hedge-ad-monetization.html.

Greenspan, R. E. 2021. Parler moves to epik, a domain registrar known for hosting far-right extremist content. https://www.businessinsider.com/parler-moves-to-epik-domain-known-for-hosting-far-right-2021-1.

Hao, S.; Syed, N. A.; Feamster, N.; Gray, A. G.; and Krasser, S. 2009. Detecting spammers with snare: Spatio-temporal network-level automatic reputation engine. In *18th USENIX Security Symposium*.

Hao, S.; Kantchelian, A.; Miller, B.; Paxson, V.; and Feamster, N. 2016. Predator: Proactive recognition and elimination of domain abuse at time-of-registration. In *ACM Conference on Computer and Communications Security*.

Ho, G.; Sharma, A.; Javed, M.; Paxson, V.; and Wagner, D. 2017. Detecting credential spearphishing in enterprise settings. In *26th USENIX Security Symposium*.

Ho, G.; Cidon, A.; Gavish, L.; Schweighauser, M.; Paxson, V.; Savage, S.; Voelker, G. M.; and Wagner, D. 2019. Detecting and characterizing lateral phishing at scale. In *28th USENIX Security Symposium*.

Hounsel, A.; Holland, J.; Kaiser, B.; Borgolte, K.; Feamster, N.; and Mayer, J. 2020. Identifying disinformation websites using infrastructure features. In *USENIX Workshop on Free and Open Communications on the Internet*.

Infostormer. 2019. Infostormer has been under a sustained ddos attack this week. https://infostormer.com/infostormer-has-been-under-a-sustained-ddos-attack-this-week/.

Jin, Z.; Cao, J.; Jiang, Y.-G.; and Zhang, Y. 2014. News credibility evaluation on microblog with a hierarchical propagation model. In *IEEE Intl. Conference on Data Mining*.

Konte, M.; Perdisci, R.; and Feamster, N. 2015. Aswatch: An as reputation system to expose bulletproof hosting ASes. In *ACM Conference on Special Interest Group on Data Communication*.

Krebs, B. 2020. Qanon/8chan sites briefly knocked offline. https://krebsonsecurity.com/2020/10/qanon-8chan-sites-briefly-knocked-offline/.

Krebs, B. 2021. https://krebsonsecurity.com/2021/01/ddos-guard-to-forfeit-internet-space-occupied-by-parler/.

Kumar, D.; Ma, Z.; Durumeric, Z.; Mirian, A.; Mason, J.; Halderman, J. A.; and Bailey, M. 2017. Security challenges in an increasingly tangled web. In *26th Conf. on WWW*.

Kumar, S.; West, R.; and Leskovec, J. 2016. Disinformation on the web: Impact, characteristics, and detection of wikipedia hoaxes. In *Conference on World Wide Web*.

Liquid Web. 2021. Liquid web acceptable use policy ("aup"). https://www.liquidweb.com/about-us/policies/acceptable-use-policy/. Accessed: 2021-01-15.

Love, J., and Cooke, K. 2016. Google, facebook move to restrict ads on fake news sites. *Reuters*.

Nguyen, V.-H.; Sugiyama, K.; Nakov, P.; and Kan, M.-Y. 2020. Fang: Leveraging social context for fake news detection using graph representation. In *ACM Intl. Conference on Information & Knowledge Management*.

Nikiforakis, N.; Invernizzi, L.; Kapravelos, A.; Van Acker, S.; Joosen, W.; Kruegel, C.; Piessens, F.; and Vigna, G. 2012. You are what you include: large-scale evaluation of remote javascript inclusions. In *ACM conference on Computer and communications security*.

OpenSources. 2017. Opensources. https://github.com/bigmclargehuge/opensources. Accessed: 2020-10-08.

OVHcloud. 2020. Terms of service. https://us.ovhcloud.com/legal/terms-of-service. Accessed 2021-01-15.

Popat, K.; Mukherjee, S.; Strötgen, J.; and Weikum, G. 2017. Where the truth lies: Explaining the credibility of emerging claims on the web and social media. In *WWW*.

Ramachandran, A., and Feamster, N. 2006. Understanding the network-level behavior of spammers. In *SIGCOMM*.

Rashkin, H.; Choi, E.; Jang, J. Y.; Volkova, S.; and Choi, Y. 2017. Truth of varying shades: Analyzing language in fake news and political fact-checking. In *Conference on empirical methods in natural language processing*.

Robertson, A., and Liptak, A. 2017. https://www.theverge.com/2017/8/20/16170370/namecheap-host-take-down-neo-nazi-hate-site-daily-stormer.

Sharma, K.; Qian, F.; Jiang, H.; Ruchansky, N.; Zhang, M.; and Liu, Y. 2019. Combating fake news: A survey on identification and mitigation techniques. *ACM Trans. Intell. Syst. Technol.*

Shu, K.; Wang, S.; and Liu, H. 2019. Beyond news contents: The role of social context for fake news detection. In *ACM International Conference on Web Search and Data Mining*.

Silverman, C.; Singer-Vine, J.; and Thuy Vo, L. 2017. In spite of the crackdown, fake news publishers are still earning money from major ad networks. *BuzzFeed News*.

Silverman, C. 2017. An ad network that helps fake news sites earn money is now asking users to report fake news. *BuzzFeed News*.

Valizadeh, R. 2018. Return of kings is going on hiatus. https://www.returnofkings.com/195790/return-of-kings-is-going-on-hiatus. Accessed: 2021-01-15.

WhoTracks.Me. 2021. https://whotracks.me.

Wong, J. C. 2019. 8chan: the far-right website linked to the rise in hate crimes. *The Guardian*.

Zeng, E.; Kohno, T.; and Roesner, F. 2020. Bad news: Clickbait and deceptive ads on news and misinformation websites. In *Workshop on Technology and Consumer Protection*.