AppNexus

Viral Content:
The First Warning Sign of Fraud

Michael Misiewicz
Manager, Data Science

Laura Yu
Associate Data Scientist

EXECUTIVE SUMMARY

As the world's leading independent ad tech platform, AppNexus sees hundreds of billions of impressions each day, across every corner of the digital ecosystem. From this unique vantage point, our data science team has gotten a first-hand look at advertising fraud. This whitepaper explores some of their observations, breaks down the current state of ad fraud throughout digital media as a whole, and offers recommendations to publishers, advertisers, and internet users.

Here are some of the highlights:

Online advertising fraud poses an existential threat to programmatic advertising. In addition to costing advertisers billions of dollars each year, ad fraud undermines brands' confidence in the programmatic marketplace. According to a recent report from the Chief Marketing Officer Council and Dow Jones, 72% of programmatic advertisers are concerned about brand safety and control in the programmatic marketplace.

Traffic acquisition vendors deserve careful scrutiny. Viral content producers are particularly susceptible to ad fraud when they purchase traffic from questionable third-party vendors. In order to eliminate fraud, advertisers and publishers must investigate the traffic acquisition firms in their programmatic supply chains.

Viral content goes hand-in-hand with online advertising fraud. AppNexus' experts have seen that generally, publishers specializing in viral content are much more likely to receive non-human traffic. We have established this link by using data science methods to find suspicious traffic patterns and overlaps in these sites' audiences, as well as similarities in the content they produce. Our analysis strongly suggests that many of these sites are operated by the same people.

"Fake News" and hate speech are often just another form of viral content. We've also seen a connection between viral publishers and the "fake news" and extremist political content that dominated the public conversation in the months prior to and following the 2016 U.S. presidential election. In fact, it appears that many of the ad-supported sites in these categories are likely driven by money as much as by ideology. Regardless of which motivation is stronger, a key goal for most of these sites is to attract lots of cheap traffic in order to maximize profits from ad revenue.

TABLE OF CONTENTS

INTRODUCTION

In digital media, few trends have generated more discussion than the recent rise of "viral" content.

Though these stories come in many shapes and sizes, industry observers generally agree that they share common traits. By and large, a piece of viral content is an easily consumable story designed to be shared on social media. From ideologically-slanted "fake news" to tear-jerker videos of sick dogs, these stories tap into psychological triggers in order to manipulate the audience into a heightened emotional state.

For years, experts have debated the impact viral content has had on the online experience and political discourse around the world. But until now, most of us have overlooked the fact that, in many cases, viral content producers have **become beneficiaries of online advertising fraud**.

Indeed, without a steady stream of visitors coming through the homepage, publishers that rely on viral social traffic experience major peaks and valleys in their revenue patterns. When a story blows up on social media, business is good. But when their content fails to gain traction with social audiences, these publishers are incentivized to buy traffic from third-party sources that aren't always legitimate.

When observing the digital advertising ecosystem, we've found that viral content is now a **warning sign** that a publisher is either knowingly or unknowingly purchasing fraudulent traffic for itself, or redirecting non-human users to other sites to drive towards consistent (and growing) monetization goals. In this whitepaper, we'll explore our findings further, in hopes of providing insight in to what they mean for the digital media industry and society at large.

By reading on, you'll learn what viral content looks like, why publishers produce it, and how bad actors are using it to buy and sell fraudulent traffic. In addition, this whitepaper arms advertisers, publishers, and technology vendors with actionable advice for fighting online ad fraud. Finally, we'll take a look at what we've learned about fake news and hate speech, shedding light on why ad-supported publishers who produce this kind of content might not be motivated entirely by ideology.

But before we do any of that, let's take a step back to review the current state of online ad fraud.

THE STATE
OF AD FRAUD

How It Happens, Where It's Going Next, and What We're Doing To Stop It

At its core, the story of online advertising fraud is a cat-and-mouse game between bad actors and those of us who work to thwart them. Every time that brands, publishers, and vendors implement a new preventative measure, fraudsters develop a still more sophisticated method of stealing ad budgets intended for legitimate human audiences.
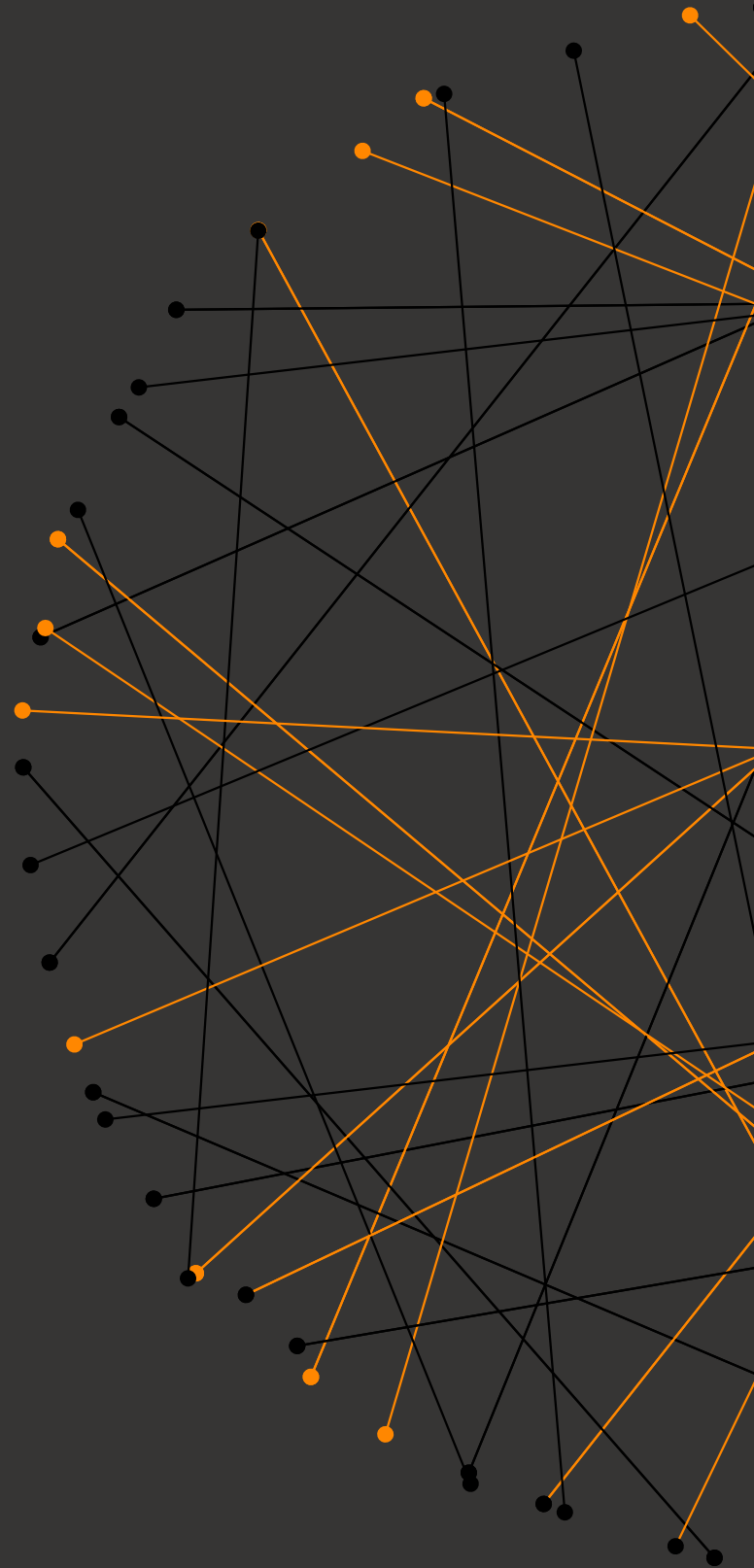
Though fraudsters use a variety of tactics to carry out their theft, the common denominator is that they profit by selling phony web traffic under false pretenses. In many cases, these schemes involve charging brands and/or publishers to access audiences that don't actually exist. While the methods are always changing, popular ad fraud methods today include:

- **Browser or device hijacking programs:** malware that takes over a user's browser and navigates to specified websites without the user's knowledge.
- **Bot networks:** server-based browsers that masquerade as human users — in recent years, this tactic for generating non-human traffic has become more widely-used than the browser hijacking malware described above.
- **Ad-stuffing:** a tactic in which the publisher fills a web page with invisible ads the user can't see.

While it's impossible to know exactly how much money the bad guys have made off with, the Interactive Advertising Bureau estimates that ad fraud cost our industry $8.2 billion in 2015. Meanwhile, the advertising verification firm Adloox predicts that this number could be as high as $16.4 billion in 2017. These numbers are inherently imprecise, as there's no one indicator that guarantees an impression is fraudulent — but these estimates make it clear that this is a huge problem for our industry.

Beyond the sheer magnitude of this theft, fraud presents an existential threat to the digital ecosystem because it undermines brands' trust in the programmatic marketplace. According to a recent report from the Chief Marketing Officer Council and Dow Jones, 72% of programmatic advertisers are concerned about brand safety and control in the programmatic marketplace.

How a complex supply chain gives cover to fraudsters

With so much riding on the line, you may be wondering why no one has developed a solution to eliminate ad fraud and end the cat-and-mouse game once and for all. The answer lies in the complexity of the programmatic supply chain, a multi-faceted ecosystem that Procter & Gamble CMO Marc Pritchard famously described as "murky at best and fraudulent at worst."

In every programmatic transaction, brands are separated from the end user by a multiplicity of agencies, technology vendors, and ad networks — all of whom provide vital information about the impression the brand is purchasing. With so many hops along the chain, it can be extremely difficult for anyone involved in the transaction to ensure that the rest of their business partners are acting ethically. All it takes is one deceitful actor to perpetuate a multi-million dollar heist.

Just recently, Buzzfeed uncovered one such scheme, in which a digital media agency used device-hijacking software to drive millions of fake visitors to a

network of low-quality websites. Along the way, the agency managed to steal from the ad budgets of Disney, Gilette, and over 100 other brands.

What AppNexus is doing to clean up the marketplace

Despite the obstacles in our path, AppNexus is deeply committed to rooting out ad fraud and building a more trustworthy digital marketplace.

Over the past three years, we've invested a substantial amount of time, money, and resources to stay one step ahead of the bad actors who threaten our clients. Today, we have 30 fraud detectors running at all times. By collecting all of the data associated with each impression – from the time the user opens a page to when the ad is served – we're able to see when something is amiss. Plus, we go the extra mile to better understand how fraud works by analyzing the malware these bad actors use in a safe, "sandboxed" research environment. This allows us to see how the malware behaves in its attempts to mimic human behavior and to track which sites it visits, which lets us ascertain that those sites are most likely buying non-human traffic.

This is just the beginning. As the world's leading independent ad tech platform, we see hundreds of billions of impressions every day, each of which is logged and analyzed by our anti-fraud detectors. In order to stamp out the next generation of bad actors, we decided to use our unique vantage point to study where fraud is happening today and which new tactics are being used to steal advertiser budgets.

Establishing these instances and assigning blame isn't easy, since there's no single, definitive indicator that an impression is fraudulent. But with the latest machine learning techniques, our data scientists can find trends and patterns that point to a high likelihood of fraudulent activity. Some of those techniques include:

- **Cluster analysis:** a foundational data analysis technique that involves grouping data points together based on key similarities. Cluster analyses helps us find suspicious commonalities between impressions generated by shady traffic sources.

- **Covisitation:** a method of identifying overlaps in traffic between different web sites, which helps us establish links between sites getting questionable traffic from the same sources. This study from researchers at NYU and Dstillery (formerly Media6Degrees) offers a more in-depth explanation of how covisitation can uncover instances of ad fraud.

- **Honeypots:** a method by which we draw bots into the open by creating a fake site and then sending traffic to it from suspicious vendors. We can then analyze for signs of non-human visitors.

Using these methods, our data science team has found that fraudsters are using

increasingly-sophisticated tactics to fool fraud detection mechanisms. Perhaps most interestingly, they've also uncovered a meaningful overlap between viral content — including the subcategories of fake news and hate speech — and fraudulent activity.

Why viral content sites are at a unique risk for fraud

While virtually all major media companies at least dabble in stories designed for social media, the most likely publishers to be engaged in fraud are those that have chosen to build their entire business off viral content. In addition to their focus on virality, these publishers are distinct in that they are more likely to be motivated by money — or at least more brazen about their pursuit of it — than media firms are.

Essentially, viral publishers use the psychological triggers of shareable content to quickly generate audiences that would take a more traditional publisher years to build. By manipulating users' love for cute animals, their political differences, and their attachment to extreme ideologies, some of these firms have been successful in widely distributing their content on social media. Even better (for the fraudsters), they've been able to generate ad revenues off of low-quality posts that require far fewer resources than a properly-reported news story or a professionally-made video.

However, even successful viral publishers often find themselves in need of a traffic boost. Since users rarely bookmark the viral sites they find on social media, these publishers are more likely to lack any sort of regular, organic traffic. As a result, they have to rebuild their audiences each and every day, especially when the latest batch of outrageous listicles fails to deliver.

What our data scientists are seeing is that viral content publishers frequently acquire traffic from third-party sources of questionable repute. In some instances, they knowingly create or purchase fraudulent traffic with the express intent of ripping off advertisers. In others, they purchase massive audiences from vendors at below-market rates, without investigating how their partners deliver so much traffic for so little money.

At the end of the day, viral content publishers are in the business of arbitrage. They acquire visitors on a tight budget and sell them to advertisers for as much money as they can get. As such, their success hinges in large part on their ability to pay as little money for traffic as they can possibly get away with. It doesn't take a genius to figure out what kind of traffic is cheapest.

TRAFFIC ACQUISITION 101

Of course, not all traffic acquisition deals are fraudulent; indeed, a great many of them are entirely legitimate. Certainly, there are thousands of direct-response brands that buy cost-per-click ads to effectively drive real human users to their product pages. But generally speaking, the less you pay for traffic, the riskier a vendor you'll be working with and the less likely it is that you're getting actual humans to your site.

Where things get tricky is in the murky space between a completely legitimate cost-per-click ad deal and an obviously fraudulent bot network.
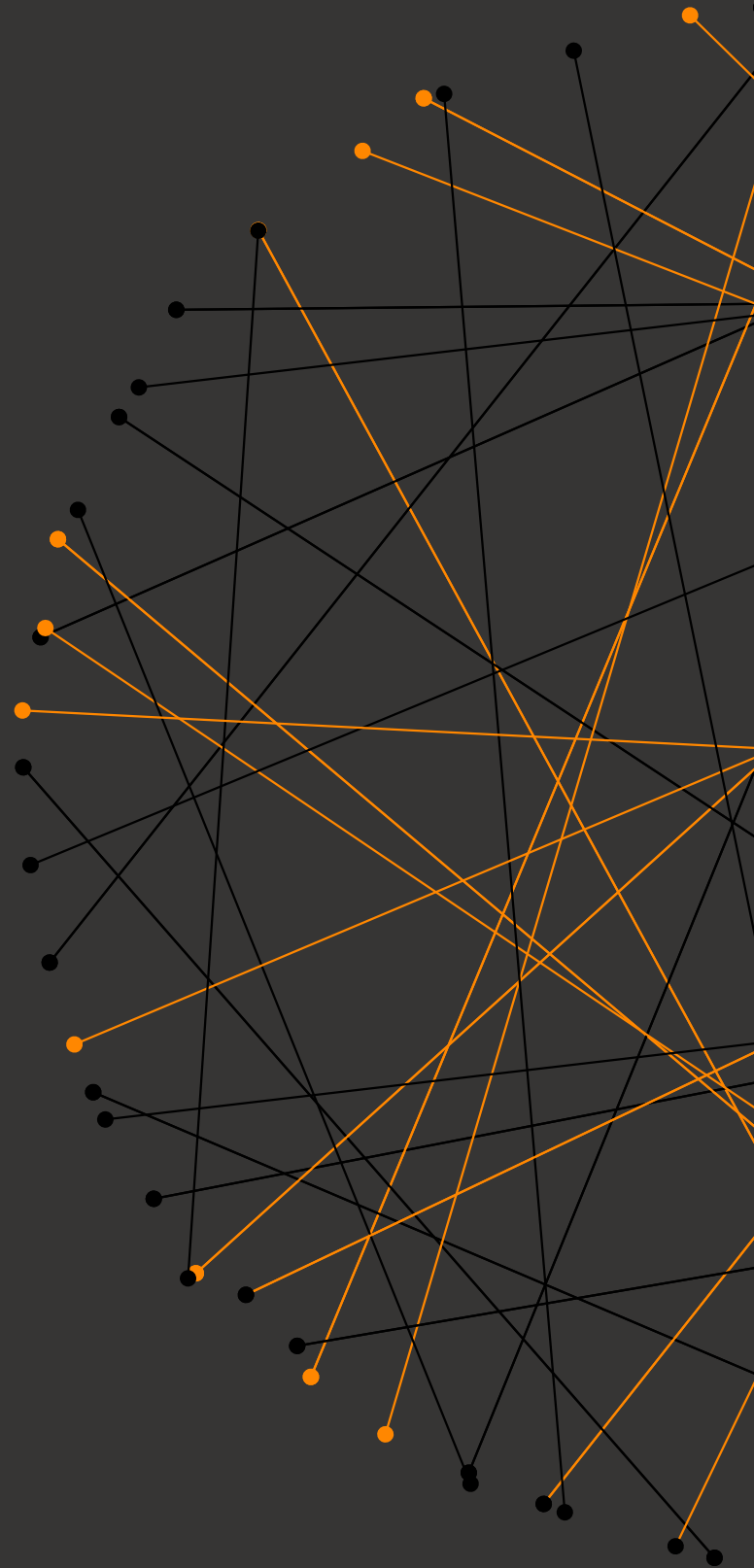
Here's a look at what we think of as the three major categories of traffic acquisition strategies, broken down by the level of perceived fraud risk they present.

Generally Considered Safe

While it would be impossible for us to guarantee that any online ad buy will be 100% fraud-free, there are certain kinds of traffic acquisition that brands and publishers can trust to deliver real human users. These include:

- **Organic marketing initiatives:** The safest way to acquire traffic is the old-fashioned way: producing great content and growing an audience around it. Sources in this category would include:

  › **Direct traffic:** traffic that comes from a human user simply entering a site's URL into their browser.

  › **Organic search traffic:** traffic that comes from human users who find a site in the non-sponsored results from a search engine.

  › **Email marketing**: traffic that is received by sending an email to an address a human user has provided to the publisher.

  None of these tactics are guaranteed to attract exclusively clean traffic, as it's still possible for sophisticated fraudsters to create bots that can mimic the human user behaviors associated with each. But the key differentiator is that in each case, the publisher isn't paying for the traffic, which means there's no incentive for anyone (except the publisher itself) to fake it.

- **Search engine pay-per-click:** In a search engine pay-per-click deal, advertisers purchase the right to promote a link in a user's search results after they enter a query. The major search engines — namely, Google and Bing — are run by large, reputable companies with dedicated fraud departments, so buyers who work with them can be confident that they're getting quality traffic. However, the same can't be said for many of the smaller players in the space.

  For the most part, firms that run search engine pay-per-click campaigns are in the retail business, as opposed to the publishing business. When they buy access to certain users and keywords, they're aiming to sell products and services, once the link is clicked. Moreover, since there are many major retailers competing for users on these popular search engines, the cost of certain keywords is quite high. Those costs mean it's virtually never profitable to use major high-quality search engines for ad aribitrage.

  Additionally, these sites can confirm the user is a real human when they purchase, since they provide some level of contact information required to finalize the exchange of goods and/or services. While it may be possible to fake an email or mailing address, making a purchase would require building a bot sophisticated enough to navigate a checkout procedure, which is more time and

resources than most fraudsters are willing to invest. Plus, you'd then have to spend money on an item, which could cancel out any ad revenue gained on the impression.

Completely Fraudulent

There are some traffic acquisition schemes that publishers should avoid entirely and for obvious reasons.

- **Cloudbots and traffic exchanges:** In this type of traffic acquisition deal, a bad actor rents a virtual private server from a cloud hosting service like Amazon Web Services, Server Beach, or Digital Ocean. Then, they run a headless, automated browser and visit several websites, either directly or through a traffic exchange. The virtual private server can be created by the publisher themselves, or the publisher can pay a third party to send them bot traffic.

  In a traffic exchange, a group of people agree to visit each other's websites numerous times, in hopes of artificially inflating their traffic. This activity is fraudulent even in cases where publishers manually navigate to sites owned by their co-conspirators.

  When AppNexus initiated a major initiative to crack down on bad actors in 2015, most of the impressions we terminated from our platform were coming from cloudbot and traffic exchange schemes. However, fraudsters have developed new tactics to evade detection. Last year, perpetrators of the Methbot fraud scheme purportedly stole millions of dollars by disguising data centers as individual home computer users. (Because of our ongoing investments in inventory quality and fraud detection, AppNexus was impacted very little by this attack.)

- **Pay-per-click exchanges and botnets:** A botnet is a network of computers that have been infected by "malware" that allows hackers to exploit their computing power for nefarious purposes. In the case of ad fraud, bad actors use the captured devices to drive automated, non-human traffic to various websites.

  Frequently, the operator of the botnet will sell clicks generated by the bots they control on a pay-per-click exchange, allowing unscrupulous publishers to sell the impressions as if they came from real users.

  Like the cloudbots we described a moment ago, botnets are egregiously bad. Unless they're highly sophisticated like the Methbot scheme, they can be detected somewhat easily by modern anti-fraud techniques.

Proceed With Caution

The focus of this whitepaper is the grey area that lies between safe traffic acquisition and obviously-fraudulent schemes.

While these traffic acquisition methods are both legal and permissible in today's ad tech ecosystem, they raise red flags that deserve serious consideration. Publishers that engage in these tactics need to be careful if they hope to remain on the up-and-up.

- **Content discovery networks:** Firms in this category promote links to publisher content in native advertising widgets on sites across the web. By now, most people are familiar with these recommendation widgets, which typically sit at the bottom of the page.

  Though these companies work with numerous legitimate publishers, it can be difficult for clients to determine exactly how much human traffic they're really getting. Even scrupulous publishers can inadvertently purchase non-human users through these programs, so long as at least one other site in the network has acquired fraudulent traffic.

  When a small publisher only sees a portion of its content discovery data, it can be very difficult to detect when another site is providing a combination of human and non-human users.

- **Social media pay-per-click:** A social media pay-per-click buy occurs when a publisher pays a platform like Facebook or Twitter to place native content links in the middle of the user's feed.

  Due to the range in prices for social media ads and consistent changes in social networks' newsfeed algorithms, arbitrageurs need to adjust their strategy to take advantage of this channel. In cases where ads are expensive, a site may lose money serving an ad to an individual, unless they overload the page with ads. But in many cases, publishers will accept the loss on an individual click in the hopes that the user shares the article with their network and attracts organic social media traffic.
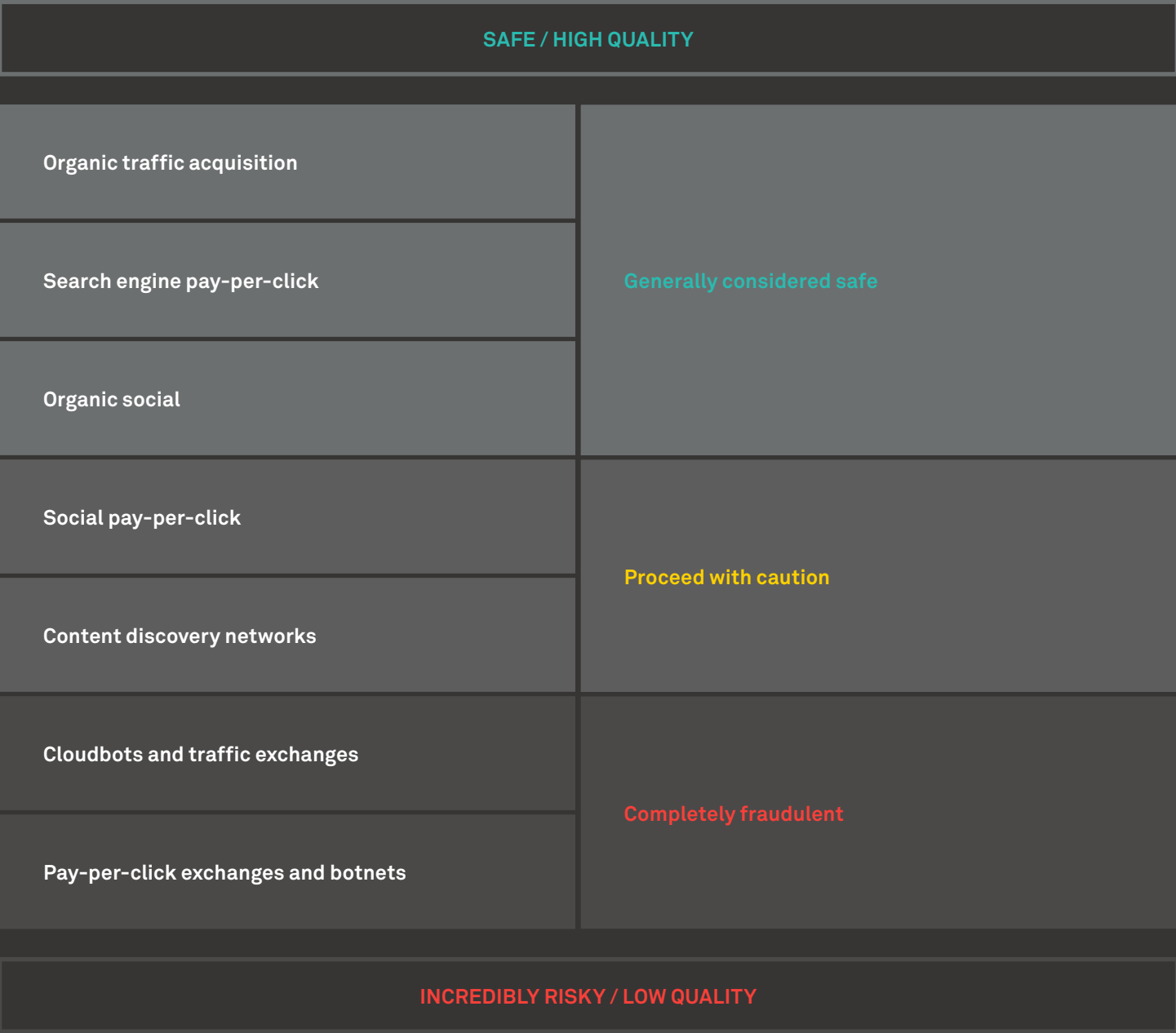
  That's one reason why social media sites are so often filled with outlandish, sometimes shocking content. The sites are unconcerned with building a long-term audience through quality, instead opting for quick and easy clicks. As such, they turn to overly dramatic "clickbait" headlines to increase the chances that users will visit and share their viral content.

  Ethics aside, publishers that are too reliant on social pay-per-click run the risk of putting their fate in the hands of the social network's algorithm. If a site like Facebook chooses to deprioritize their content, these firms could be forced to change their whole business model at the drop of a hat.
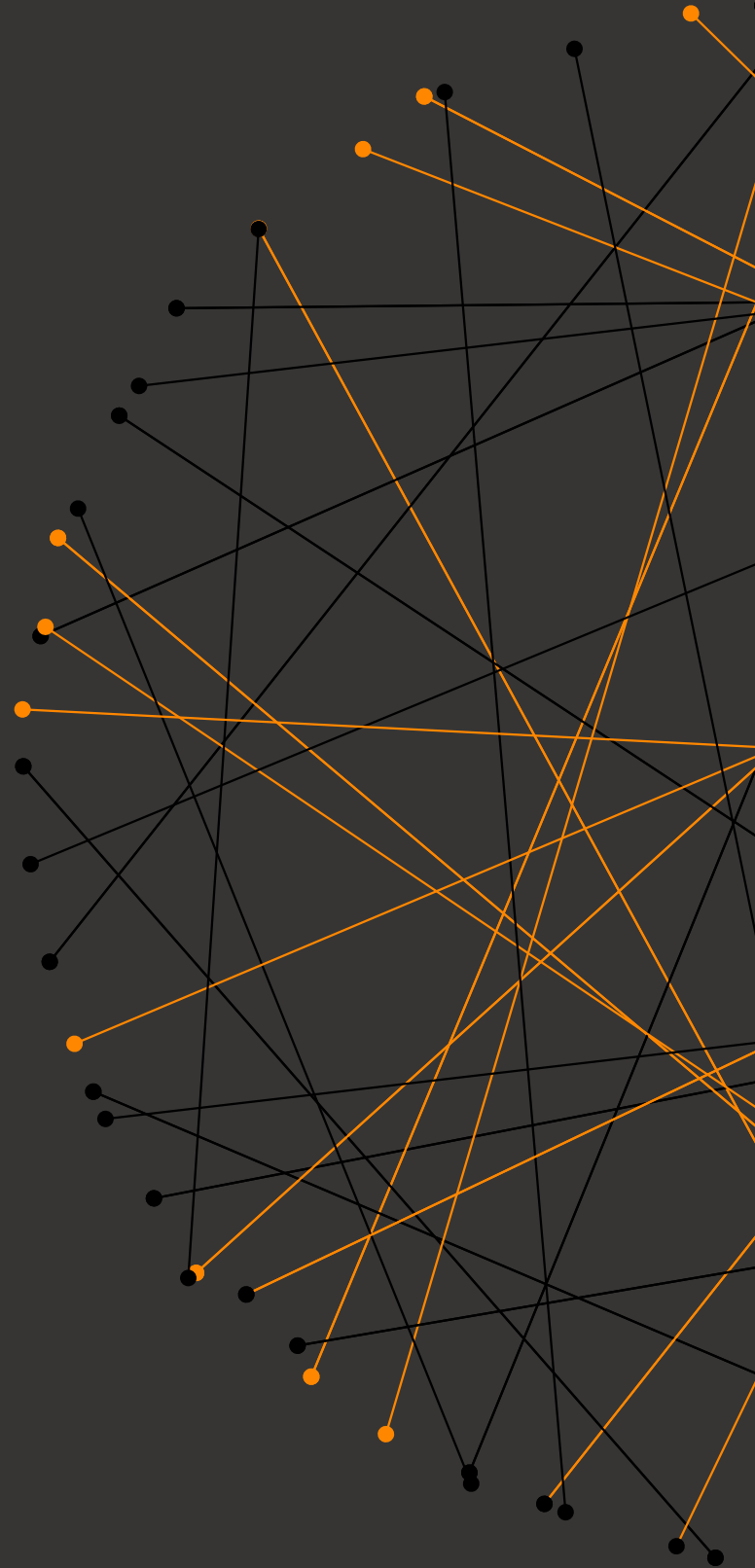
Traffic Source Quality Continuum

SAFE / HIGH QUALITY

Organic traffic acquisition	
Search engine pay-per-click	Generally considered safe
Organic social	
Social pay-per-click	
Content discovery networks	Proceed with caution
Cloudbots and traffic exchanges	
Pay-per-click exchanges and botnets	Completely fraudulent

INCREDIBLY RISKY / LOW QUALITY

WE TOOK A DEEP DIVE INTO VIRAL CONTENT'S FRAUD PROBLEM

Here's What We Learned …

Though we've described "viral content" on several occasions throughout this whitepaper, it may be useful to provide a more thorough definition of the term before we dive into the results of our analysis.

First and foremost, viral content is material that is created to be spread quickly online, particularly on social media. These stories take advantage of network effects to multiply their audiences every time a user shares them with their friends. Crucially, they encourage this sharing by using psychological triggers to whip the audience into a state of emotional arousal. As *Harvard Business Review* puts it, "Viral content tends to be surprising, emotionally complex, or extremely positive." However, the article also notes that "if arousal is high enough, viral content can be primarily negative."

With that definition, viral content includes a wide breadth of internet stories. It can be a surprising video of a cat and a dog riding a water slide together, or it can be a piece of hate speech or "fake news" that drives the reader into a state of excitement or outrage. The key is that all pieces of viral content manipulate the user's mood in order to generate clicks and shares.

In terms of content strategy, viral publishers are primarily focused on how many page views they receive, as opposed to how much time the user spends on the page once they get there. Rather than building a devoted audience that returns to the site day after day, these publishers bring people to their content as quickly and as cost-effectively as possible.

How Viral Publishers Attract Users

According to our research, bad actors in the viral content space typically use multiple traffic acquisition strategies, sometimes in combination.

On one end, they slap together low-quality content with just enough effort to get the stories through a quick human review. Upon closer inspection, visitors find that the content doesn't match the headline they saw on social media, or that the story contains plagiarism, or that parts of it were auto-generated by a computer program. Essentially, these publishers fool people into clicking on their stories by wrapping worthless content in the trappings of viral social media promotion.

Additionally, both legitimate and illegitimate viral publishers expand their audience reach by purchasing clicks through traffic acquisition programs. As noted earlier, these deals may bring in loads of traffic, but the publisher is left with little insight into who the users are, where they came from, or whether they're even human.

The Anatomy of a Suspicious Viral Publisher

Though viral content isn't necessarily an indicator of ad fraud, our data science team has come to see it as a warning sign that something might be amiss. Looking across the digital ecosystem, our experts have found that low-quality viral publishers are most at risk for non-human traffic.

In order to create a safe environment for brands, AppNexus applies vigorous scrutiny to the viral publishers on our platform, using automated and human detection methods to ensure our sites are acquiring their traffic organically.

We found that suspicious publishers typically use a handful of cheap tricks to pass themselves off as legitimate websites. Here are a few red flags to look out for.

- **"Borrowed" Content:** Overt plagiarism is a glaring sign that a publisher might have questionable ethics. For instance, on the site depicted below, every article is lifted from a mainstream publisher and filed under tabs for NPR, The New York Times, and Science Daily.
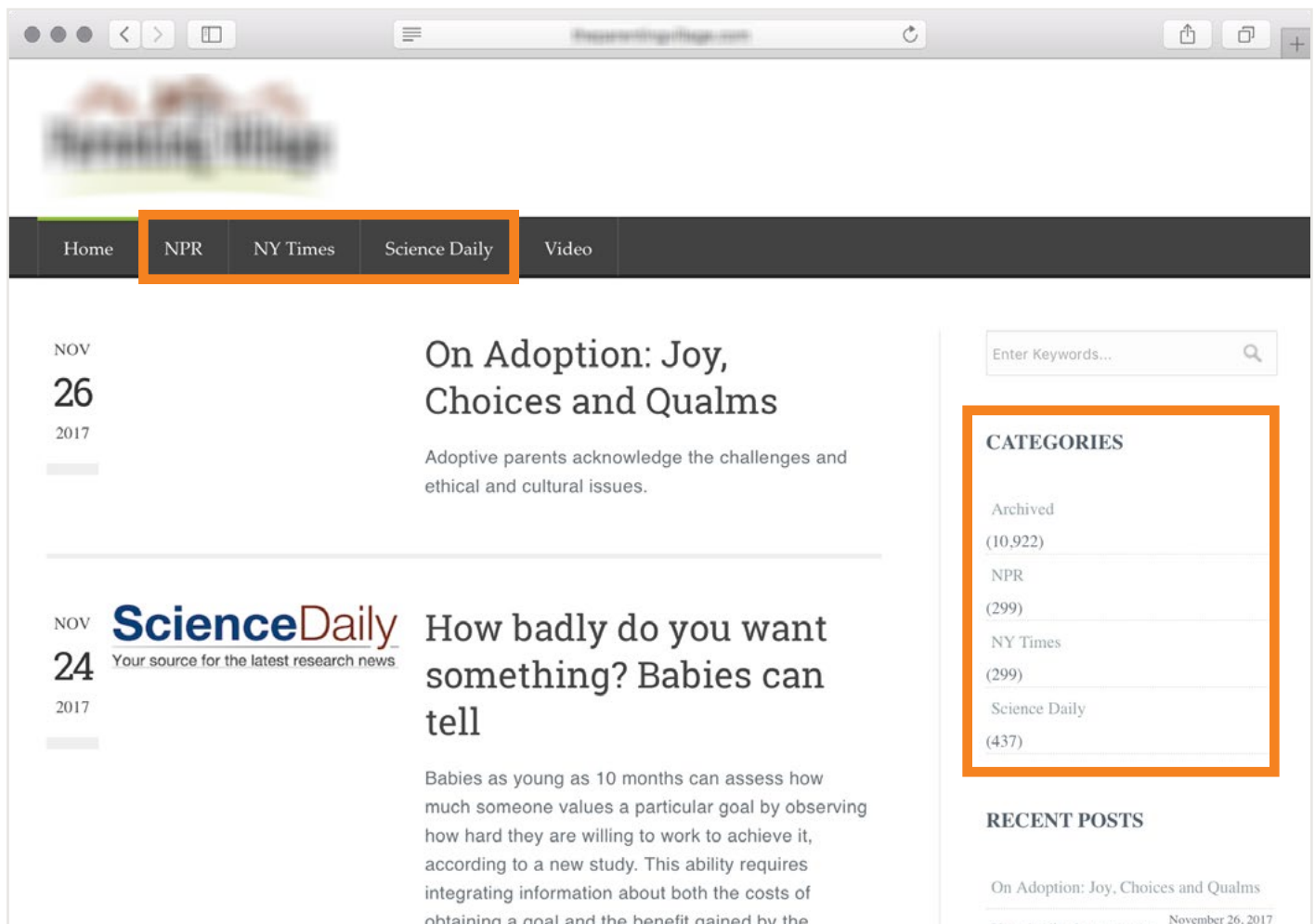
- **Awkward Formatting:** A tell-tale sign of lifted content is that parts of the story look out of place on the offending publisher's website. This happens when fraudulent viral publishers scrape a story off another site without reformatting it to fit their own design. For instance, the story below was lifted from NPR. It also has photo captions shoehorned into the regular article text.
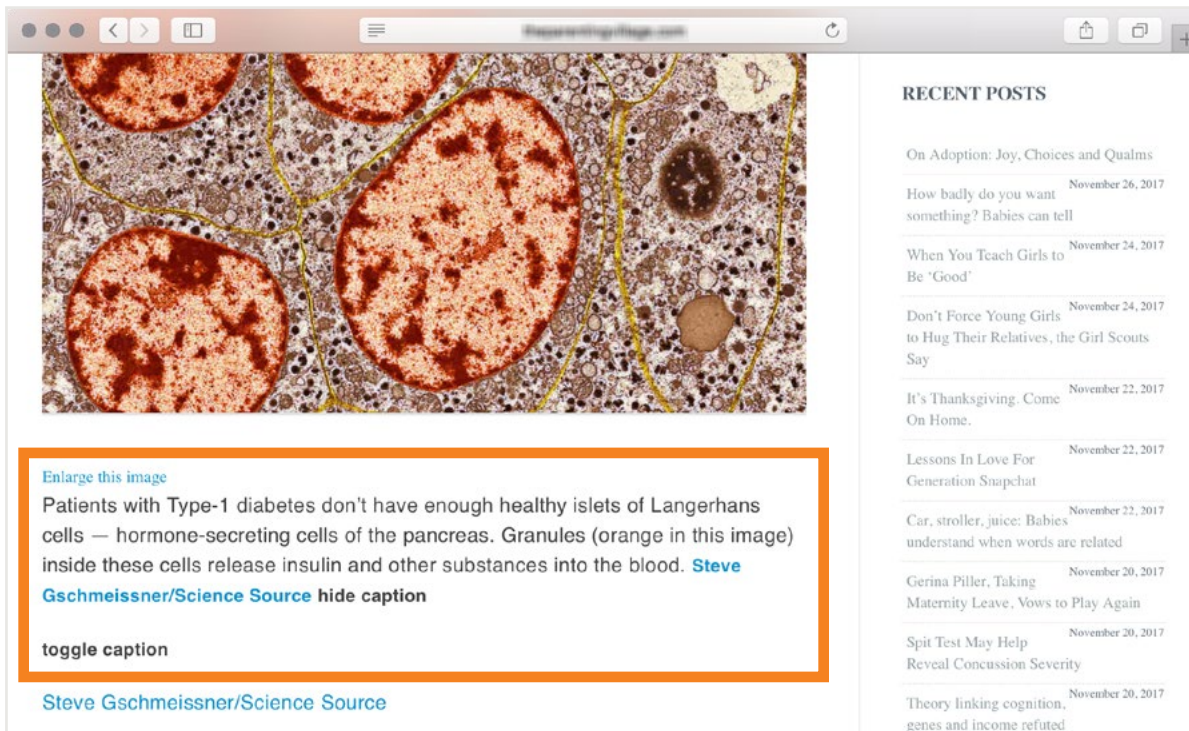

- **Partially-Scraped Content:** In some instances, fraudulent publishers don't even bother to lift an entire article. The article shown below contains just a single sentence after the headline.
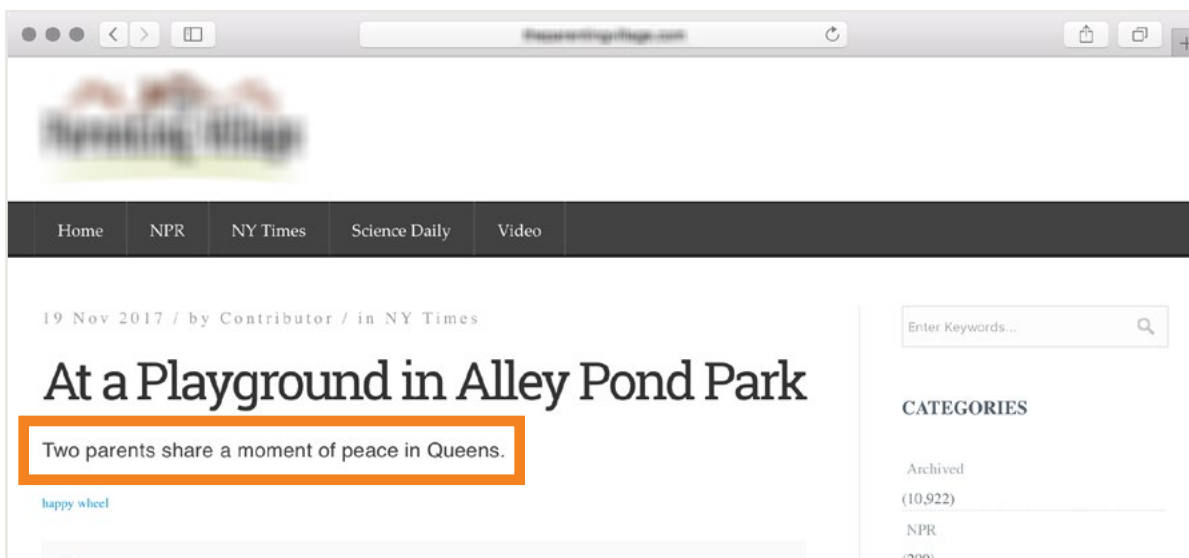
- **Out-Of-Place Links:** Links to random, unrelated pages are also common on suspicious, viral content sites. For instance, the page below has a link at the bottom bearing the text "happy wheel." When you click it, you wind up on a landing page that contains boilerplate writing about a free online game called Happy Wheels.
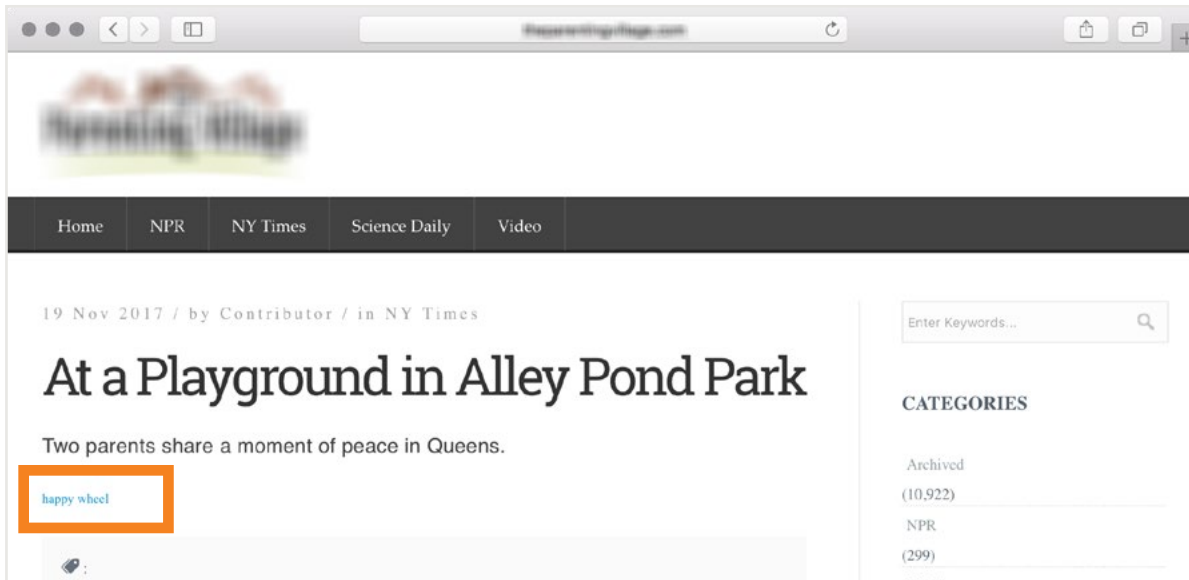

- **No Author Pages:** Real websites have real writers, and most of them contain author pages where readers can learn more about the site's contributors. Stories on sites like the site depicted below, which plagiarizes content from ESPN, are published without bylines.
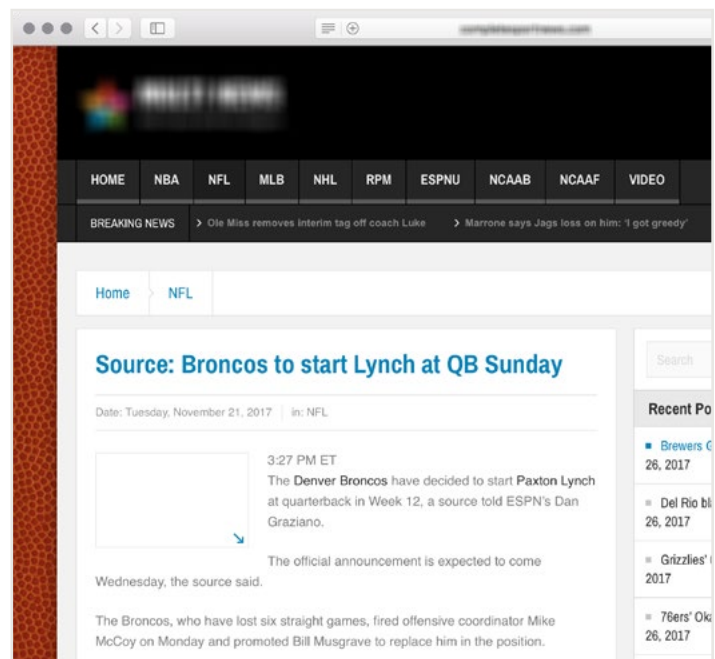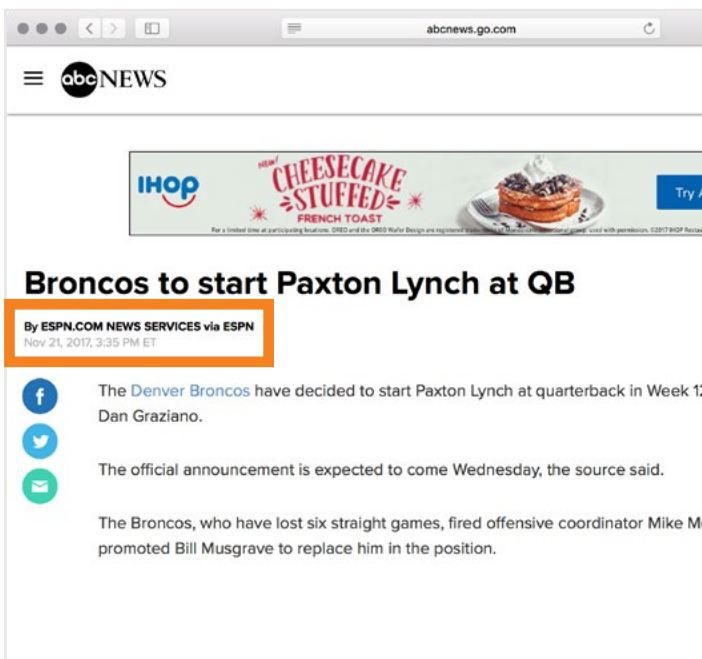
- **Heavy Ad Loads:** Quality publishers balance their desire to generate ad revenues against their concerns that an ad-saturated user experience will drive away their audience. But if you're not expecting human beings to consume your content, you don't have to worry about degrading the user experience with an inundation of advertising. As such, these sites will often look like the one depicted below, which is overflowing with autoplay videos and other intrusive ad units.

- **Missing Sections:** Low-quality viral sites frequently come off as if they are still in the process of being built. For example, take a look at the site below. It's built to look like a conventional online newspaper, but it only posts lifted stories and low-quality viral content. When you click a link to its sports section, nothing comes up.
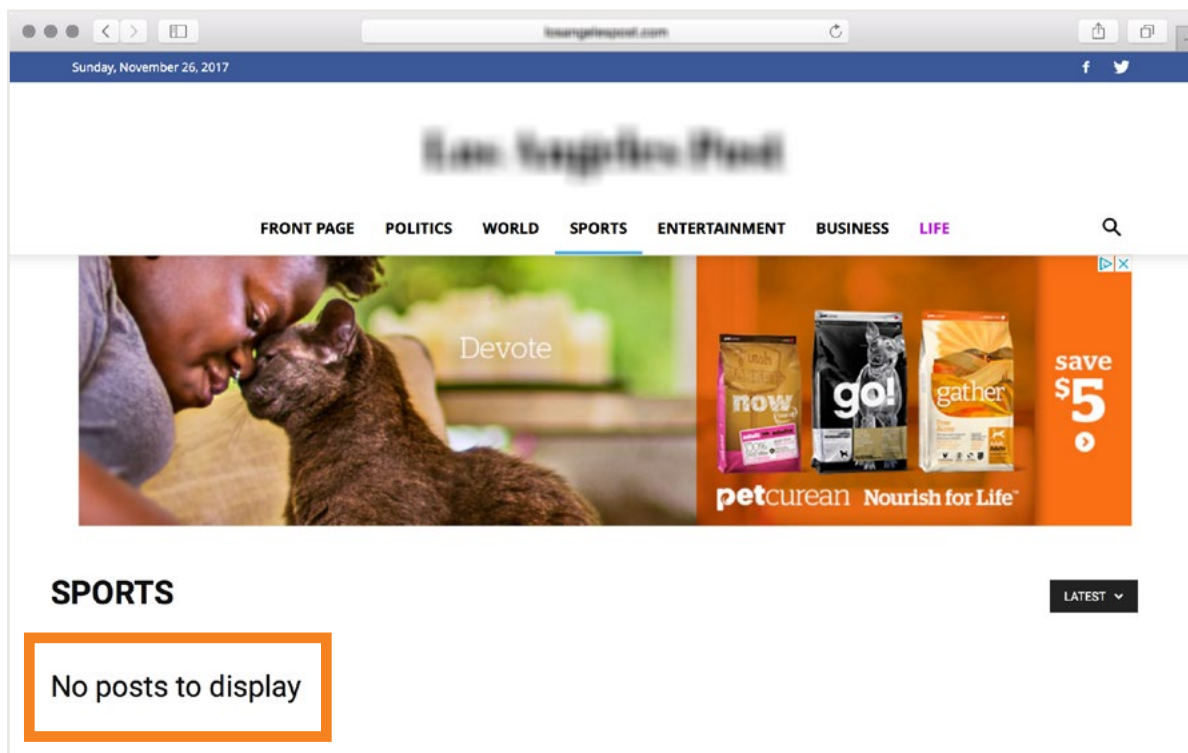
- **Reliance on Embedded Video:** Embedded videos are a nice way for publishers to present viral content without having to do much work. Because legitimate publishers serve their own ads beside their videos, they sometimes do not place restrictions on other sites who want to republish their content. As a result, sites like the fake newspaper depicted below are full of viral videos packaged alongside shoddy, slapdash writing.
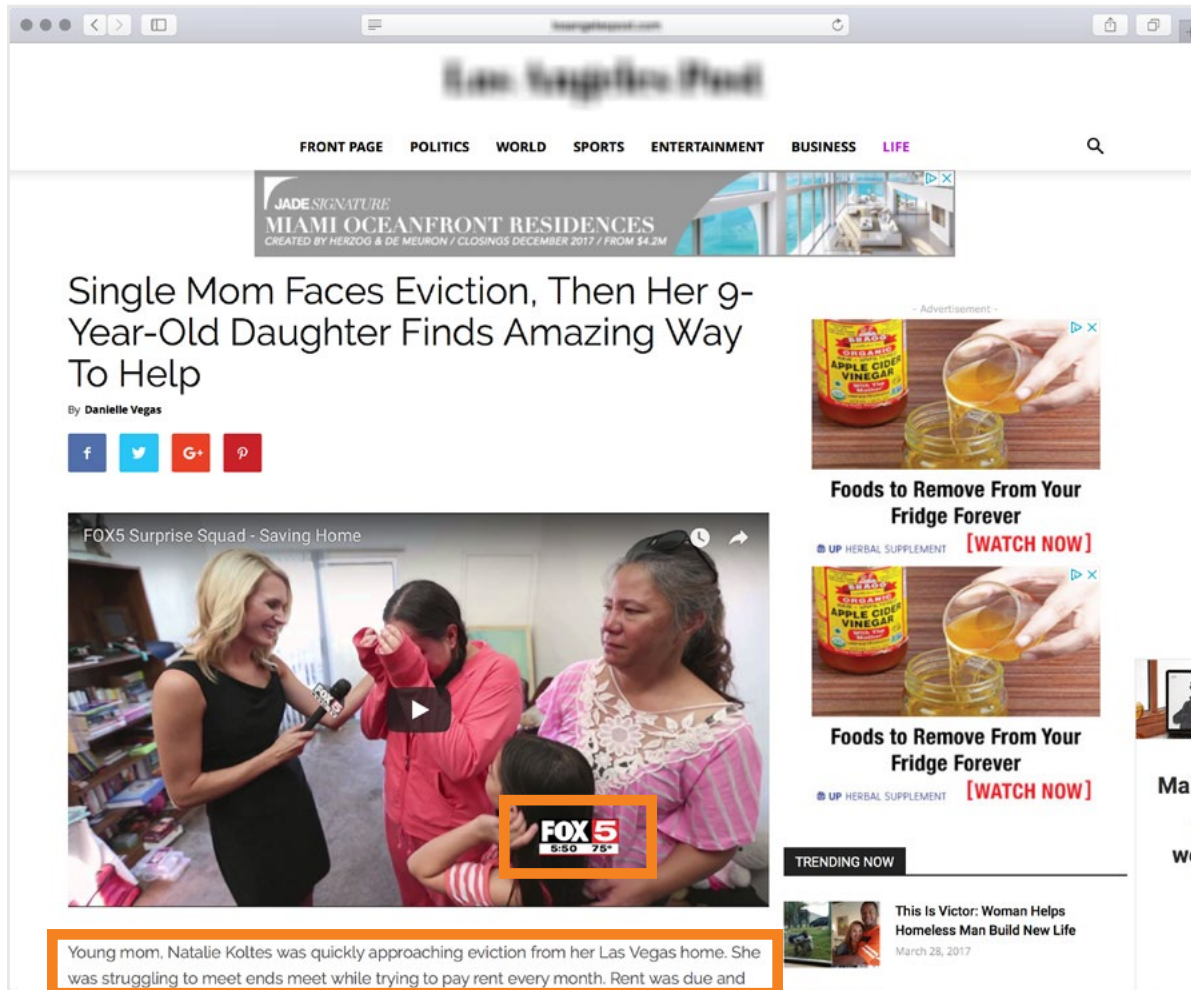
- **Tab-unders:** In addition to monetizing their low-quality content with display and video ads, fraudulent viral publishers also sell traffic to other sites. One way they do this is by using tab-unders, a tactic that opens a new browser tab beneath the one the user is looking at — without the user's consent or knowledge. **Pop-unders** are another variation of this tactic, in which an entire new browser window opens underneath the window the user is actually looking at.

How Fake News and Hate Speech Play Into the Viral Equation

In addition to the overlap between viral content and ad fraud, our data scientists found a significant connection between viral publishers and the "fake news" content that dominated the conversation in the months prior to and following the 2016 U.S. presidential election. In fact, our data scientists ultimately came to the conclusion that fake news is a subcategory of viral content rather than its own breed of storytelling.

Like other viral content, these ideologically-slanted, factually-dubious stories exist to get readers fired up, engaged, and clicking. Some of these publishers don't care why a reader would click; all they care about is that they did click. Another commonality our data scientists found is that many ad-supported fake news publishers appear more motivated by advertising money than by ideology — what they care about most are clicks. In some instances, we found publishers that appear to own and operate both liberal and conservative fake news sites. By contrast, online destinations that are considered to be bonafide, hardcore hate sites usually sustain themselves on donations rather than advertising.

One of the most fascinating things we found was that fake news and hate speech content usually operate in a grey zone between truth and fantasy, with stories rarely created entirely out of whole cloth. Instead, a kernel of reality is embellished and wrapped inside hyper-partisan commentary designed to shock and excite the audience. In other instances, conspiracy theories and factual inaccuracies were included in stories, but they were presented inside articles that also discussed real events.
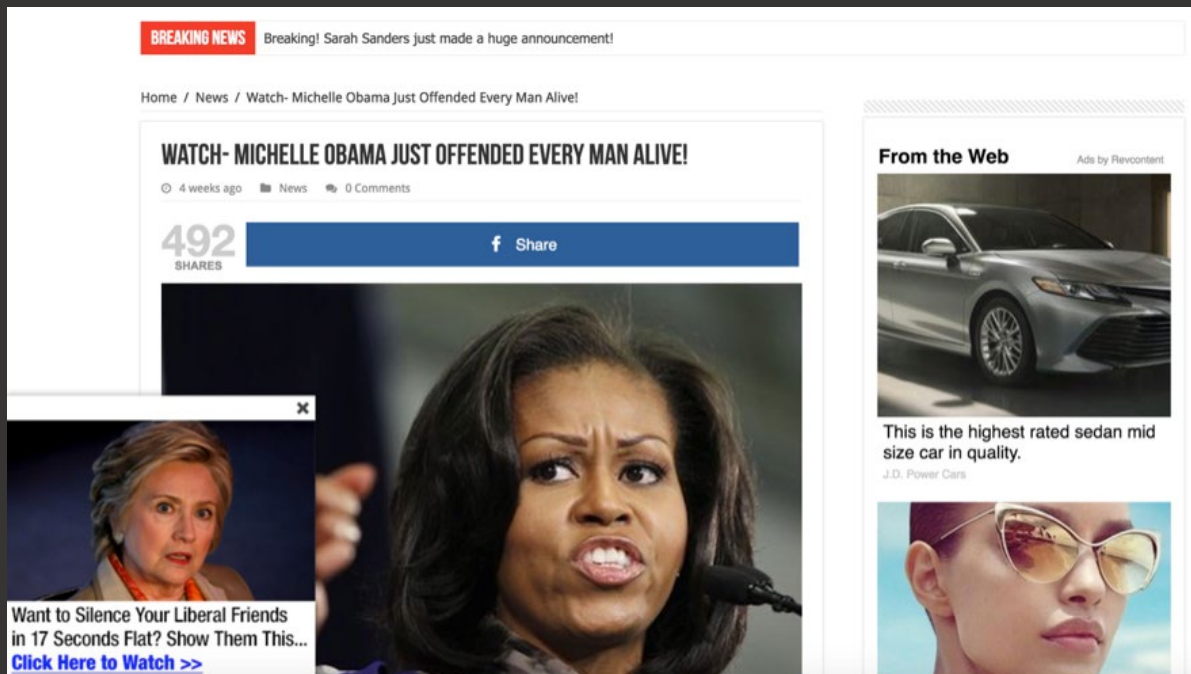
For instance, the post depicted above discusses a real video of Michelle Obama speaking about gender issues, beside commentary declaring that "MICHELLE OBAMA JUST OFFENDED EVERY MAN ALIVE!" In addition to the incendiary spin on her remarks, the site slipped in the patently-false claim that her children were adopted. And, of course, it's filled with ads, some of which are political in nature.

Other posts urge readers to share them with their friends, claiming that doing so is an important act of resistance against social platforms that are "suppressing stories from conservative sites who wholeheartedly support President Trump." However, according to our data science team, we've seen cases where one specific ad network will hawk impressions for both right- and left-leaning sites.

All of this puts ad tech platforms like AppNexus in a bit of a bind. While we're committed to being a non-partisan, open ad tech platform, we also want to make sure we're not allowing bad actors to sell inventory with us. Furthermore, we enforce strict platform rules barring the monetization of sites or apps that promote hate speech, deceptive acts in commerce, and graphic violence.

In order to take the bias out of our analysis, we use consensus methods to measure fraudulent traffic, identifying ad fraud based on traffic patterns and other non-ideological factors. If we find low viewability, irregular content, and unusual traffic patterns, we believe there may likely be something suspicious happening. Any one of those elements on their own? Not necessarily.

THE WAY
FORWARD

Combatting Ad Fraud in the Age of Viral Content

An important thing to remember is that viral content is just the latest front in an ongoing battle between fraudsters and anti-fraud actors.
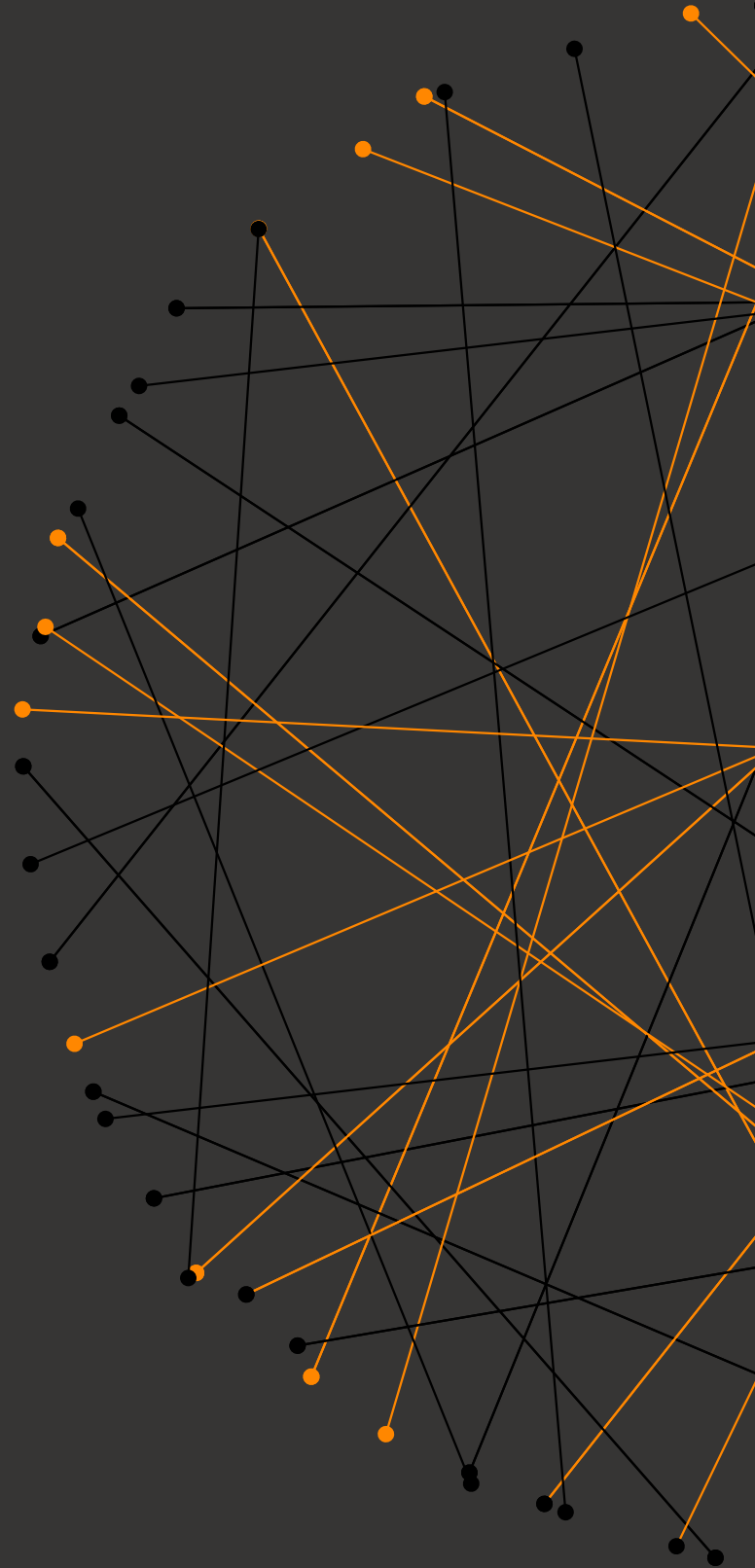
An important thing to remember is that viral content is just the latest front in an ongoing battle between fraudsters and anti-fraud actors.

When our industry became better at sniffing out data center bot fraud, bad actors in the Methbot scheme spoofed the IP addresses of their data centers to impersonate subscribers to major home internet providers. Eventually, ad tech will discover effective tactics for fighting viral content fraud, and bad actors will be pushed out to other parts of the internet economy. Already, industry-wide efforts like IAB's Trustworthy Accountability Group and ads.txt have shown our resolve to work together to eliminate shady publishers from the supply chain.

But for now, advertisers and publishers need to begin developing a roadmap for evaluating viral traffic. Beyond identifying obviously fraudulent actions, industry stakeholders must figure out how they're going to handle the grey areas of viral content fraud. What happens when viral content attracts a mix of human and non-human traffic? What happens when an audience composed entirely of real humans flocks to low-quality "fake news" content? Answering these questions will prove a valuable first step in our efforts to solve this problem.

At AppNexus, we're chipping away at the problem from multiple angles, using multi-pronged strategies – some even ranging from pest management and epidemiology – to cut down on fraud wherever we can. In addition to listening closely to customer feedback, we're investing in more rigorous content analysis to determine the quality of a given website. This includes scraping page content to check for plagiarism and assessing whether the body of a post matches its URL.

In the meantime, we'll leave you with a number of useful steps that internet users, publishers, and advertisers can take right now to fight fraud in viral content.

Internet users

- **Educate yourself in media literacy.** Learn to distinguish quality content from questionable content.

  Before you share an article, try clicking around the website to see what other stories they've published. Are there other sources that confirm their reporting? Does it look like they've put time and effort into getting things right? Do you recognize any of the cheap tricks we listed earlier as you explore the site?

- **If you see something, say something.** Don't hesitate to complain to publishers when you encounter ads, sponsored content widgets, or websites you don't like. Legitimate media companies should be very sensitive to the user experience.

Publishers

- **Use extreme caution when buying traffic.** You need to vet any traffic vendor you work with and keep in mind that the cheaper the clicks are, the more likely they are to be fraudulent — if you're getting a deal that sounds too good to be true, it probably is. If you're not careful, you could wind up damaging your reputation in the marketplace. One good question to ask yourself is whether or not this partner has a dedicated fraud team, which would be an indication that they take this issue seriously and don't traffic in fake users. A lower bar to clear would be to see if the company's web site resembles that of a reputable vendor. Many traffic acquisition companies dealing in fraudulent traffic have web sites containing only generic text and no information about employees, which would indicate their business may not be entirely above board.

- **Build an engaged, devoted audience.** It's more important to build a steady, stable stream of visitors than it is to go for clickbait and quick eyeballs. Ultimately, a viral-only approach is unsustainable, as it leaves you forever at the mercy of social algorithms.

- **If you traffic in fraudulent impressions, know that your days are numbered.** More and more ad tech players are working on ways to privilege quality content. Time is running out for plagiarizers and viral manipulators.

Advertisers

- **Dig into your data.** The best weapon against ad fraud is to audit your supply chain with every tool at your disposal. When JPMorgan took the time to manually audit the sites its ads were running on, the company eliminated all but 5,000 of the 400,000 publishers it was working with.

  Notably, removing these sites had no impact on performance, meaning that the company had previously been buying ads from 395,000 publishers that provided zero value.

- **Ask publishers where their traffic comes from.** If a publisher buys traffic from third-party sources, avoid purchasing their impressions.

- **Talk to your vendors and programmatic partners.** At AppNexus, we read every single case that's submitted to us by our clients. We're always looking to learn more about ad fraud, and we're always happy to help whenever there's a problem. If there's something on your mind, we'd absolutely love to hear from you.

DISCLAIMER