EPISODE 332

[INTRODUCTION]

[0:00:00.7] JM: The online advertising industry is a giant casino. Giant technology companies are the casino owners. Online publishers are the Casino employees. The brand advertisers are the victims who keep returning to the casino to lose their money. The small ad tech companies are the sharks who make lots of money exploiting the inefficiencies of the system.

One of these smaller ad tech companies is called eZanga. eZanga sells pre-filtered traffic. Pre-filtered traffic means traffic that will pass through bot detection filters. A publisher can purchase traffic to their website so that the ads on that website get viewed. eZanga describes this technology as marketing and has won a giant contract with the United States Government to handle the advertising for the GSA.

Advertising fraud does not just promote misinformation. It is now taking our tax dollars and spending it on paid traffic. If any of these is confusing to you, don't worry, we explain it all in today's episode with Shailin Dhar who is the advertising fraud expert who wrote a detailed report about eZanga and its contract with the U.S. Government.

Shailin was previously on the show to give an overview of ad fraud and what his work as an ad fraud investigator entails. He works at The Dhar Method, a company he started to do consulting in advertising fraud, and it's a great business because almost nobody talks about this stuff, and the brands really want to hear about it.

Also, Shailin will be a speaker at our third meet up, which is Wednesday, May 3rd at Galvanize in San Francisco. The theme of this meet up is Fraud and Risk In Software, and you're going to hear about some fraud today in this episode and you're going to hear about fraud from Shailin when he talks at the meet up. We're going to have some great food. We're going to have engaging speakers including Shailin and a speaker from Coinbase. We're going to have a friendly intellectual atmosphere and we hope to see you there.

To find out more, you can go to softwareengineeringdaily.com/meet up.

[SPONSOR MESSAGE]

[0:02:20.7] JM: Software engineers know that saving time means saving money. Save time on your accounting solution. Use FreshBooks Cloud Accounting Software. FreshBooks makes easy accounting software with a friendly UI that transforms how entrepreneurs and small business owners deal with a day-to-day paperwork. Get ready for the simplest way to be more productive and organized. Most importantly, get paid quickly.

FreshBooks is not only easy to use, it's also packed full of powerful features. From the visually appealing dashboard, you can see outstanding revenue, spending, reports, a notification center, and action items at a glance. Create and send invoices in less than 30 seconds. Set up online payments with just a couple of clicks. Your clients can pay by credit card straight from their invoice. If you send an invoice, you can see when the client has received and opened that invoice.

FreshBooks is also known for award-winning customer service and a real live person usually answers the phone in three rings or less. FreshBooks is offering a 30-day unrestricted free trial to Software Engineering Daily listeners. To claim it, just go to freshbooks.com/sed and enter Software Engineering Daily in the How Did You Hear About Us section. Again, that's freshbooks.com/sed.

Thanks to FreshBooks for being a sponsor of Software Engineering Daily.

[INTERVIEW]

[0:04:01.6] JM: Shailin Dhar is an advertising fraud consultant with The Dhar Method. Shailin, welcome back to Software Engineering Daily.

[0:04:07.7] SD: Thanks for having me back, Jeff.

[0:04:08.6] JM: Advertising fraud is a serious problem as we both know. It takes money from internet publishers and it fills our newsfeeds with false information. On Software Engineering

Daily, we've done many shows about ad fraud including a previous one with you as a guest, but some listeners don't know about advertising fraud. Give us an overview for what advertising fraud is.

[0:04:33.8] SD: Advertising fraud, the general scope entails all types of devious, nefarious practices where advertising dollars are spent on media that will never return any benefit to the advertisers. There's no ROI for the marketer that's spending the dollars on media. That can happen with robotic traffic visiting a site. That can happen with ad network falsely taking credit for an app install that happened organically, or it can happen with — Let's say an advertiser is targeting a specific demographic, but the ads are actually shown on sites that don't serve that demographic. All these different things come under the umbrella of what we call ad fraud.

[0:05:19.1] JM: That sounds like something that only affects marketers, it only affects big brands that are blasting their advertisements all over the internet. Explain what ad fraud actually affects all of us.

[0:05:31.2] SD: The reason that we get to consume so much content online for free is because a lot of it is advertiser supported, and if advertisers begin to have lack of faith in the integrity of online media, there's obviously the danger that they start pulling their dollars, which means that our content providers have to start charging us. I think that's something that definitely affects the wide consumer base.

The other thing is what affects these websites with ad fraud, let's say specifically with these fraudulent sites that run on robotic traffic, is that those are dollars that should be going to the legitimate sites that have real human visitors and they are making less money than they truly should. That creates a different dynamic of incentives that can lead to bad practices in the long run.

[0:06:24.9] JM: There's almost nobody in the technology industry who has any incentive to report the truth about advertising fraud. Explain why that is.

[0:06:37.2] SD: We're trying to change that because the reason that people don't feel the need or feel an incentive to report it is because they feel that they would be individually blamed for exposing or letting the market know that they was in their network or their platform, but I always push people to kind of look at it as, “This is a systemic problem and you can actually look better than your competition if you come out ahead of it and acknowledge that the problem is there and then create a process to find a solution rather than kind of hiding in the back and saying, “Oh, no. This is not an issue,” and kind of sticking our heads in the sand.”

[0:07:15.7] JM: There are some bot detection companies who sell products that claim to be able to detect and stop bot traffic. Do these companies have an incentive to tell the truth?

[0:07:28.1] SD: It depends on which company you’re talking about. There’re a lot of companies in that space that do operate with a lot of integrity. They have the best intentions both on their sales side and their engineering teams. Both sides are working hard. It really just becomes a case by case basis of what is the true operation in the background of the company.

[0:07:51.8] JM: The mainstream press has done some reporting on advertising fraud. There was the story from the New York Times about methbot, and this story described ad fraud as a scheme that is committed by Russian hackers and cybercriminals. As you and I know, that’s an incomplete version of the truth. In reality, ad fraud actually involves mainstream publishers, it is involves big name advertising networks, giant tech companies, these every day names, the tech companies that we know, other organizations. The Russian hackers are just a symptom of the institutional structure of online advertising.

I grew up playing poker. I understand the concept of a whale, that is the sucker who is gambling a lot and losing a lot. The brands are the whales. The giant brands that are blasting the internet with their advertisements, they are the whales. Online advertising is a giant casino built to exploit brands with a lot of money to spend, and Russian hackers are just hustlers that are trying to make a buck. They did not invent the games that everyone is playing. The technology companies invented the casinos. How accurate is that analogy?

[0:09:15.2] SD: I would agree with that.

[0:09:16.2] JM: Okay. You’re an ad fraud consultant. Go ahead.

[0:09:19.8] SD: This need for extra traffic without having scrutiny or oversight practices in place to check whether it’s human or not started way before we had all these — The invent of bot detection and all these discussions about fraud. As soon as pay-per-click started, you see the rise of affiliate marketers who deliver clicks and there are four steps of different traffic vendors away from the actual advertiser that are delivering these “results”. Everything was based on clicks, and that was such an easy metric to fake that the existence of these fraudulent traffic kind of got built into the foundation of advertisers. I feel I’m very, very confident that we have a way overinflated idea of the true supply of advertising online.

[0:10:18.8] JM: Meaning the true supply of people that are viewing ads, human traffic.

[0:10:23.2] SD: Yes.

[0:10:24.6] JM: You’re an ad fraud consultant. You make your money by going to brands and explaining to them that their ad dollars are spent on bots. I am a journalist, so I benefit from salacious stories. You and I are telling a story that is very much out of alignment with any story that a large tech company tells. For that matter, any story that, frankly, a journalist tells. I’ve really tried to read the journalism around this topic and it’s vapid, frankly.

Do the listeners have any reason to trust us over the claims that are made by the big tech companies and the writing from journalists who have surveyed this topic?

[0:11:04.1] SD: This whole trust thing goes back to; can the claims be supported? I don’t like to put out any claims that I don’t backup with evidence, and if there’s any evidence with previous reports that I’ve put out that the report doesn’t include all the details, I always welcome people to come reach out to me and I will give them the extra details that I just can’t have publicly available online.

I get that there is skepticism, definitely been called a fearmonger in the past, because people say that I try to make the problem look bigger than it is, but it’s not that I have any incentive to create fear, because I’m not selling you a perfect solution. I’m telling people that there’s also going to be fraud in advertising just like there’s always fraud and deception in any industry. It’s just about minimizing your risk. Anybody that says they have 100% fraud-free guarantee, I think is either misinformed, or is consciously making false claims.

The reason people can be skeptical on both sides, and I understand it. I think skepticism is healthy, but I think if somebody's making a claim, you should always look into it, and I've made it a point to always backup my claims with evidence.

[0:12:24.0] JM: It's important to note that neither of us is saying that advertising cannot work. Just like in a casino, there are some winners and sometimes the whales win. Sometimes you blast out a million dollars worth of advertising and $50 worth of it actually hits a human. In that sense — "Hey, the advertising actually is working."

[0:12:46.9] SD: This is the mentality and this is kind of the resistance to addressing fraud is, like I said before, we have this artificially inflated level of supply, a real human valuable supply. I think there's an artificially deflated return on investment based on that. So many advertisers had said, "Okay. This is — I'm used to this number that I'm spending on advertising and I'm used to getting this supposed return on it, and I'm okay with that as long as it stays that way."

What my issue with that is when you do an audit and you find out that you're actually spending X-amount on very suspicious advertising supply, or inventory. It's like saying — It's like a person saying that, "I ordered two stakes whenever I go out. Now, that I thought I was only wasting a little bit, and it was fine, 'cause I still end up full, and I'm showing that you could order half of what you're ordering and still end up full. Why would you continue to order the other half of that meal if it just ends up getting thrown away?"

[0:13:53.3] JM: Because my manager is paying for it.

[0:13:54.1] SD: Right. It's not your own money, it's your client's money, or your complex. This really — When it really hits down is when you see small business being affected by this, because global brands end up doing fixed marketing budgets every year. The corporation says, "Okay. We're going to allot $10 million to advertising this year," so the CMO takes that budget. He says, "Okay —" She says, "We're going to take 5 million and spend it on digital." Now, that 5 million needs to be spent. Otherwise, they get a lower budget the next year. They are incentivized as a marketing department to spend all of it, and then they take that to the agency who takes a margin of that 5 million and they are incentivized to make sure it all gets snapped.

There's not anybody in that chain, in that supply chain that's incentivized to make sure — At least currently, make sure that it's being spent safely and spent conservatively. It's not part of the structure right now. Year over year over year, people continue to spend money on advertising, because it does seem to work. It's just that, "Can we have a higher ROI? Can we make sure that publishers are — Real publishers, real human audiences, are being rewarded properly and just given higher rates based on the same budget and cut out the spend that goes to the fraudulent practices?"

[SPONSOR MESSAGE]

[0:15:28.7] JM: Deep learning is at the forefront of evolving computing and promises to dramatically improve how our world works. In order to get us to that bright future, we need new kinds of hardware and new interfaces between this AI hardware and the higher level software. That's why Intel acquired Nervana Systems, a platform for deep learning.

Intel Nervana is hiring engineers to help develop this full stack for AI, from chip design to software frameworks. Go to softwareengineeringdaily.com/intel to apply for a job at Intel Nervana. If you don't know much about Intel Nervana, you can check out the interviews that I've conducted with engineers from Intel Nervana, and those are available at softwareengineering.com/intel as well.

Come build the future of AI and deep learning at Intel Nervana. Go to softwareengineering.com/intel and apply to work at Intel Nervana. Thanks to Intel Nervana for being a sponsor of Software Engineering Daily, and I really enjoyed the interviews I have done with the Intel Nervana stuff. I think you'll enjoy them too.

[INTERVIEW CONTINUED]

[0:16:52.0] JM: The big tech companies that we could name off the top of our head, they're certainly complicity here, but they are really hard to scrutinize from the outside looking in. I wish there were an insider from one of these companies that would come forward and talk about this. I've tried to talk to these big companies, they basically will not discuss this at all. There is a giant ecosystem of small ad tech companies and these companies are a little more porous. You published a report recently about a company called eZanga, and eZanga has won a contract from the Federal Government of the United States for online advertising.

In your report, you described how eZanga "generates sources and sells fraudulent web visits and clicks. This is a huge concern. Tax dollars are not potentially going to be spent on fraudulent advertising activity."

We're going to unpack this whole scheme and let's start with eZanga. What is eZanga?

[0:18:01.8] SD: eZanga as far as I've known it, and until this year, has been an online advent work. They sell web visits on a per click basis. My first exposure to the company was in the XML search feed time. Basically, you get a search feed which means that — Let's stake a Yahoo search feed for example that's based on XML Code, I put that code on my page and now whenever a user visits my page and types a keyword into a search bar, it will return the Yahoo results of advertisers for that area, wherever that user is searching from.

I have shailinslocaldirector.com and I have visitors coming to the site, and whenever they — Let's take San Francisco for example, somebody in the outer Richmond of San Francisco comes to my site, types in local dentist, and they get the results that Yahoo had given them for local dentists if they search from the same IP address.

That was basically a great channel or people to send keyword-based bots that would go to a search page, type in a keyword and look at the results and click on a few of them and generate a pay-per-click that, let's say, Yahoo is paying $1.50. They're getting $2 from the advertiser. They're paying $1.50 to the basic ad network. Their next in line is getting a dollar and somebody at the end is getting 50 cents. That money changes hands, everybody takes a cut and nobody is complaining. That was my first exposure, is seeing them as a pay-per-click network for search ads. Then, got into the rise of pre-filtered traffic, and they were selling pre-filtered traffic for websites. You can buy traffic that's geared to pass the Integral Ad Science filter, or the Double

Verified filter, or Forensic, or Moat, or Pixelate, and they sold that at sub-penny prices; less than one cent per click. The lowest I saw for 4/10ths of a cent per click.

Sub-penny prices for, basically, unlimited amounts of traffic, and they include this in the report, screenshot from their platform, where it looks like it's a search ad. There's a title. There are subtexts, and there's a click URL, and that's all well and good, but my basic test just as a — Not like a technical test of the traffic, but my basic test was I'm just going to type in gibberish into the text ad and see if I get the same amount of clicks, and so I just bang on my keyboard and put it nonsense as a text search, and I still got the same amount of clicks. That's a very easy test to do, is, "Okay. I'm posting a supposed search ad for film reviews and I'm getting hundreds of thousands of clicks every day, but somebody could still, with the stretch of their imagination say, "Okay. That could be legitimate," so I just typed in nonsense and I was still getting the same level of clicks, and so it's very easy to tell that way that it's all robotically generated.

[0:21:05.5] JM: Lets' breakdown a little bit further what pre-filtered traffic is. You mentioned some things like Moat and some other filtering companies. Describe in more detail, what is pre-filtered traffic also that term traffic. Even that term may be unfamiliar to people. Why do you use that word traffic?

[0:21:26.6] SD: Web traffic is basically the term used for success in web sessions on a page. When you're buying web traffic, you're buying visits to your page.

[0:21:38.8] JM: Okay. Talk about pre-filtered traffic.

[0:21:40.8] SD: Pre-filtered traffic is it's a reaction to the rise of bot detection filters. This is a big misconception on the advertiser's side at least, and I'm sure there's a lot of people confused about this generally is it's not that these companies say that this is definitely a human, and so we're going to let it through.

[0:22:06.3] JM: You're talking about companies like Moat, or —

[0:22:09.7] SD: Moat, or Integral Ad Sceince, or WideOps, Double Verifier, all the leading bot detection companies.

[0:22:17.1] JM: Right. You have a user that hits a page and you have a blob of JavaScript on that page and it will check if that user is probably a bot, or hopefully a non-bot. Does is it give you some probabilistic estimation?

[0:22:31.1] SD: What they are able to do is based on the information and the research that they do is come up with a way to say, “This is definitely a bot.” If it’s not definitely a bot, then it passes through. It can be labeled as some debris of suspicious, but that’s what the blocking is based on is, “Is this definitely a bot by our standards?” Not, “Is this is not a human.” Do you get it?

[0:22:59.5] JM: Right. There’s like a 20% chance that it’s a human then it will get through.

[0:23:04.1] SD: That varies for every company. I can’t say that they let it in for certain degrees of confidence, but that’s the difference is they’re letting things through if it’s not definitely a bot. If, by their standards, it’s definitely a bot, that’s when it gets blocked. It’s not that they’re letting things through based on verifying that it’s a human.

[0:23:22.7] JM: You mentioned being able to buy this pre-filtered traffic for a specific website, can I buy this traffic to come to any website that I have?

[0:23:31.9] SD: Any website, yes.

[0:23:32.9] JM: Okay. I can go to eZanga and I can say, “I want this much traffic directed towards my site.”

[0:23:39.8] SD: You could send it to the Wall Street Journal, or you could send it to CNN, you could send it to Bright Park, or you can send it to eBay. You can send it wherever you want as long as you’re the one paying for it.

[0:23:54.7] JM: EZanga Just says, “This is pre-filtered traffic. It’s traffic that if this traffic hits a page, it will evade the filters that Moat, or Integral Ad Science, these bot detection companies have put up," and so you can have these bots — I'm sorry. These traffic click on anything on this webpage. You could just tell it to do stuff.

[0:24:23.7] SD: Yeah. Basically, they are purporting that we have traffic that if you are being detected by a certain filter, this traffic will pass that filter. That is mostly done in the hopes of generating ad revenue based on these real time betting ad tags that each ad tech vendor uses one or two different bot detection vendors. If I'm a website that's selling into a platform that uses bot detection A, I'm going to go buy traffic that is engineered to pass filter A. If I'm selling into a platform that uses bot detection vendor C, then I'm going to go buy traffic that passes vendor C. There's entire market or pre-filtered traffic, which is traffic engineered to pass specifically —

[0:25:17.5] JM: With this traffic, am I telling it to click on specific ads, or is it just hitting my page and —

[0:25:22.6] SD: It's just hitting. Just hitting your page, yeah.

[0:25:25.1] JM: Okay. Presumably, this type of traffic, when you buy it. — Actually, let's just talk about that. You have a WordPress site, ecelebnews.com, this is a site with basically the most useless content you could have. It never changes. I know that because I've gone there a couple of times because I like the name. It never changes. It only has five actual articles and it's clear that this is just a default WordPress template. You've got some stuff about Cameron Diaz, and Angelina Jolie, and 50 Cent. It's eceleb news. It's what you would expect from eceleb news, except that it's even worse, because it never changes. It's not even news.

You purchased traffic from eZanga to visit your horrendous WordPress site; ecelebnews.com. Explain why you did that and what happened.

[0:26:19.2] SD: I wanted to look at the Google Analytics and what showed up, and I also had three different detection organizations, some are independent and some are bot detection companies. We consistently, over a period of seven months at different test periods, found that this traffic is anywhere, at any given time, between 60% and 90% confirmed bots depending on the vendor.

I wanted to see that and get independent verifications that this is overwhelmingly very suspicious fraudulent traffic, but I also wanted to see what shows up in my Google Analytics site, and it all just says, “Filtered by eZanga.” Every source, every refer, all. It’s all coming from their servers. It’s not coming from a previous website.

If you’re putting search ads somewhere, you would assume that somebody’s on a website where these ads are showing up; Google search, either a search engine or a search engine results page within another site, whatever that is. The refers and the sources always came back as an eZanga server.

[0:27:31.9] JM: To be perfectly clear, you are buying traffic. Why would a site buy traffic? Why would this be useful for somebody to purchase eZanga traffic?

[0:27:42.4] SD: If they have ads on their page. If they have programmatic ad tags on their page, they can sell this traffic and generate it — Sell this traffic. Basically, converting the clicks into impressions that generate revenue from ad exchanges.

The basic formula is you look at the number of ads on your page and you look at what you’re paying per click and you balance that out with the CPM that you’re getting, cost per mill, cost per thousand impression. You basically come up with a breakeven point.

[0:28:16.3] JM: As long as every click on my site generates more money for me from the advertising network, then it costs to pay eZanga to send bot to me than I am going to make money off of that transaction.

[0:28:33.9] SD: Right. Exactly. Your cost per click, your breakeven cost per click is your CPM divided by a thousand over your number of ads per page. If you have four ads on your page and if you’re getting $2.50 CPM. So $2.50 for ever thousand ad impressions, that means that the most you can pay per click is one cent. Four ads on a page, $2.50 CPM, you can pay one cent per click. Obviously, we know that there’s lots of sub-penny clicks available in the market. There’s a lot of people making a profit on this model.

[0:29:11.3] JM: How much leverage can you put into this? How much money can you make?

[0:29:16.4] SD: It's unlimited. You don't have to just scale the volume of traffic on a specific site. You can just scale the number of sites that you have in operation.

[0:29:25.6] JM: Right. Okay.

[0:29:26.8] SD: Once you'd get a certain level of hits, you have the Google Analytics to back it up. You just move into getting video ad tags, which were switching from $2 CPM to $12, $15 CPMs with video.

[0:29:43.2] JM: Just to be clear. The scalable model of an ad fraudster is set up ecelebnews.com, set up advertisements to be displayed on that site. Pay for traffic that is cheaper than the ads that I am displaying on my page pay out. I've got an arbitrage there. I lever that up as high as I can go, and then I setup beefnews.com and I do the same thing. I lever it up and then I switch to chickennews.com and then lever that up and you just keep going.

[0:30:19.2] SD: Yeah. Now, you've got gourmet food recipes network and now you operate as the hub of an audience that is interested in gourmet food.

[0:30:31.7] JM: Right. We've explained by WordPress is 40% of the internet now.

[0:30:38.0] SD: Yes.

[0:30:39.1] JM: When you had this traffic visiting ecelebnews.com, how did that traffic perform?

[0:30:47.7] SD: Okay. We did some tests with live ad tags, but I did that experiment last year and the problem with that is you're kind of — You can justify it as a test where you kind of become part of the problem. You're stealing advertisers' dollars and you're not the one that has the opportunity or the capability to go and pay them back.

We did as small scale, but mostly I was interested in just showing what type of traffic this was. We know that it does potentially generate a profit in terms of ad revenue for people if they have enough ads on a page. Their site is set up correctly. I was mostly looking at — Since they are selling clicks to the U.S. Government. They won a service's contract from the General Services Administration of the U.S. Federal Government, and the General Services Administration basically handles things like transportation, and communication, and basic office things for all of the federal agencies.

Let's say, for example, the state departments wants to put a commercial out for recruitment for employment, they will go through the General Services Administration to put that out. The General Services Administration, it's a huge organization. They've got a budget of over $20 billion every year to handle things like this.

[0:32:13.4] JM: Before we get into the GSA stuff, the government contract stuff, you had some interactions with the sales reps at eZanga. What did you talk to them about?

[0:32:27.4] SD: My conversations with them were very frank. They know that I, at one point, was on the fraudulent network side. When I reached out them and asked for traffic, there weren't a lot of questions asked. I basically sent an e-mail saying, "Hey, I want to get traffic." They said, "How many accounts to you want?" I said, "All five."

Down the list of the five prominent bot detection vendors and then I sent an email saying what are the bid landscapes, meaning what are the ranges of the cost per click that I would have to pay for each? I immediately got a response, very helpful. Then, I asked — Again, this is me being devious on my investigation, I asked, "What are the best places to monetize each type of traffic?" Obviously, they sell a lot of these to different networks and they get feedback on what's working and what's not. He was able to give me that information as well.

[SPONSOR MESSAGE]

[0:33:29.7] JM: Do you want the flexibility of a non-relational, key-value store, together with the query capabilities of SQL? Take a look at c-treeACE by FairCom. C-treeACE is a non-relational key-value store that offers ACID transactions complemented by a full SQL engine. C-treeACE offers simultaneous access to the data through non-relational and relational APIs.

Company's use c-treeACE to process ACID transactions through non-relational APIs for extreme performance while using the SQL APIs to connect third part aps or query the data for reports or business intelligence. C-treeACE is platform and hardware-agnostic and it's capable of being embedded, deployed on premises, or in the cloud.

FairCom has been around for decades powering applications that most people use daily. Whether you are listening to NRP, shipping a package through UPS, paying for gas at the pump, or swiping your VISA card in Europe, FairCom is powering through your day.

Software Engineering Daily listeners can download an evaluation version of c-treeACE for free by going to softwareengineeringdaily.com/faircom. Thanks to FairCom c-treeACE for being a new sponsor of Software Engineering Daily, and you can go to softwareengineeringdaily.com/faircom to check it out and support the show.

[INTERVIEW CONTINUED]

[0:35:11.5] JM: Do you have a picture for what the cost of the traffic that you brought in would have been relative to the revenue that you would get serving ads to that traffic? I think you said you don't monetize ecelebnews.com because you don't want to take advantage of the ad network, so you don't want to be part of the problem. Do you have an idea for what you would have gotten paid if you had those?

[0:35:35.9] SD: The small, small test that we did was about $2 to $2.50, depending on the day, CPM, which would have made a profit. Depending on the volume and how much we ended up buying, it could have easily made — You could turn $200 into $300 on a daily consistent basis. That was with a very small scale test.

[0:36:00.9] JM: Incredible. This is one day turnaround time to arbitrage $200 into $300.

[0:36:06.8] SD: Yeah.

[0:36:07.1] JM: Okay. Let's talk about the U.S. Government contract. The U.S. Government has contracted out all of its pay-per-click advertising to eZanga. Explain why this is problematic.

[0:36:20.5] SD: I can't say for sure whether it's of. I know that they are one of the vendors, because I don't have the full insight into — I'm sure somebody internally is still using Google, or Bing, for pay-per-click ads, maybe even Facebook. They are the one ad network that is outside of those giant companies that has won a contract. That's based on a lot — Basically, a year of press that this company did on how they are a click filtering agency and a fraud prevention network. All the while, they are selling pre-filtered traffic at sub-penny prices.

I know that given the size of the GSA, something is bound to slip through at some point. This only caught my eye, because I had been watching and I had a Google alert for eZanga. When they released — They had that press wire in September of 2016, it showed up immediately and I couldn't believe what I was reading. I've been watching it since then.

I looked over the services contract, everything, and they are selling pay-per-click advertising to the government at 10 cents per click, 10 to 20 cents per click. That's very different. Assuming it's the same sources of traffic that they were getting me, that's a huge bump up in their margin, because I'm getting 4/10ths of a cent to 1.2 cents and they're selling for 10 to 20 times that to the government. It's a big, big issue in transparency.

[0:37:56.6] JM: You're saying that not only did the U.S. Government make a huge purchase that presumably would get them some sort of economy of scale on that purchase. They were actually just paying more for each click.

[0:38:11.9] SD: For what it's worth. You could spend a dollar, or $2 per click on Facebook or Google and say that you're getting, whatever the market value is. Obviously, with this ad network, there's the same type of traffic available for much, much less if you know how to approach the conversation. Obviously, sub-penny click prices are inherently suspicious, and 10 to 20 cents seems like you're doing some public service by giving them a discount.

[0:38:40.6] JM: Is the product that the government is purchasing here the same product that you bought for eceleb news?

[0:38:48.2] SD: That's what it looks like. They're not selling — They don't have a separate safe network of network. I've gone through all of their different sources that they have and none of it is legitimate in any sense.

[0:39:05.1] JM: What is the internal narrative for the person at the U.S. Government who made this purchase. Is it literally eZanga is saying that they're going to get more people to view our ads. Is that the narrative, or do they also —

[0:39:21.1] SD: Or visit something like a recruitment page, or an awareness campaign, or anything like that.

[0:39:26.8] JM: I see. In this sense, the U.S. Government probably use eZanga as an advertising agency. The U.S. Government probably sees it as, "Oh! eZanga is going to market our recruiting webpage. That sounds great to us." When in fact all that's happening is eZanga is sending "traffic" to that page, which is probably bots sitting in an Amazon web services server somewhere.

[0:39:54.5] SD: Datacenter traffic, yeah, all coming from a server.

[0:39:56.5] JM: It's datacenter traffic.

[0:39:58.1] SD: What we've found is that the general trend we see is about 90% of it is kind of very blatantly suspicious, and 10% of it looks like it's human, but we think that's just pop-unders from either unknowing sites, toolbars, or porn sites.

[0:40:21.2] JM: Right. Here, we're talking about the traffic laundering problem, and I talked about this recently on Software Engineering Daily where you've got this issue where — This is actually where I feel it impacts the average user, because the thing that people want to believe, or a lot of people do believe is that, "Oh, the recent election caused these shadowy Russians to set up propaganda, because they wanted to influence the election." Maybe that's true, but it doesn't actually have to be that complicated. All it has to be is people set up link bait. People get emotionally — Russians, or Americans, or whoever set the Denver Guardian, sets up link bait that Hilary Clinton is a lizard alien and they get people to go to that because they get emotionally charged. Then, they're able to get additional leverage out of that organic human traffic by also sending bot traffic. You've shown me a variety of these schemes. You've shown me a very complex Facebook scheme that I'm hoping to talk to you about in the future that is horrifying.

[0:41:36.7] SD: Sometimes, it's that you get some human traffic and then you'd throw in robotic traffic to fluff up the numbers. A lot of times, especially with social media sharing, is humans are more likely to interact with it if they already see that, "Oh, this has been shared 10,000 times and it's got 50 likes," or on Twitter you see that, "Oh, it's been retweeted 25,000 times and it's been favorited another 25,000." That will make you more likely to engage with it if you see or feel, if you're under the perception that it's being engaged with by lots of other people, it gives it some illusion of legitimacy, and that's what fueled a lot of that problem.

There are that scheme that we've gone over several times where they're just constantly throwing in robotic traffic as engagement fluff, but there's also where you don't even pursue human engagement until you've gotten to a point where it looks legitimate.

[0:42:42.0] JM: Just to take a step back here, because I feel I need to keep doing this, because at this point I've been reporting on ad fraud for like six or nine months or something. I don't know if you felt this way when you started reporting, you're like, "As soon as I start talking about it loudly enough, people will listen. People will start to care. They'll start to realize this is a giant arbitrage, and the way that the market treats arbitrages as we saw with the mortgage crisis, is it levers them up insanely, in such a dramatic fashion that is collectively horrible for society. Yet, nobody seems to care."

I've talked to people in mainstream press in the edge of "press". I guess Software Engineering Daily now has to be the press. Not to laud myself, and not to laud you too much, but nobody else is talking about this. You talk to people at the government to try to alert them of this. What was their response?

[0:43:36.4] SD: The response was obviously great concern. Some organizations were worried about; do they have the jurisdiction to do anything about this? Others were wondering; is the money big enough to really pursue? It's sketchy territory, and then it's also — It's an unpaved path. So there's not a lot of precedence in here. Somebody who's going to be choosing to pursue a legal action on any type of ad fraud case is going down a path that hasn't been walked before.

I do understand the hesitation, because a lot of legal proceedings rely on precedence which makes things a lot easier, but with ad fraud, it's just not there. I do want to keep beating the drum. I did take this story initially. As soon as I saw that press release of the government contract, I took it to a lot of the major press outlets. Obviously, they were interested, but they were like, "It's not fully there. It's not fully juicy."

Now, I don't think we got to this yet, but the new development is that eZanga is launching their own bot detection service, and that's when some people have definitely — And it's piqued their interest, because that is sending a much — It's blaring a big, big horn right in their faces that they're going to be on the — They're going to create the supply and also be in the business of verifying that it's good.

[0:45:06.2] JM: The fox is guarding the henhouse.

[0:45:08.3] SD: Yeah, basically.

[0:45:09.5] JM: You got this report that you shared with me. It's like 16 pages documenting this stuff. Have you talked to eZanga about this yet?

[0:45:17.7] SD: I have not.

[0:45:19.6] JM: I know you have done — There's been other cases in the past where you've looked at these scams, one of the other scams that I'm hoping we can discuss on Software Engineering Daily at some point in the future, where you have emailed people at the company, you've been like, "This is what I'm seeing when I analyze your traffic." Only from the outside looking — You do analysis from the outside looking in, and you look, "Here's what happens when I buy your product. Here's what happens when I interact with it from off the shelf services. Here's how much I can make off of it. What do you think of my findings?" Basically, their response is, "Ahh —" Nothing.

[0:45:51.6] SD: Yeah. I’ve had somebody on the board of advisers of one of these big traffic scams and we reached out to him on LinkedIn and said, “Hey listen, you’re obviously a very well-respected person in advertising community. You’re on the board of this company, which is a very, very sketchy. I’m an ad fraud researcher. I’d like to share what I found with you,” and I got a simple message back saying, “No. Thanks.”

With eZanga, they obviously will see this report at some point. They know I’m in the business of ad fraud research, yet they have shared their information with me. I do have to say that they probably could do a really good job of bot detection. Given that they’re in the business of selling traffic. They have seen the worst of the worst of traffic quality. I don’t doubt that they would be good at bot detection, but it’s the — With any type of verification, you’re selling trust, and you’re supposed to depend on the integrity of the company, and I just — I can’t get behind the fact that somebody who is going to verify the quality of something also is in the business and has been in the business for years of selling extremely low quality web traffic, non-human.

[0:47:13.2] JM: By the way, the way that these companies work — I don’t know if you ever saw the Enron movie, but it’s like these companies — It’s often like one or two people at the top, or three people at the top who really understands what’s going on. They really understand how sleazy their business is, and then you’ve got 100 to 1000 people who are just like, “Hey, we’re a bot detection company, or we’re a marketing integration company,” and they only understand some subset of the company, because the broader landscape is so complicated. There are people who are probably listening to this episode, they’re like, “Wait. What the heck is going on with traffic? What is traffic? What is an arbitrage—” They don’t understand the big picture, because it takes a long time to really understand it. You could talk to people at eZanga, they’ll probably be like, “What are you talking about? We’re just a marketing company.”

[0:48:03.7] SD: For sure. That’s happened with sales people there. I’m as close to 100% sure as you can be that these people at the top know what’s happening.

[0:48:14.1] JM: The same was probably true for — What were these horrible mortgage companies during the mortgage crisis. I’m trying to remember the names of them, but these companies that were just — I'm sure some of these sales people were probably just like, "We're just giving people cheap mortgages. We're giving them a home."

[0:48:28.3] SD: Have you seen The Big Short? There's a scene where the guys — Steve Carell's character, his firms goes down to Florida to basically meet mortgage brokers and they're talking about how they get these applications approved, and no down payments, the fudge the numbers, they get these bonuses. One guy is like, "I was living paycheck to paycheck, and now own a boat." Steve Carell's character was like, "What are they doing? Why are they admitting to all of these?" One of his employees were like, "They're not admitting to anything. They're bragging." They don't see the bigger picture, the problem here.

I've seen this with — You see brilliant engineers working at companies that do massive amounts of ad fraud, and I just think they spend their days coming up with amazing up engineering solutions to little issues. They don't understand the implications that this is serving the mass tsunami of ad fraud, but they're smart engineers and they are solving problems that get presented to them on a daily basis.

[0:49:31.0] JM: We were just talking about eZanga. Now, you're saying that the type of culture that we're describing for eZanga where you've got just a few people in the company who understand the big picture, understand how much fraud there is. It's the same thing that is true for companies that begin with a G or begin with an F.

[0:49:51.5] SD: Yeah, sure. Conceptually, yes.

[0:49:53.8] JM: Conceptually.

[0:49:54.7] SD: I can't say that that — I've not seen any direct evidence of that obviously, but yes.

[0:49:59.8] JM: Sure. Okay. This gets to the two types of companies that conduct ad fraud as part of their business model. We've got blue collar companies and white collar companies. Explain the difference between blue collar and white collar ad fraud companies.

[0:50:16.4] SD: I always try to come up with things that make this easy to understand, just like something in ad fraud that you can relate to. I basically tried to separate ad tech companies, whether they're traffic vendors, or publishers, or exchanges, into white collar and blue collar crime.

Blue collar crime, we think of as drugs, or theft, or violent crimes, and white collar we think of financial fraud, embezzlement, investor fraud, those types of things. White collar ad fraud is when you are not in the direct business of transacting on non-human traffic, or fraudulent advertising, but you do in the end benefit from it, and the majority of your revenue is not from non-human traffic.

Let's take a digital ad agency that takes 15% of their client spend as a management fee. They are what I consider generally white collar ad fraud companies, they are involved in the consumption of fraudulent advertising, which means that they are the ones buying it. They have direct brand and budget access. They sort of remain intentionally blind to the possibility of the existence of non-human traffics. They are the ones kind of sticking their head in the sand.

Generally, the larger companies, generally, they have public exposure and a big name. Again, the majority of the revenue is not from non-human traffic. The common thing that they have with blue collar ad fraud companies is they profit from the buying and selling of this fake traffic. Blue collar ad fraud companies are not involved in the consumption of the advertising supply, the fraudulent advertising supply. They're involved in the distribution, or the creation of it.

They have lowered direct access to the brand budgets, which means that they're sitting father away from the actual advertiser. They're fully aware of the possibility of non-human traffic. Generally, there are smaller companies that do tens of millions of dollars in revenue, but you've probably never heard of them because there's no real reason for them to go out and seek public exposure. This is the other difference between them and the white collar ad fraud companies, is the majority of their revenue comes from non-human traffic.

[0:52:42.4] JM: As you know, the main source of revenue for these advertising fraud schemes is the brands who are ultimately paying for this. I'm not just referring to brands like Nike, or American Apparel, but also "small business". My company for example, I have another software company, a different company from Software Engineering Daily. When we launched recently, we did some ads on Facebook, and we targeted white micro — I thought we did very narrow targeting. I guess I'm not a Facebook expert, but I've done so many shows about ad fraud, I thought I understood what we would need to target to some degree.

Every single one of the interactions that I saw looked like it was a bot account, it was from some very strange name from India. I was targeting people in the United States. Very quickly, it was like, "Okay. Turn it off. This is useless. It doesn't get us anything. It gets us nothing." Maybe I'm just unsophisticated. I don't know.

The brands that you speak to on a regular basis, is the tone changing? Are people starting to say, "Oh my God! We've been wasting money for such a long time."

[0:54:03.3] SD: I always caution them on overreacting. What ends up happening is somebody in the marketing department of a brand — A light bulb goes on in their head, they're like, "Oh my God! We need to control this ad fraud problem. We need to cut off our exposure to ad fraud and we're going to say that we have a zero tolerance policy for fraud," which I think is an emotional reaction and it's an overreaction to the problem, because you're not actually addressing the issue, which is fraud, which fraud exists everywhere and will always be there. You should try to make sure fraud or at any industry is a minimal percentage.

You see estimates of fraud in advertising, digital advertising of — Some people say 10%, some people say 20%, as high as 30%, 35% of the advertising supply, they say it's fraudulent. That would not be acceptable in any industry where you say 10% or 30% of the commodity is fraudulent. We should try to get to, what I'd say, is 2% to 5% is a healthy range, because there's always going to be sophisticated fraudsters who will always find a way to commit it, you just have to make it as hard as possible.

With advertisers waking up to the existence of this problem and how widespread it is understand that it's a systemic problem. It's not that you need to immediately change agencies, or you need to immediately shut off your campaigns. One of the general overreactions I see is, "Oh, we're just going to shut off programmatic, because that's where the fraud is." It's like, "No. That's not where the fraud is. You're going to see fraud in direct buys too, but you need to make sure that you're operating safely and asking the right questions. Having processes in place. Fixing fraud within your advertising campaigns is not a flip of a switch. You can't just say, "Okay. I'm detecting with this vendor and now I'm all okay," or "I've shut off these sites and blacklisted them. Now, I'm all okay." That's not that simple.

[0:56:14.6] JM: You said direct buy. What is direct buy?

[0:56:17.3] SD: Direct buy would be instead of me going to an ad exchange and typing in on my whitelist, meaning the sites that I want to target and show ads on; New York Times, and Fox News, and Huffington post, and rather reaching out as an agency, reaching out directly to Huffington Post and saying, "I would like to buy 2 million impressions over the course of the next two weeks."

Direct buy to the publisher, which — I'm not saying Huffington Post has a lot of fraudulent traffic, I'm just — Even if you're doing direct buy, you have an exposure to the fraud.

[0:56:56.3] JM: You could even imagine Software Engineering Daily, we air and ad for some giant — Let's say it's SAP. Let's say SAP buys a million listens on Software Engineering Daily, that would be a good reason for me to go to a bot network and say, "I would like a bot network to listen to a bunch of podcast episodes please." It would say, "Okay."

[0:57:15.6] SD: There are traffic vendors that offer that exact service. SoundCloud listens, podcast listens. That service is available.

[0:57:23.1] JM: This is something — Recently, I released an album on Spotify, and I had some bots listen to it, because I was like, "I wonder how many bots —" I didn't buy that much fake traffic, and I hope Spotify doesn't blacklist me for this. It was just an experiment. I only bought, I think, 500 listens for $5. I just want to try to it.

[0:57:44.5] SD: It's not bad, one cent per listen.

[0:57:46.0] JM: I think that's what the deal was, but I was like — The experience made me think, "Is this what Drake is doing? Is this what's running the internet?" I listen to some of the music at the top of the charts, I'm like, "This is garbage." Is it just bots?

[0:58:01.6] SD: I've never actually looked into the amount of bots on these streaming services, but that's definitely a service that's available. A lot of traffic networks and bot softwares you see, "I want YouTube likes, or YouTube views, or SoundCloud listens, or podcast listens." Whatever it is. Yeah, it's definitely available. I can't say based on any evidence, how bit it is, but it's definitely available and there's obviously the incentive for creators of the content to access it.

[0:58:38.4] JM: Okay. Just to wrap up, one more question, and I know we're up against time. This YouTube stuff recently where advertisers are pulling their ads off of YouTube because of scandalous content. For example, there's a Jihadi video and then you get an Estee Lauder ad served against it, and Estee Lauder is pulling their traffic from YouTube because of this problem. Is there any chance that this kind of thing spirals into a closer investigation of programmatic traffic, or fake traffic, because these are totally different problems, but you could imagine one spiraling into another.

[0:59:13.8] SD: I'll defend Google on this. Free consulting everybody that's listening, is it's user-generated content on YouTube, and you have a choice in any online advertising platform to run on a whitelist, which means I'm going to only show ads on specific YouTube channels in this example, or have a blacklist where I'll serve anywhere where my user targeting fits, except for these specific channels.

Now, what my advice to everybody whenever I talk to someone is you have to have both. You have to have a whitelist campaign where you know that these are vetted channels that I like. I vetted the content. Overtime, it's performed well. It gives us return on our investment by showing ads to these users, and then I have another campaign where I'm obviously paying a lower rate, but I'm blacklisting sites and channels as I find that they're bad. Then, taking the good ones and adding them to my whitelist.

The issue with user-generated content, and this is a YouTube problem, is nobody that is uploading hate speech or extremist violent content is going to label it hate speech. They are going to label it as entertainment, or education, or news, or whatever they want to do. That's the issue with user-generated content. One of the policies that Google introduced was you can't start monetization until you have 10,000 views, which is not a bulletproof policy, but it does make it more difficult for these channels to monetize.

Again, the issue with this extremist content, this was the big spark plug for all of these public discussions was extremist content, I think it was ISIS related that was getting advertisers' ads on it. They're not doing this to buff up their number and generate $15 from YouTube ads. They're doing this because that's just a place to post videos and people can watch them there. Advertisers, I think, need to be weary of the fact the some of it is an overreaction.

You have to give Google a little bit of leeway on user-generated content. I think they're doing the best that they can. I don't think they want to be in the business of monetizing hate, or violence. Yeah, I just say caution on the overreactions, because we create policies as a react and we always do an overcorrection, and I think that's going to be dangerous in this sense.

[1:01:42.5] JM: All right, Shailin. I want to thank you for coming back on Software Engineering Daily.

[1:01:45.5] SD: Of course, my pleasure. Thank you for having me.

[1:01:47.1] JM: It's a pleasure. Please as always, and I look forward to hearing about your schemes, or other schemes you're investigating in the future.

[1:01:57.0] SD: I definitely look forward to coming back and sharing that.

[1:01:59.5] JM: Okay.

[END OF INTERVIEW]

[1:02:02.7] JM: Thanks to Symphono for sponsoring Software Engineering Daily. Symphono is a custom engineering shop where senior engineers tackle big tech challenges while learning from each other. Check it out at symphono.com/sedaily. That’s symphono.com/sedaily. Thanks again Symphono.

[END]