

EPISODE 788

[INTRODUCTION]

[0:00:00.3] JM: Advertising fraud occurs when a brand pays for an advertisement online and that advertisement is not shown to an actual human. The advertisement ends up being shown to an automated bot account that has been created to view ads.

Advertising fraud is rampant on the internet. It's not possible to know how much money is lost to fraud, but the costs are in the billions of dollars. Praneet Sharma and Shailin Dhar are the founders of Method Media Intelligence, a company that builds solutions around improving advertising quality. In previous shows, Praneet and Shailin have described the online advertising ecosystem in detail. They've told stories of bot farms, replay attacks, ad tech companies.

In today's episode, Praneet and Shailin return to the show to discuss how advertising fraud is getting worse, not better. Praneet and Shailin worked with BuzzFeed reporter Craig Silverman to produce some of the remarkable findings about mobile advertising fraud, which Craig came on the show to talk about a few weeks ago; the mobile advertising fraud schemes account for hundreds of millions of dollars of theft every year. It's great to have Praneet and Shailin back on the show and I hope you enjoy it.

[SPONSOR MESSAGE]

[0:01:29.3] JM: HPE OneView is a foundation for building a software-defined data center. HPE OneView integrates compute, storage and networking resources across your data center and leverages a unified API to enable IT to manage infrastructure as code. Deploy infrastructure faster. Simplify life cycle maintenance for your servers. Give IT the ability to deliver infrastructure to developers as a service, like the public cloud.

Go to softwareengineeringdaily.com/hpe to learn about how HPE OneView can improve your infrastructure operations. HPE OneView has easy integrations with Terraform, Kubernetes,

Docker and more than 30 other infrastructure management tools. HPE OneView was recently named as CRN's enterprise software product of the year.

To learn more about how HPE OneView can help you simplify your hybrid operations, go to softwareengineeringdaily.com/hpe to learn more and support Software Engineering Daily.

Thanks to HPE for being a sponsor of Software Engineering Daily. We appreciate the support.

[INTERVIEW]

[0:02:52.0] JM: Shailin Dhar and Praneet Sharma, guys welcome back to Software Engineering Daily.

[0:02:56.7] PS: Thanks for having us, Jeff.

[0:02:57.7] SD: Thanks.

[0:02:59.3] JM: You're the founders of Method Media Intelligence. You've both been on the show. Shailin's been on several different times. I love having you guys on, because you are very well acquainted with advertising fraud. We'll talk about plenty of new material. Just in case there are people who are listening who haven't heard the previous episodes about advertising fraud, why don't you start by just giving us an overview of what advertising fraud is.

[0:03:26.1] PS: The universal umbrella definition of ad fraud, there's two sides; one is the industry accepted version, which ad fraud as the umbrella term covers non-human traffic, lack of view ability and also brand safety violations. For example, an ad appearing next to content that can be deemed unsavory, or inappropriate, like a car ad next to a new story about a car crash, or airline ad next to a new story about an airplane crash, so those types of things. What we focus on is making sure that advertisers are getting the commodity that they intend to purchase and anything outside of that can be deemed ad fraud.

[0:04:08.0] JM: In the process of advertising fraud, who gets defrauded? Who suffers from advertising fraud?

[0:04:13.4] SD: Mainly advertisers, right? Advertisers that are actually trying to maximize their impact, their marketing dollars that are being impacted most definitely. There are also publishers that are trying to write meaningful content and they have to compete and get diluted with garbage publishers that write fake news articles and don't really have any real meaningful content.

[0:04:34.1] JM: Let's say I am a black hat business person on the internet today, I want to build a business around making money from advertising fraud. What would be a good strategy?

[0:04:45.0] PS: I think there's various good strategies. I mean, Shailin can talk about a few. The one strategy that I've seen proliferate is just driving invalid traffic, right? Getting a bunch of ad tags and driving invalid traffic to let's say just a fake site that you've created, just to load those simple little banner ads that could really add up. There's so many ways to monetize ad tech now. You have cost per install and cost per acquisition, you have cost per click of course through search ads, but there are many other channels that allow cost per click.

Yeah, any action on the web that is deemed monetizable, you can essentially gain with invalid traffic, or even just real users that are just unintentionally loading ads in the background. You see this a lot with the Android ecosystem is a lot of malicious apps loose run in the background and load ads.

[0:05:29.1] SD: There's always two sides to an ad fraud operation. One is the actual generation, or creation of the invalid traffic so actually creating those ad requests. The second part is finding one place, or multiple places to actually monetize those things. There's the creation of the website, there is the funneling of traffic to that website or just generation of ad request, and then there's doing the black hat business side of getting onboard into ad exchanges, which is actually where I consider the fraud to be. Because there's nothing immoral, or legal about running a bunch of ad requests to your own site. I mean, it gets done in the QA process by web developers and app developers all the time, but it's the actual pushing of those ad requests into a live at exchange, where marketers are spending money that is the fraudulent part.

[0:06:25.8] JM: What percentage of ads today do you think are ultimately viewed by bots, or are presented in a way that I don't know, ad stacking, or basically presented in a way that undermines their usefulness? Do you have a perspective on how many ads are actually getting viewed by humans, or are reaching humans in a meaningful way?

[0:06:50.1] SD: Yeah. If you bucket in what we quantify as waste, we've seen data that has about just from our own product is from 25% just across the board. Ads that never show up is the biggest one, right? Ads that never rendered. You can't have any view ability or any measurable view ability on anything that is a render. Then there's invalid traffic where we've seen 5% to 10% just across the board. Then you get an in-app where we've seen upwards of 70% invalid traffic sometimes.

[0:07:18.2] JM: Let's talk about some of the stack of technologies that go into rendering an ad. Let's say I go to the washingtonpost.com, an ad loads on the screen. Describe the stack of technologies that have led to that ad being purchased and rendering on the page.

[0:07:37.2] SD: Yeah. Washington Post is a good example. Many of the premium publishers, they started using what's called header bidding and that's just a set of JavaScript libraries that are on the page. Those JavaScript libraries essentially make requests to the ad tech ecosystem. Those requests first go to what you call supply side platforms and these are the servers that a supply side platform could be something like automatic, that's an example, open X. Then they also go to the exchanges, which is double click and AOL, like these big exchanges that connect the sellers and the buyers. Those are also just a set of servers.

Then those requests go to finally be buying platforms. Buying platforms, these names might sound familiar, there's TV 360 owned by Google, there's Medium app, Trade Desk. These buying platforms essentially run the auction process. This happens in about 50 to a 100 milliseconds, but there's a lot of room of error and a lot of room for failure.

What will happen is these buying platforms based on some simplistic algorithm will bid a certain amount what's called a CPM is okay, what's a cost per thousand that I want to actually bid? Finally, when someone wins the page which is Washington Post finally gets a win notice. Like

hey, this advertiser from this buying platform from this exchange, one here is their ad tag, here is their creative that you need to render this HTML and JavaScript.

Then the actual page Washington Post, their JavaScript libraries will have to load this ad tag and it will then make requests to what's called ad servers. As you can see, this is an oversimplified version of this, but there are many round-robin requests that go on and essentially, that opens up this avenue for ways to exist, because at any one point, a system could fail and the ad would never render, or it would bug out. This is a very common thing in ad tech, because you can essentially say these systems are duct-taped together.

[0:09:32.1] JM: Right. It's like, if you were to design a web page that had a bunch of images associated with it and you're making a bunch of asynchronous calls to remote servers that host those images in different places, some of those requests might fail occasionally and that would lead to the image not being rendered. Then the more hops that are needed between any of those individual image requests increases the likelihood that one of those is not going to fail, or one of those is not going to work and that discrete image will not render and what is in ad, but an image ultimately.

If we talk about that whole bidding process, the ad is going to get rendered on Washington Post, what about the tooling for knowing that I'm a human? What tools does an organization like Washington Post using to detect that I'm not just a bot or a web scraper?

[0:10:26.1] PS: Typically across major publishers, as well as the advertiser side most companies will be using a bot connection technology that uses behavioral signals to assess whether a user is displaying human-like behavioral qualities or automated bot-like behavioral qualities.

What ends up happening is there's always a traffic creator who feels threatened by updated bot detection. What they do is either they'll get access directly to these platforms and start reverse engineering them, or in some cases they'll just block these technologies from even scanning them when they visit the page.

[0:11:07.3] JM: We have a chicken-and-egg game between the bot detection companies. If I'm Washington Post, maybe I'm purchasing a provider that has a script on my page that is protecting me somehow from these bots. We have a chicken, or a cat and mouse game between the bot detection companies and the companies who are – or the individuals who are developing these bots, who are learning to replicate behavior that evades the bot detection software.

[0:11:39.1] SD: I don't really like the cat-and-mouse game analogy. It's more like a blind lion trying to catch 200 mice. The issue with this equal back-and-forth and a cat-and-mouse example is most of the datasets from any bot detection vendor or technology are going to be sampled. They have very heavy checks that they do and there's – they take significantly longer than any programmatic or digital ad transaction.

To avoid actually hindering ad delivery, they have always sampled impressions, or page views that they actually scan and make decisions on. When it comes to that, you don't even need to let the bot detection run a 100% of the time, that's why a lot of these traffic vendors will just block the technologies on the page.

[0:12:29.0] JM: Let's say we're talking about a random game from the app store. Let's say we're talking about like a animal trivia game. A trivia game about animals, there's some ads that run at the bottom, you see a different ad every time you have a different animal trivia challenge. How do these ads get to me? Is the stack of ad technology on mobile app ad deliveries, is that basically the same as what we discussed with Washington Post?

[0:12:58.8] SD: It's the same, except with mobile apps. In-app advertising, you just have a – you have a lack of measurability, because instead of using JavaScript libraries, they're using two things; they're using SDKs that are bundled with the apps, but the actual ad creatives are loading in web views. Those are componentized versions of Safari and Android Chrome.

Unfortunately, those web views don't have the same level of measurability that they do on the web. They don't know when app they're landing on and then they don't know where, so there's a mismatch standard. Also, the privileges that the app has is way higher than any web view would possibly have, just because of the operating system layer.

What you have in in-app is a big, I would say just a big breach in a sense, because there's no – there's no cross-origin security. The app can do whatever it wants with the web view, can trigger clicks, it can do click injection, they can load ads wherever it wants in the background and then spoof view ability. There's a ton it can do. There are steps by Apple and Google to alleviate this. One is standardized view ability, which is called intersection observer. Hey, is this truly in view? Now you have a very native functionality, where web view could say, “Hey, I am in view or not?”

Still, it doesn't defeat the other problems around especially something that affects install attribution is click injection and click spamming. This is where in-app advertising, the whole bidding stack is the same, but the problem is the actual client-side stack is all problematic.

[0:14:29.5] JM: What's the perception of mobile app advertising versus desktop display advertising? My sense is that people think that mobile app advertising is perhaps somehow safer, or is a better way to target users. Am I wrong about that? How did this –

[0:14:50.9] PS: It depends on who you ask.

[0:14:52.3] SD: Yeah, and with mobile – what happened with mobile is what you'll see in general is the CPMs, the cost will be higher almost all the time, because they think it's just – because maybe it's closer to your face, I don't know. That's the perception that people have is since it's a mobile device that you should bid twice as high for some reason.

[0:15:12.6] PS: The perception is that this is more captive human attention. There's not as many distractions on your three and a half, four-inch phone screen as there is one a 15-inch computer monitor. There's that perceived extra quality factor there. Then there's also a misconception that well if it's in-app, it actually has to be on somebody's device. What we cover a lot of times in any initial education seminar, or even just introductory session is the topic of device emulators that can be run in the cloud.

[0:15:49.2] SD: Yeah, and the Android emulator is the one that sticks out, because believe it or not, and many Android devs will know this already is you have access to the Play Store. I can install from the Android emulator Facebook, Pandora, any app from the Play Store and then just

run it. I can run that in the cloud. What's the reason for that? I don't know, but it's a QA testing tool that everyone uses and you can definitely run it in all the traffic.

[0:16:15.2] PS: Any QA testing tool that's available, you can always bet your money that somebody on the ad fraud side is using that to go there.

[0:16:22.4] JM: Are there tools for both Android and iOS to do emulation in a data center?

[0:16:27.4] SD: Only Android. iOS emulator requires the Apple ecosystem. Apple has made a lot of efforts to actually restrict advertising with ITP, intelligent tracking prevention. They have no sense of what install attribution even is. Whereas, the Play Store has promoted install attribution now to appease everyone who's buying CPI traffic. It goes to show you that Apple does not make money through advertising. I mean, even though they're getting into it, the bread and butter revenue is coming from selling products.

[SPONSOR MESSAGE]

[0:17:04.1] JM: Failure is unpredictable. You don't know when your system will break, but you know it will happen. Gremlin prepares for these outages. Gremlin provides resilience as a service using chaos engineering techniques pioneered at Netflix and Amazon.

Prepare your team for disaster by proactively testing failure scenarios. Max out CPU, black hole or slow down network traffic to a dependency, terminate processes and hosts. Each of these shows how your system reacts, allowing you to harden things before a production incident.

Checkout Gremlin and get a free demo by going to gremlin.com/sedaily. That's gremlin.com/sedaily to get your free demo of how Gremlin can help you prepare with resilience as a service.

[INTERVIEW CONTINUED]

[0:18:03.2] JM: Earlier, you guys mentioned that term waste, ad waste. There's waste that can happen as a result of just Byzantine failures. You've got a distributed trace of requests that you need to work in order to get your ad rendering on the page. Sometimes servers just fail. Packets

get dropped. It happens. A lot of this waste seems to be coming from a more deliberate, more perhaps malicious series of events that is actually what we term ad fraud and what I've had you guys on to discuss so much.

What makes Method Media Intelligence, the company you guys run so interesting is that you are devoted to analyzing ad fraud, studying it, publications around advertising fraud. You've done some work that has helped people who are journalists at BuzzFeed and New York Times. I think of you as investigative journalists/engineering activists/ruthless capitalists who are building your own business. I find you very intriguing characters. Can you just tell me a little bit about Method Media Intelligence? What do you do? What is your modus operandi?

[0:19:22.0] SD: Yeah. Method Media Intelligence was founded by Praneet and myself in the beginning of 2017. One of the first things we did as a company was a few commissioned research reports for clients, basically investigating the effectiveness of ad platforms, buying platforms that they were using.

One of the things that happened soon after that after a lot of internal research was our filing of a bot detection patent that focuses on what can be proven objectively about the machine, rather than any behavioral signal analysis that we had seen with most bot detection vendors in the past. The significance of that is that this check for me was able to do in 5 milliseconds and that led us to eventually create our main flagship product, which is proactive auditing, which was essentially ad monitoring for digital advertisers.

Making sure that we combine the principles of objective bot detection using the actual device signals and combining that with basic accounting principles of I need to have receipts of every transaction that I engage in.

That's what we're working on. Probably I would say 80% of our time gets spent on working with clients around this product, which essentially it's a merger of a few different ideas that have always been around in ad tech, but this is the first time that advertisers have had a very low-cost, non-volume based solution to track all of their digital ad transactions and then also have corresponding quality metrics around each of those transactions.

Our goal when we were working on this was how can we move away from only telling advertisers that they had 5% bot traffic, or 10% bot traffic, or 1% and actually tell them how much money went to these categories. We do that by collecting our own logs, because we can make decisions in 5 milliseconds on every impression and merging that with the demand side platform blogs, which have the price information for every single impression.

Now rather than getting a report that says 5% bot traffic, because that doesn't always equal out to 5% of your campaign budget. There are some ad impressions that get sold for \$10 CPM and then there's others that get sold for \$190 CPM. It's not always easy to back up how much money was actually lost to invalid traffic, unless you're merging those two data sets.

[0:22:02.6] JM: Can we do some jaw-dropping anecdotes now, because I know you guys have worked with some companies who have – I'm not sure to what degree you can talk about your particular clients, but just hearing anecdotes from you guys, even anonymized ones, there just some insane malfeasance that goes on that – I mean, that's ultimately the reason that drew – what drew me to this topic is just how much money is being stolen by some people and the machinations, the sometimes hilarious, sometimes frightening, sometimes existentially threatening machinations that lead to theft. Can you just tell us some stories about perhaps clients you've worked with, or people you've seen and how they have gotten defrauded and I guess, perhaps what you can do to alleviate that.

[0:22:56.3] SD: I mean, this is not anonymized for any specific client. This is what we see across clients and across the industry and any initial conversations we have as well, is just a complete lack of access to data in the first place. Forget things about analyzing the data to assess the quality of how horrendous that is. We've seen top 10 advertisers who their ad vendors will a lot of times just blatantly refuse to share transaction data of where their money was spent.

One of the things that we I mean, joke about, but also it was very shocking thing to hear from a vendor was if we show you where we're buying, then you can just buy it yourself; the motto for having a black box technology business.

[0:23:42.6] JM: Another aspect of this ecosystem is that you can end up dealing with the brand. If you talk to the brand, so let's say there's a big brand that's advertising I don't know, let's say it's Applebee's. Let's say it's Applebee's. Applebee's is running a bunch of ads and you go to Applebee's and you say, "Hey, Applebee's. You've got this fraud going on." Even if Applebee's says, "Yes. We know we're the victim of fraud." You may still have trouble politically getting through the Applebee's organization, because you encounter different people in the organization and there are actually going to be people in that large organization that may be okay, or maybe even be happy that fraud is occurring, because it can help – there can be marketers within the giant organization whose KPIs, whose key performance indicators are actually helped by the fact that there is fraud that is negatively affecting the overall spend of the company. Tell me about the difficult politics of working within this this ecosystem.

[0:24:56.6] PS: Yeah. We find it very adversarial. I mean, there's a lot of big advertisers want this ecosystem to change and are taking very public stances against it. Then you have advertisers and their marketing teams that at the end of the day is just analytics and that analytics gets turned on to a new report and that report is tied to a bonus structure and incentive structure.

There are cases where even though the organization as a whole is suffering, the marketing team is not. They're actually benefiting from invalid traffic, or invalid data, or lack of data whatsoever. They can actually just start fudging numbers. This happens and it's slowly peeling away. Yeah, it is shocking to us to see that at the advertiser side initially was very shocking, but now we're used to it. The biggest advertisers in the world, they take a total complete hard stance against ad fraud.

[0:25:46.9] SD: That's again as an organization. What we encountered – I guess it's been two years now where we've been facing this is when we're approaching this problem, we have to understand that this marketing organization, the entire media operation is assigned a budget at the beginning of the fiscal year. They have certain goals that they need to meet with that budget. If that's a certain number of products sold, a certain number of new customer signups, whatever that is, if you tell a marketing department that you can achieve those goals for less money, that's not as attractive as you can double that performance for the same amount of money.

I think that it's an overlooked nuance that actually ends up being very important, where marketing departments are not incentivized to save money at the end of the fiscal year. They are focused on getting better results. Does that make sense? That small ones where saving the marketing department money is not always helpful to them as individuals. They'll get a smaller budget next year. If you can help them increase the return on that media investment, that's always going to be more tracked.

[0:26:59.8] JM: By the way, when we're talking about results we don't necessarily mean results in terms of driving customers going to your Applebee's restaurant. We're talking about results that you can report to your manager and say, "Hey, this number of people viewed our ads. We drove engagement with Applebee's."

[0:27:19.0] PS: Right.

[0:27:19.9] SD: Yeah, it's a lot of means.

[0:27:21.7] JM: We did this show with Craig Silverman recently. Craig has been doing great work investigating the fake internet. I don't know if you guys – to what degree you guys can talk about this, but you worked with Craig some on the investigation, right?

[0:27:35.9] SD: Yeah. I worked with Craig on the Cheetah Mobile investigation and other investigations as well. Yeah, and we still work together on a few investigations.

[0:27:45.6] JM: Okay. Recount what Cheetah Mobile is and why that was worth investigating.

[0:27:51.8] SD: Cheetah Mobile, Craig has a lot more details in the business, but they're a Chinese company that has a lot of apps, like utility apps. You think of just cleaners, antivirus, utility apps, some flashlight apps. They are a major investor in Kika Keyboard app. All these apps are on the Play Store downloaded by a lot of American citizens.

What we found alongside with Kochava was they were running a lot of ads in the background, they were actually gaming install attribution, so actually making money through that. Then there was some other stuff that we found that just why would an app like a flashlight app need so

many permissions, or what else is it logging and where is that information going? We focused more on the advertising angle. Now it is becoming apparent that many apps on the Play Store, you can't even find the business details of the developers and they have 50 million installs. It's incredible to find a selfie app that has so many installs and has so many permissions and Google doesn't really know about it.

[0:28:57.2] JM: The example of a flashlight app; you open it up, you're just using your phone as a flashlight and then in the background it's loading ads and ostensibly would be showing you those ads, but you might not even see the ads rendered on your screen. They're just hidden from you.

[0:29:15.9] SD: Not only that. Why does a flashlight app need to load ads? I mean, obviously you're planning something – you're looking for something and you're probably not looking at your phone, because you're looking for something else using the flashlight. Common sense things that advertisers sometimes forget, or just people that are buying all these things don't get is why does a utility app need ads? If it's supposed to be keeping you safe and everything, it just runs in the background. That's the problem. It runs in the background and it runs persistently.

[0:29:45.0] PS: Just these philosophical things of media buyers need to realize that buying ad impressions on a flashlight app, or a meditation app does not make any logical sense.

[0:29:58.5] JM: The other thing about the Cheetah Mobile story that was so shocking was the idea that you could have these apps that not only are they stealing advertiser money by rendering ads that nobody is looking at, they're also potentially funneling information that is basically surveillance data on US citizens in order to envelop US citizens in the same dragnet surveillance information that is just the norm across Chinese internet users. Am I getting it right? That's the allegation that Craig took, right?

[0:30:37.9] SD: Yeah. I mean, that's the implication there because of the permission level that they have, right? They can record your voice, read SMS, they can get access to your credential information, your account information. I mean, it's all an open game.

[0:30:51.8] JM: To what degree has Google cracked down on this, or either before or after this great long-form journalism piece on Cheetah Mobile came out?

[0:31:01.5] SD: What Google has done is they've responded to these allegations and actually, what they did was I believe they kicked out a few of those apps. Then something changed with those apps and they let them back in. It's been we'll just wait till it disappears.

[0:31:17.8] JM: Do you guys get the sense that Google is essentially overwhelmed by the challenges of the App Store?

[0:31:27.2] SD: No, because then you have this other company Apple that does not have as many issues as them when it comes to this. They still have issues, but not as much as the Android ecosystem.

[0:31:36.3] JM: Why aren't they putting more limitations on their ecosystem?

[0:31:40.4] PS: Apple doesn't own an ad exchange where the banners from the malicious apps will get monetized potentially.

[0:31:46.4] SD: Yeah, plain and simple.

[0:31:48.0] JM: You're suggesting that there is a profit motive that drives Google to indirectly encourage advertising fraud?

[0:31:57.7] PS: It's not that they encourage it. There is no financial or reputational punishment that gets put on them for having these things around. It's not that they're encouraging it. I mean, that would be a pretty big scandal if they were making any suggestions that hey, if you want to be in the Play Store, make sure you got these extra ad units. It's just that there's no harmful financial impact to them from having this in there.

[0:32:23.9] JM: I'm not sure I see the line between encouraging behavior and opening up a casino where the incentives encourage you.

[0:32:35.6] PS: It's active versus passive, right? They're not encouraging it, but they're not doing everything in their power to discourage it.

[0:32:42.4] JM: Let's say I open a casino –

[0:32:44.4] PS: It's relative.

[0:32:45.3] JM: Let's say I open a casino, it's slightly off the Las Vegas Strip. It's you can't see it as well. It's really boring. It's the most boring-looking casino in the world. There's no flashing lights. It basically looks maybe like a doctor's waiting room, it's that level of boredom. You go in and there are some games and the games actually have really good returns. You go in there and you have to gamble, but it's you actually have 60% chance of winning money. It's like wow, yeah, you just go into this somewhat boring casino and you might lose, but most of the time you actually win. I mean, to me that's essentially what the advertising ecosystem is. It's like this casino where it's boring, it's technical, like most people don't really care about it, but the people who go there they're like, "This is just the best casino ever." I mean, is that a fair analogy?

[0:33:39.0] SD: I would say so, because when you compare to a boring casino that isn't known by a lot of people but it makes a lot of money, when you get to ad tech, a lot of people are very protective of this casino. They would tell everyone else that visits that, "Hey, there's no money in here. Go away. This is a boring doctor's office." Then they would try to keep this all amongst themselves.

[0:34:01.2] JM: Here's the thing is ultimately, the operator of that boring casino is responsible for it. I sense every time I love Google. I'm complete completely in love with Google. I'm a tech – I'm so pro-tech, but it's just not to throw stones, it's just this is a casino that's basically operated by Google. I just find it hilarious that for all of the criticism that Google gets, nobody really focuses on the issue that makes Google money.

[0:34:32.6] PS: Exactly. Yeah, this is a big thing because Google, their public image is they do a lot of projects. They have so many research initiatives and everything and those are great. They have DeepMind which takes everyone's hearts and the research that DeepMind comes up with

is great. Then what funds all that is AdWords, double-click, ad sense, all the advertising products.

One might say that hey, they have GCP and GCP is up-and-coming as a revenue stream. It's nowhere close to Amazon or Microsoft, so it's mainly the advertising. Yeah, it's the secretive society that is just funding Google. You will even talk to people like Google and they're like, "Oh, yeah. The ad products. What do they do?" Well yeah, they have very little knowledge of them.

[0:35:13.8] JM: Let's take the generous case towards Google. This is their business. They have to do it. I get it. Apple doesn't have to do this, so Apple's hands are completely clean. Maybe it's just such that this is a really hard problem. Maybe Google is trying to develop a completely honest ad ecosystem. What's the most generous case towards Google?

[0:35:34.1] PS: Yeah. If we were to take the actual generous case, like are they actually trying and is there a lot of effort around this? There are teams at Google that do monitor traffic quality. There's an immense amount of traffic that goes throughout AdWords and double-click in the entire ecosystem. The only thing is they're not very open with other sets of data. The problem is you can go to double-click and say, "Hey, I found out this invalid traffic. Can I get my money back?" They'll be like, "Okay, well we'll give you some make goods. We'll give you this much amount of money and it's this is how much we found without actually disclosing okay, this is what our methodology was and this is why we think this is so."

Then they'll claw that back from publishers and there might be publishers that do deserve that money, but Google will take it and just keep it in the exchange. A lot of times, Google is then pocketing the difference and the actual make good that they decide is deemed necessary. The reason they can do this is at any point in the advertising, programmatic advertising ecosystem, you have to touch Google. There's the DFP service, which is used by publishers that's just their ad server. That's what they have to use across the board.

There's a double-click ad exchange. Everything goes to the double-click ad exchange. Then finally, a lot of people – a lot of big advertisers use DB 360 and that – if they're using a bidder that's not DB 360, it's most likely going through double-click anyways and it connects to the exchange. Once you're using DB 360, you have Google at every point. They know what's going

on, or they can sometimes tell there's just so much data to analyze. No advertiser can really, really essentially tell what is going on.

I mean, the silo is just not on the teams, it's also the data. That's been one of the main problems with Google. Then their response is basically, "Well, what are you going to do? Spend on the other monopoly, which is Facebook?" That's the reason they call it a duopoly is because you can't escape Google on the open internet. They know that. You're not just going to go to some other small exchange, or something. It will always touch Google at the end.

[0:37:31.5] JM: How does the Facebook ad buying process contrast? Is it any cleaner? How does it compare to the doctor's office casino? What's the mirror analogy for Facebook?

[0:37:44.8] PS: It's closed. It does not –

[0:37:47.2] SD: It's a speakeasy.

[0:37:48.9] PS: Yeah.

[0:37:50.1] SD: It's a speakeasy.

[0:37:50.6] JM: It's a speakeasy.

[0:37:53.0] PS: Speakeasy where you go in blindfolded and they tell you how many clicks you have.

[0:37:59.5] JM: I guess on Facebook, do we have any reason to believe that there are more or less invalid traffic, or I guess wasteful traffic expenditures?

[0:38:10.6] PS: I can't go into too much detail here. My thing that I've said about Facebook for a long time now is there's a very simple way for them to address the issue of fake profiles, is they can very easily set up a group of people internally that goes out and buys followers and likes from all of these services that sell it on a massive scale, and use that to constantly remove whatever accounts do that. I mean, it's just setting up a honeypot type situation. I don't think that

the problem is too hard to address within Facebook, because it is closed. It's much more difficult on the open web.

[SPONSOR MESSAGE]

[0:38:58.3] JM: We are all looking for a dream job. Thanks to the internet, it's gotten easier to get matched up with an ideal job. Vetterly is an online hiring marketplace that connects highly qualified jobseekers with inspiring companies. Once you have been vetted and accepted to Vetterly, companies reach out directly to you, because they know you are a high-quality candidate. The Vetterly matching algorithm shows off your profile to hiring managers looking for someone with your skills, your experience and your preferences. Because you've been vetted and you're a highly qualified candidate, you should be able to find something that suits your preferences.

To check out Vetterly and apply, go to vettery.com/sedaily for more information. Vetterly is completely free for job seekers. There's 4,000 growing companies, from startups to large corporations that have partnered with Vetterly and will have a direct connection to access your profile. There are fulltime jobs, contract roles, remote job listings with a variety of technical roles in all industries, and you can sign up on vettery.com/sedaily and get a \$500 bonus if you accept a job through Vetterly.

Get started on your new career path today. Get connected to a network of 4,000 companies and get vetted and accepted to Vetterly by going to vettery.com/sedaily. Thank you to Vetterly for being a new sponsor of Software Engineering Daily.

[INTERVIEW CONTINUED]

[0:40:42.9] JM: Not to take words out of your mouth, but the environment is massively inflated by people who are creating accounts, selling fake follows, fake likes. If they really wanted to shut that stuff down more aggressively, they certainly could take steps, all your words. I mean, my words, not yours. Are any of these dynamics changing? Any of the dynamics of ad tech changing? Is anything getting better? Anything significantly different?

[0:41:10.8] SD: No. Things are getting worse, primarily because there are other formats coming around that limit measurability. You look at OTT and people are crazy about OTT. We have to –

[0:41:22.0] JM: What is OTT?

[0:41:23.5] SD: OTT is over the top. Connected TV, Roku, Chromecast, yeah, all these connected TV devices. You now have programmatic sense in them and they can load the same video ads that you traditionally got in your desktop, but now everyone watches Hulu on Roku or Chromecast or something like – or Apple TV. We're going to load ads there. The problem is the measurable quality has just got down. There's very limited measurement you can do. You don't have access to the same APIs and SDKs that you did on the web, so people just have to blindly trust.

Also, it's very, very expensive because they've basically taken a TV look at it is hey, this is a 1080p screen, you're going to pay a \$100 for this right now and it's juiced up the price.

Whenever you get a lack of measurability, just lack of access to data, just obscurity, like that's a recipe for fraud. If you look at the stat, Forbes actually had a good stat, was a quote from a guy who said within one year, the request volume went up by 20 times for OTT. Everyone has been going there.

What's also funny is since you compare it to TV, you would assume that there is just channels like ESPN, NBC and CBS and etc., and just that. There are random OTT channels, nothing like TV, where you have random channels – just like apps, there's millions of them. I'm pretty sure that people aren't – they're watching one channel or two channels. They're not watching multiple channels.

[0:42:49.5] JM: It's like the infomercials channel. I've seen your particularly weird channels.

[0:42:56.8] PS: Yeah, it's foreign channels, or –

[0:42:58.4] SD: As far as the advertiser knows, but here's the problem is as far as the advertiser knows once they've spent that money, it could have just been the blank screen channel. You have no way to verify what type of content was this before or after. They're buying it with the

expectations of how TV makes an impact on viewers, TV ads and not taking into account that this can be manipulated and counterfeited the same way that any open web page you can.

[0:43:29.2] JM: Again, just to get into the politics of an organization who is buying this traffic a little bit, the account manager who is buying this traffic still gets to say to their manager, “Hey, I bought OTT traffic. It's the cutting-edge thing. We got this number of views and we aired the toothpaste ad, or the restaurant ad.” The manager of that account manager is completely satisfied, because they get to go to their manager and say, “Hey, my underlings reported a 50% boost in views. Therefore, we'll probably see in six months that there's going to be a lot more toothpaste sold,” and you just have this chain of hierarchical middle managers that are all satisfied, because these numbers that are ultimately just useless are making them look good enough.

[0:44:19.2] SD: Yeah. You also have to consider that it's not just an accountant man. It's someone at the agency, right? Many big advertisers, many advertisers have agencies and it's the agency that has no affiliation with the advertiser whatsoever and has no – I mean, let's be honest, they don't care if you sell more Gillette or more cars. They just want to run their advertising and get the numbers.

You have them internally say okay, this is account manager for a certain big advertiser, like this is what we saw on this channel. Great. They'll go to their manager, their manager will give a thumbs up. All the way to the top they'll be like, “All right, great. This account will be happy, because we got 50% views from this channel and it was good.”

[0:44:58.7] PS: Increased our reach by 50%.

[0:45:00.5] SD: Yeah. Then they'll go back to the marketing team and give them a slide deck and everything in the marketing team if they're suspicious enough will be like, “Well, can you actually show me the data?” Then that's where the agency will start scrambling and actually start getting the data. Or the marketing team will be like, “You know what? Check the box and then I'll go to my higher-ups.”

[0:45:16.8] JM: What's the most disturbing thing that each of you has learned about ad tech in the last year?

[0:45:21.1] SD: I think for me, it's actually not a tech thing. There's all these tech things that come up with the in-app advertising and all the broken measurement and the bugs and everything. It's not tech related. It's the relationships between an advertiser and who they work with and how much money can just disappear without any tech involved at all.

[0:45:38.6] JM: Like billing. Billing related stuff.

[0:45:41.6] PS: Yes. Verification of billed amount versus goods and services delivered. It is very disconnected.

[0:45:48.9] JM: Oh, man. It makes you wonder about hospital bills and stuff like that, I'm sure. Not to go down a separate rabbit hole. You guys unfortunately moved out to the big city and left me behind in small town San Francisco. You've struck out to the big city New York to build out Method Media Intelligence. How's that going?

[0:46:12.2] PS: Going well. We are settled into New York now. We got some office space. We have two hires that we're very excited about that are going to be starting in the next few weeks. We have a few advisors that we're going through the process of onboarding, that are going to be strategic advisors just in terms of evangelizing the philosophy that we have, as well as the product.

Yeah, I mean, we were traveling to New York from San Francisco almost once a month anyways, so this just gives us a lot more time to be closer to where most of the large marketing organizations have home-based of some sort.

[0:46:53.1] JM: What's the difference between the New York and the San Francisco ad tech ecosystems?

[0:47:00.1] PS: New York is most of ad tech, San Francisco is Google and Facebook and a few small –

[0:47:06.3] SD: Yeah, it's almost funny because you think of the duopoly size. There's more ad tech in San Francisco than there will ever be in New York. It just feels more advertising related here, just because the ad industry started in you Madison avenue and everything.

[0:47:20.5] PS: This is where the checks can come.

[0:47:21.8] SD: Yeah. This is where a lot of the agencies have big offices and just the media industry is very prolific here. It's funny to think just in San Francisco while we're there, we're just surrounded by tech, SaaS, everything. The two biggest tech firms are in the area.

[0:47:38.5] JM: How are you guys playing to scale Method Media Intelligence? Because when you started, when I was just hanging out with you guys, it was very much a – it felt like a consultative business, very good consultative business, but eventually you developed these technology solutions around running tests. You can detect whether ad impressions are viewed by a user who has a GPU. This is actually a technology solution that you were able to build and you have customers for it. What's the vision for scaling MMI?

[0:48:08.0] PS: The main metric we're looking at internally is how much money are we protecting for our clients? The way we're looking at growth and scale is how do we make that number as big as possible in terms of – this is related to how we price, because we didn't do volume-based pricing in terms of how many impressions and advertiser purchases. We do it based on the size of their budgets. We do it for very large advertisers on somewhere around 1% to 3% of their budget. That's knowing that the waste impact that we see across the board in these exchanges across the web is 25%.

What we want to do is I think this year's estimate is somewhere close to 60 or 65 billion to be spent just on programmatic advertising. We want to monitor as much digital media spend as possible. That starts with going after the largest advertisers in the space.

[0:49:07.8] JM: All right, final question. We've talked about the duopoly of Google and Facebook. Amazon is getting into the ad stack quite heavily. How would you characterize Amazon in contrast to Facebook and Google?

[0:49:22.2] SD: Since it's a tertiary business for them and it's growing rapidly, they have a lot more leverage that's coming up. They're building ad products within AWS that facilitate this. Think of the – there's a product called AWS media tailor that allows server side stitching of ads, that's an example. Yeah, they're taking it very seriously. I mean, Amazon retail and AWS are the big money makers, so they can be very exploratory. That gives them considerable leverage on advertisers, because they can just say hey, this is a third going at it for us. You'll play by our rules. Unlike Google and Facebook, where they do have to play the sign dance.

[0:50:01.6] JM: Okay, guys. Well, always a pleasure to talk to you. I'll let you go. Thanks for coming on Software Engineering Daily once again and for doing interesting work.

[0:50:09.9] SD: Yeah. Thanks Jeff.

[0:50:11.0] PS: Thanks Jeff.

[END OF INTERVIEW]

[0:50:15.0] JM: This podcast is brought to you by wix.com. Build your website quickly with Wix. Wix code unites design features with advanced code capabilities, so you can data-driven websites and professional web apps very quickly.

You can store and manage unlimited data. You can create hundreds of dynamic pages, you can add repeating layouts, make custom forms, call external APIs and take full control of your site's functionality using Wix code APIs and your own JavaScript. You don't need HTML or CSS.

With Wix code's built-in database and IDE, you've got one-click deployment that instantly updates all the content on your site. Everything is SEO-friendly. What about security and hosting and maintenance? Wix has you covered, so you can spend more time focusing on yourself and your clients.

If you're not a developer, it's not a problem. There is plenty that you can do without writing a line of code, although of course, if you are a developer then you can do much more. You can explore all the resources on the Wix code site to learn more about web development wherever

you are in your developer career. You can discover video tutorials, articles, code snippets, API references and a lively forum where you can get advanced tips from Wix code experts.

Check it out for yourself at wix.com/sed. That's wix.com/sed. You can get 10% off your premium plan while developing a website quickly for the web. To get that 10% off the premium plan and support Software Engineering Daily, go to wix.com/sed and see what you could do with Wix code today.

[END]