EPISODE 359

[INTRODUCTION]

[0:00:00.5] JM: Search engines are run our lives. The path we take to information is dictated by Google, Facebook, Amazon, and other forms of search. Search engines feel objective and truthful, but they're built through ongoing experimentation and subjective decision-making. That's what has kept Danny Sullivan writing about search engines for 20 years.

The content Google prioritizes, the ads that we see, the way that a product review changes how highly a search result appears on a screen. These are the topics that Danny studies. He's the founder of Search Engine Land, an invaluable resource for news and updates about search engines and marketing. I've been reading Search Engine Land since college, so it was a treat to sit down for a conversation with Danny.

Software Engineering Daily is looking for sponsors for Q3. If your company has a product or a service or if you're hiring, Software Engineering Daily reaches 23,000 engineers listening daily. Send me an email, jeff@softwareengineeringdaily.com if you're interested.

Thanks for listening to this show.

[SPONSOR MESSAGE]

[0:01:25.5] JM: Blockchains are a fundamental breakthrough in computer science, and ConsenSys Academy is the place to learn about block chains. The ConsenSys Academy developer program is a free, highly selective, and carefully designed 10-week online curriculum where you will immerse yourself in blockchain development and earn a ConsenSys blockchain certification.

By completing the program, you will be eligible for immediate hire by ConsenSys; a leader in the blockchain space focused on the Ethereum platform. If you want to learn about blockchains and become a developer for blockchain technology, check out the ConsenSys Academy by going to softwareengineeringdaily.com/blockchain. The graduation ceremony is a spectacular all-

expenses-paid trip to Dubai where you will meet block chain developers working on real-world solutions.

Dubai has announced ambitious plans to become the first city to run on blockchains by 2020. If you like the idea of immersing yourself in blockchain technology, graduating, and going to Dubai, check out softwareengineeringdaily.com/blockchain. Build your new career on the block chain with the ConsenSys Academy developer program. Applications are open from now until July 1st, so you want to apply soon. For more details and to apply now, go to softwareengineeringdaily.com/blockchain.

To learn more about ConsenSys and Ethereum, please visit Consensys.net, and sign up for the ConsenSys weekly newsletter for everything you need to know about the blockchain space. Thanks to ConsenSys for being a sponsor of Software Engineering Daily.

[INTERVIEW]

[0:03:26.0] JM: Danny Sullivan is the founder of Search Engine Land. Danny, welcome to Software Engineering Daily.

[0:03:30.8] DS: Thanks. Thanks for having me.

[0:03:32.4] JM: I've been reading your work since I was in college, and I think that was like 7 years ago. One thing that you have captured in your writing, even back then, is how engineering and marketing have become intertwined and this started with email marketing, I think, but it got much more interesting with search engine optimization. What did search engine optimization? What did that term mean in the earliest days?

[0:04:00.8] DS: In the very earliest days, we didn't even have it. It emerged around, I think, '97 or so, and the idea was really that the search engines were going to interact with websites and look at the information on them, and depending on what they could gather, would have an impact on whether or not your listings would rank well in the top results.

At the time, there were things like people publishing websites that were all made of images and the search engines couldn't understand anything that was in the images, and so part of search engine optimization, or SEO, was understanding you need to use real words or people would build websites out of frames and the search engines couldn't crawl all the frames properly. Again, you had these architectural limitations as to how well your content would perform, because in some cases it is just literally wasn't visible.

[0:04:51.6] JM: In searching an optimization, it built into a term that meant trying to outsmart the search engine, but it sounds like in the earliest days, like if you have a site that's all images, that's not necessarily something that is shouldn't be indexed and yet there's not an obvious way to index it. That's something where you might actually — It's beneficial to the end user to add some search engine optimization to it.

[0:05:17.2] DS: Absolutely. In fact, it depends on who you talk to and where they were and what they were doing, but a lot of people didn't view search engine optimization even from the start. It's trying to outsmart the search engines, but rather understanding what did you need to do to build content to be friendly to the search engines to ensure that they could access your content.

It was akin to — I remember when I was writing about this, when we had many browsers and people used to say, "Is your browser going to work with NCSA Mosaic. Will your website work with that? Are you going to work with Netscape? Are you able to work with Internet Explorer?" You'd have developers who would test all those sorts of things. One of the things I used to talk about with the idea of search engines being the biggest browser of all and that you had to ensure that this stuff could work with them as well.

Another example too was the idea that search engine optimization is very much like a PR effort. When people undertake PR, typically, it's not that they're trying to somehow outsmart a newspaper, but it's the idea that if you understand how journalists operate, how they gather information, what's helpful to them, what's not help to them, you increase the odds that your coverage will be better. That sort of a different thing then, say, trying to outsmart or game a search engine where typically people might do things that are not within the search engine's guidelines themselves. They go into what's called black-hack tactics and then it really is, "I just want to be number one no matter what."

[0:06:49.4] JM: There was this period that the spam versus anti-spam began to emerge. I guess — I don't know if that was '97 or '98 or if it was later than that, but that's when people started doing things that were on the malicious side, where just to promote your page you would embed lots of transparent text or white text against the white background. What was that battle like to watch in the early days as SEO was decreasing the relevance of some Google results?

[0:07:26.1] DS: It was interesting. We had it just I think as long as people were dealing with the search engines, say, '95 or so you. One good example was at the time when Infoseek was a popular search engine, and you had the ability to have it reindex your page just as soon as you wanted to submit a URL and push submit, it would do this instant indexing. People would go through and they would submit their page and they would then see if they rank better and if they didn't they would immediately make a bunch of changes and just kind of keep going out in that way.

You had other people that started doing what was quickly known as doorway pages which were that they would create content that that was meant to rank just for the search engines and really didn't have an intent on trying to show up for human beings at all. They would add to that the concept of cloaking, which was that you would create a page just for the search engine, only the search engine would see it and if a human being then tried to go to that same page, they would see something completely different.

You had a lot of cat and mouse stuff that was going on. You had a sense of it sort of order being the Wild West. You had one case really early on that even went up to the FTC where people were taking content from other publishers and putting it on their own servers and cloaking it and trying to get top rankings. It really was like there was no sheriff in town to some regards, and people you would get frustrated that they couldn't get the best results.

Although I would say that the bigger issue wasn't really just that the SEO stuff or the black-hack tactics were polluting the results as much as you just had more and more information coming on to the web and it was becoming increasingly difficult for the search engines to find the needle in the haystack, if you will, just from all the noise that was out there whether it was intentional or just a consequence of the growth of the web.

[0:09:15.0] JM: What happened since then? What are the main trends in how SEO has evolved?

[0:09:20.3] DS: The big shift really was when Google came out and they started making use of the link graph and link analysis and looking at how people were linking from one page to the other and trying to determine the popularity and the relevance of these pages, so that they were able to kind of get that signal out from the noise because they were bringing in the better step up to the top.

Then falling from that, you had a huge growth and people just trying to buy links or game links and do that sort of stuff, and so even more tactics have evolved since then where they try to go after the artificial link networks and such.

The link usage was certainly one of the biggest change that we saw happen. Other changes that we saw influence stuff were the growth of the vertical searches so that if you're looking for news content, you'd have a news search engine, and then news content might be better for search that seem to be news related, and you had image search and you had shopping search.

Around, I think it was 2009, we had a shift towards personalized search results so that the stuff that seem more relevant to you based on your browsing habits and your past search history started giving you a more personalized result and it made it more difficult for people to sort of game because instead of gaming the search engine and getting a top ranking for everybody, you might not be getting it for everybody because so many people are getting personalized results.

Then we've had changes to increasing efforts to fight spam to the degree of things that some people may have heard of like the panda update that was designed try to figure out what was low-quality content and keep that from ranking better. More lately, just the growth in the mobile web and trying to better reward factor loading a more mobile friendly pages. Every year, every few months even, you get these changes that come along that keep kind of raising the bar and that are things that are important for both marketers and even developers to understand.

[0:11:19.6] JM: When I think about mobile, I think about taking out my phone and sometimes navigate, I go to Facebook or I open up the Facebook app and then I go to this some new site that I get link to on Facebook and then the ads are just like loading like crazy and the ads are such an affront to my experience and it just make me think, "Man! If Google were responsible for this piece of quality, it would never make to the top of the search rankings."

The fact that a lot of attention has shifted to this newsfeed model of consumption rather than a search-based or I email you link sort of model, it seems like it's really changed the content consumption of the internet.

[0:12:03.7] DS: Certainly. You got two things that kind of happen there. You had previously, to say, Facebook coming along, especially the idea that the search engines, and Google in particular, were somehow the gatekeeper in the way that people would get all of their content. Then you had this new growth in consumption where social media came alongside it. It wasn't that people were consuming lesser the search engines, it's just that they were spending even more their time now also consuming for social media and the social media having their own challenges.

It's been interesting watching people at Facebook having to constantly update their algorithm to try to fight low-quality content in the same way that we watch the search engines , and Google in particular, do that over the years. Even as you've seen this rise in ads, you've seen Google itself try to do things, they've had any number of efforts where they've said, "If you have ad heavy pages, you might not do as well. If you have experiences where you're interrupting the user too much and your pop-ups might be too interfering, especially on mobile, that you might not perform as well in the search results." It's this continue evolution of trying to get people to good content.

[0:13:15.5] JM: Facebook has algorithms that promote certain posts. Google has algorithms that promote certain search results. I know you're on the outside looking in, so you probably can't say with super significant authority. How do these compare? How do the algorithms of promotion compare?

[0:13:34.8] DS: It's much different, because with Facebook, their algorithms are heavily tied into the personal experience and showing you stuff that seems to be related to what you're already viewing or what your friends are liking or what may be popular among your network. Whereas Google, while it has the personalization, I think, is far less prone to do that kind of filtering.

You also will have Facebook, I think, try to look at what this community of known users is doing in the amount of time their spending and so on. Google can do that to some degree, but the connections are not as strong and they don't know see everything, let's say, that's going out on the web to whereas Facebook will see everything that's happening in terms of a click from its own network. Google can do that based on how people are clicking on their own search results, so do have some degree of that as well, but I think that Facebook just might have better insight in that regard.

[0:14:35.7] JM: Does Google also get more comprehensive insight because they are the browser that is the most popular at this point?

[0:14:43.4] DS: It gets debatable. They get data from how people are browsing. I have even a long track at this point to the degree that they say that's used or not used within their search results. I think their last dance was, "No, we're not. We're not using it." People don't necessarily always believe that, and maybe there was an announcement along the way where I've missed that sort of stuff. They certainly can see a lot of the browsing that's going on. They have the potential to understand how long people are dwelling on certain websites and what they're doing there. Certainly, on a personal basis, they'll use that information. We know that they'll use that as part of your personalize search results, the websites that you're going to.

In aggregate, potentially, they're using it as well and it does give them a lot of insight in that regard. As well as the idea that they have many sites out there that run things like Google Analytics, which means that they see a lot of the tracking that's going there. They typically have said that they don't use stuff like Google analytics, and that kind of makes sense because you can't rely on something that not everybody is using. If they've some sites, they don't use it and they don't have insight into it, then it makes it difficult to decide you want a reward or not reward based on that.

[0:15:58.5] JM: Dice.com will help you accelerate your tech career. Whether you're actively looking for a job or need insights to grow in your current role, Dice has the resources that you need. Dice's mobile app is the fastest and easiest way to get ahead. Search thousands of jobs from top companies. Discover your market value based on your unique skill set. Uncover new opportunities with Dice's new career-pathing tool, which can give you insights about the best types of roles to transition to.

Dice will even suggest the new skills that you'll need to make the move. Manage your tech career and download the Dice Careers App on Android or iOS today. To check out the Dice website and support Software Engineering Daily, go to dice.com/sedaily. You can find information about the Dice Careers App on dice.com/sedaily and you'll support Software Engineering Daily.

Thanks to Dice for being a loyal sponsor of Software Engineering Daily. If you want to find out more about Dice Careers, go to dice.com/sedaily.

[INTERVIEW CONTINUED]

[0:17:19.5] JM: It was interesting because, personally, I feel more tracked by Facebook even though I know that that's not true, because every time I go to Facebook — I don't know. It's just that I guess the way the perception I have for how I consume the Facebook product convinces me that I am being tracked more by Facebook and Google as more of this ambient all-seeing eye that is like, "Oh! The internet is tracking me," but actually Google is probably the biggest broker of the tracking there.

[0:17:50.5] DS: Yeah, I would say that, certainly, Google has greater insight across the overall web than Facebook. Although Facebook has a lot because when you consider how many people will have a Facebook widget on their pages, and that's data that all goes back to Facebook if someone's logged in. Ultimately, you have to have a Facebook account, and while so many people do have a Facebook account, Google is sort of agnostic. They don't care if you have a Facebook account or not. They're still seeing the data from a wide range of resources.

[0:18:20.7] JM: There was an article today on Search Engine Land, the site that you run. It was about reviews, and these could be Yelp reviews, they could be Amazon reviews. We're going to be talking about A-number of reviews sites. How do reviews play into search rankings?

[0:18:35.5] JM: When it comes into the local results, Google has talked about the idea that they'll try to take into account whether sites are reviewed well and perhaps reward them in that regard to the degree that they can gather that data. Certainly, they have a better sense of it just for their own Google Maps product and when people do reviews in that way. It really seems to have more of an impact when you're talking about local search results.

In particular, it's something that they've wanted to do and had made an effort a few years ago when you had someone who had a — That had like an eyeglasses site that got focused on in the New York Times because the person was convinced that any kind of review, whether or not it was negative or positive, was a citation that would help them rank better. The whole story was about how they were happy to be horrible to customers. They didn't care. They would still always win.

Google went through and said, "Well, it's not quite how it works that way, but we're still going to take a stronger emphasis on ensuring that if we are showing sites, that they are sites and local that are better reviewed."

[0:19:38.7] JM: There's a lot of critical press around Yelp. Some people accuse Yelp of essentially strong-arming businesses into paying for Yelp's services, or else they get de-indexed from Yelp. Whether or not this is true, I kind of see it as like this is what would happen if Google was evil or if they — Maybe if they had the internal mantra of do evil or intentionally do evil. This is the kind stuff that they would do.

What do you think of those accusations around Yelp? Do you think there's any — Is there any truth to that?

[0:20:18.9] DS: They have been concerning an and certainly there have been some cases where you're like, "Isn't it funny somebody did poorly on Yelp and then suddenly they got a call

from a Yelp salesperson." Yelp has been steadfast in saying it doesn't work that way and that they're not trying to target people and that it's not that you're just going to buy your way up to the top. I think it's one of the things you still want to keep an eye on and that it still needs to be watched.

Yeah, when you hear things like, "I got a call from a salesperson as soon as I got a bad review." It's like, "Wow! Okay. What's going on there?"

[0:20:52.3] JM: There's a lot of cases where people have said that happens. Yeah, it is all disconcerting. It I could be one of those things that's like driven by internal incentives that are pretty opaque to the outside where it's maybe one of these things like the Wells Fargo pressure to do cross-selling where maybe there was necessarily an explicit directive from upper management, but it's just one of these things were like, "Oh, the way that the incentives are set up, it just permeates the company."

It could even be where the incentives were not initially set up to do this kind of thing that it's a malicious outgrowth, and because they haven't anything to quash it, it's just a facet of the company.

[0:21:41.8] DS: Yeah. I think the hardest thing I think for a business is that you — Especially, I think, a few years ago, you have these local businesses that were just local businesses and being very busy doing their business. They're not internet savvy. They're not — Or they weren't up on the idea that necessarily they should even be on Yelp or that they already were on Yelp and they didn't know it, or how they should be responding to customer inquiries. Much less thing on things like Google Maps. They were local and dealing with people because they walked in through the door and they dealt with them that way.

You kind of feel for them when suddenly they've had this whole internet kind of injected and they've had to figure out how we make time, how we come up with the expertise, how we deal with this sort of stuff. I think these days though, if you're a local business, you understand something like yelp is a fact of life and you really can't just sort of ignored in the way that you could've in the past. You had to have a strategy about it. You have to understand that if there

are people complaining, that you're interacting with them in that way. You're never going to please everybody. I think the more that you can be proactive, the better it's going to be for you.

[0:22:51.6] JM: Yelp is a kind of product where there is not necessarily going to be as much consumer scrutiny. There may be some consumer scrutiny, but when I go on Yelp and I find a restaurant, I usually find the restaurant, I go and eat there and I'm like, "Okay, that was good, decent restaurant."

In contrast, Amazon — Amazon, it has a bigger ownership of the end-to-end experience. If people are returning products on Amazon, Amazon has a better view into the product quality. I think maybe that's led to Amazon having a better reputation among customers. Do you consider Amazon a search company?

[0:23:38.3] DS: Yeah, to some degree. Certainly in terms of shopping search. If you are looking to buy a product, Amazon is a shopping search engine, a first choice for so many people at this point. I certainly know it is for me that if I'm purchasing something, the first place I'm going to go to will be Amazon rather than, say, a few years ago where I might have gone to Google and I did a Google shopping search.

I'll look through Amazon and see the prices that I'm getting there, and if I feel they're reasonable or if they're good, and if Amazon fails me, then I may turn and go back to Google and see if there's a product elsewhere, see if the price is better in some other place. That's a real flip-flop.

Again, it's always dangerous when you talk about your personal experiences versus, say, what people are doing by and large. I don't think that's probably too far off the mark that you have lots of people who turn to Amazon when they need to buy things, and that means that they're conducting a search to see what's out there that they want to find.

[0:24:31.4] JM: I've heard there are these people who are just voracious reviewers on Yelp. There are voracious product reviewers on Amazon. I've certainly seen some of these people where you see these beautifully written reviews and it's like a work of art, and these people are certainly legitimate, but it's of course hard to know who out of that group is a legitimate user and who is maybe a paid user. What do you think about that paid reviews business? Do you have

any perspective for what percentage of reviews on the internet are legitimate and how that plays into what gets promoted in terms of products or restaurants?

[0:25:09.3] DS: I don't have any good statistics on that, so I couldn't say with any great deal of authority. Certainly, you can see, say, on Amazon and anybody who's purchased something from them, may have been through this experience where you get the product and they're like, "Please leave us a review and we'll perhaps give you 10% off. "What's more noticeable is when you go in and you look at some of these things that are highly rated on Amazon and then you look at the reviews and they're all like, "I was given this product in return for my free and unbiased review." The more I see things like that, the less likely I am to trust that the product is actually that good because they've clearly made an effort and perhaps too much of an effort to go out there and do that.

On Yelp, I don't use that as much as Amazon, but I personally tend to trust the reviews that look like they were from just human beings that went there and weren't trying to be the greatest Yelp person in the world, but just saying I like this and on.

The difficulty, I think, with all of these cases is that the people you may trust the most may not have the incentive to go there. I depend a lot on Amazon reviews and then I'm fairly rare that I'll actually go and leave a review. When I look at Yelp stuff, I'll depend on that and then I'll forget or just not make the time to go back and put my own review of a place that I like. It's kind of unfortunate, because I think you do have a lot of businesses that deserve to get that kind of credit when they've had good experiences, but for, if you will, the real people, the nonprofessional people, etc., their incentive to leave reviews probably tends to happen when they have bad experiences, because now you're like, "I'm going to do something," as supposed to, "Yeah, I just had a great meal. That was a wonderful restaurant," and I'm done.

[0:26:55.4] JM: Yeah. Back to Google, you have much insight into how decisions around search are made internally at Google?

[0:27:06.4] DS: They have a bunch of meetings among the different search teams. They've even had some that they've put out on lines of people themselves that actually watch one of these search quality meetings that they've had where they've talked about some of the changes

that they're considering doing, "Should we make this kind of shift? We make that kind of shift. What would happen if we do or if we don't do?"

I think you have a variety of people that are involved, and it kind of goes up the line as to whether something is troublesome or not in terms of they need to make a change and then trying to evaluate what kind of change they should make. They certainly are constantly experimenting. I can't remember if they've said there, always running like 100 or 1,000 experiments at any time, but they're certainly not sitting back and just saying, "Well, we're done.." They do all sorts of things. I think especially the more that they feel like something has a public focus, the more that they may fixate time on trying to solve those sorts of issues.

[0:28:02.4] JM: You had a recent post about how Google is penalizing websites who distribute content with link schemes. Explain what a link scheme is.

[0:28:13.6] DS: Google considers a link scheme anything where you're trying to get links back to your website in a way that they would consider unnatural or unearned. Links, in a lot of ways to Google, are like votes in an election. It's not exactly the same because just because you get a lot of links isn't the same as you get a lot of votes. Google tries to understand the importance of those links, and so they may give greater weight to say a link that came to you from the New York Times than they might give to a link that comes from a site that has relatively little authority.

In addition, they're looking at things like the context of a link. Amazon could have all the links in the world at once, but if all those links are talking about, "This a great place to buy books," then it does better for searches related to books rather than to searches related to cars, because you're not buying cars. What's in the link or around the link or near the link, the contest link has a role.

When it comes to these linking schemes, people understand links are important, so they go out and they try to obtain them. There are things you can do that Google would consider to be acceptable and then there are things they feel like are more extreme. In this example, I was writing about you would have people who would write content for new sites, and the content that they write they would give to the new sites for free and it would contain many links back to the originating site to the to the person who was distributing it. The links oftentimes would be very

what are called anchor-rich, or text-rich. They have lot of words in the links, lot of words that perhaps the site is hoping that it will rank better for. The hope is that if these places publish this content, then they've obtained all these links from people across the web and that they may do better for those words.

Google feels like taken to some extremes, that that's unnatural. That you didn't really earn those links because someone editorially reviewed this content and decided it made a lot of sense to really reward you for it as much as you just kind of gamed the system, almost kind of like ballot stuffing.

[0:30:17.1] JM: I want to actually talk about search advertising. We've been reporting heavily on online advertising and ad fraud and trying to get a really deep picture for the different players in the advertising space. I want to talk about these. I want to talk about online advertising, and then I want to talk about ad fraud. I believe 70% of Google's ad revenue comes search advertising. I think the breakdown is like 70% and then like roughly 30% AdSense, something like that. Explain how search advertising works.

[0:30:54.7] DS: It's pretty straightforward. When you're talking about search advertising, someone is placing an ad so that when a search conducted for certain words, their ad will appear. A basic example you do a search for, say, cars and you've got someone like Toyota who wants to show up and so they buy ads that are links, so if someone searches for cars, their ads shows up at the top and it's like, "Hey, how about buying one of our Toyota cars?"

When someone clicks on that ad, Google gets a per click fee, and that's the basics of it. The more complicated aspects are that you pay more depending on the popularity of the competition around a term, but you also may pay less and still do better over other people who want that term if your account is shown as having a good history so that even if there's somebody out there who says, "I'll pay $100 per click," and then you got another person who wants to pay $50 per click. If that person who's willing to pay $50 per click has an ad that actually draws more clicks overall, then Google may reward them because they're going to make more money off of that in the long term.

Then the mechanisms of buying against words are greatly evolved now where you can actually just buy against certain kind of concepts, or even as Google announced last week, you can have your car ads show up only for people that Google's determined are actually in market to buy a car. They determine these people who've done a whole bunch of searches related to car purchasing, and now they'll show you're ad only to those people as supposed to maybe people who are searching about cars generally including things like, say, Cars the movie.

[0:32:30.8] JM: The customers of the search advertising, are the customers the brands themselves, like Toyota? Is Toyota directly purchasing the ads, or does Toyota go to an ad agency, and the agency goes and makes the purchase?

[0:32:46.8] DS: It can be both or either. You have brands that buy things directly and companies will buy things directly and manage their own accounts. That's very common. You also have companies that outsource to what are called SEM firms, search engine marketing firms, that will go through and run the campaigns on their behalf. You even have some companies that do both, that they may allocate some kinds of efforts to an agency to run and then they may decide some of the stuff they want to continue to do in-house.

[0:33:18.0] JM: When there's outsourcing to agencies, does it misalign the incentives? Because I've heard stories where a big brand goes to an agency, they pay for a bunch of search advertising and then the agency — May be the campaign doesn't do so well, so they purchase bot traffic to go and search that term and then click on the ads so that the agency performance looks better. Does that happen?

[0:33:47.4] DS: I mean, certainly, those kinds of things can happen, but they don't tend to be, I think, long-lasting, because one of the things about search advertising is it's incredibly accountable. Unlike, say, buying a brand ad that shows on TV and you could say, "Look! A million people look at your ad." A search agency is going to be accountable on whether or not those ads have actually performed in terms of sales.

Someone doesn't just need to see the ad nor do they just need to click on the add and then go to the landing page. They actually need to be converting by some agreed metrics. If those conversions aren't happening, then they're still going to be accountable. Now, the conversion

could be that someone have gone to a landing page and filled out an interest form. Potentially, you take it out as far as they've actually purchased a car, but they're really a good agency and certainly good company shouldn't be handing over money for a very long period of time if they're not actually seeing real world results rather than just a bunch of clicks.

[SPONSOR MESSAGE]

[0:35:01.5] JM: For years, when I started building a new app, I would use MongoDB. Now, I use MongoDB Atlas. MongoDB Atlas is the easiest way to use MongoDB in the cloud. It's never been easier to hit the ground running. MongoDB Atlas is the only database as a service from the engineers who built MongoDB. The dashboard is simple and intuitive, but it provides all the functionality that you need. The customer service is staffed by people who can respond to your technical questions about Mongo.

With continuous back-up, VPC peering, monitoring, and security features, MongoDB Atlas gives you everything you need from MongoDB in an easy-to-use service. You could forget about needing to patch your Mongo instances and keep it up-to-date, because Atlas automatically updates its version. Check you mongodb.com/sedaily to get started with MongoDB Atlas and get $10 credit for free. Even if you're already running MongoDB in the cloud, Atlas makes migrating your deployment from another cloud service provider trivial with its live import feature.

Get started with a free three-node replica set, no credit card is required. As an inclusive offer for Software Engineering Daily listeners, use code "sedaily" for $10 credit when you're ready to scale up. Go to mongodb.com/sedaily to check it out. Thanks to MongoDB for being a repeat sponsor of Software Engineering Daily. It means a whole lot to us.

[INTERVIEW CONTINUED]

[0:37:02.9] JM: This takes us to the complexities of attribution, because let's say I'm a car company and I air ads on television, I am paying for search advertising, I'm paying for all these different things, and then the customer converts. It could be hard to disambiguate what the channel is, what the channel is that actually caused the attribution. I mean I guess this takes us

to that classic saying that, "Oh, half my advertising budget is wasted. I just don't know which half." Is attribution an unsolvable problem?

[0:37:43.1] DS: Last week, Google kind of announced that it was solved at least according to them. Their Google attribution thing that they've rolled out was designed to say, "Look, we're going to help you really understand what was the conversion or what were the steps along the way including trying to measure things like in-store sales and whether or not that person had exposure to your ads online or had exposure to your search ads and what roles those kind of played.

How well that actually plays out, we'll see. I was in the room when this was announced to a bunch of Google advertisers and there was little cheering, and it wasn't just cheering because they had a bunch of Googlers out there told to cheer. It was like, "Yes, these advertisers would love to have better attribution."

It becomes even more difficult when you start talking about online advertising or off-line advertising, because, as you say, how do you know whether or not it was the ad on TV that caused someone to go in as supposed to the ad that you ran on Google?

There are things that can be done that have been done. You can, for example, segment out the kind of stuff that you're doing. If you're running an online — If you're running a television campaign, you should be immediately able to tell whether or not your traffic is spiking because people are doing Google searches. That is very typical that you have ad campaigns that are run and then people go through and they do some searches. You can then try to look at what happens to that period of time to understand whether or not it was the brand awareness that generated that and if you should do more, as supposed to just the ads that you might be running anytime when you're not running a big huge off-line campaign as well.

That's something that's savvy advertisers would looking at and trying to understand. I think the reality is that it all feeds together. If you've ran an off-line campaign and you weren't prepared for the people who were then were going to do searches, then you're probably wasting some of the opportunity that you have. This is very common, say, during the Super Bowl where you have all these ads that might run and then you could go when you do searches and you'd find that the

advertiser didn't actually run any ads on Google itself. Sometimes you'd even have competing advertisers running against whatever terms that were being used to try to kind of capture the interest of people who might come in that way.

It really is trying to understand how all these stuff filters in together and account for, but also realizing there's probably not any single channel that's ultimately the, I don't want to say the blame, but that's ultimately generating the whole conversion.

[0:40:13.2] JM: In my reporting for advertising fraud stuff, most of the stuff that I've investigated has been display advertising and there's a lot of fraud in display advertising. Of course, it's impossible to audit exactly how much there is, but I've talked to a lot experts and the numbers, even the lower bound of their numbers is pretty alarming. I don't understand search advertising as well. How much fraud is there? Do you have any perspective on that?

[0:40:46.5] DS: I don't have the figured off the top of my head, but Google usually does an annual report. The numbers unsurprisingly are very small. I think they're like single digits. They usually report a lot about how much they've actually caught before someone build on it. It was a very big issue years ago and in fact there was a whole case that was filed against Google about fraudulent clicks. You would get people especially who might click on their competitors ads or do something like that that drive their competitor prices up.

Some people felt like that was going to kill Google. In fact, the lawsuit was settled and it wasn't settled for that amount of money. It really hasn't been that much of an issue. Even with the display ad stuff that's going out there. Google, I think, certainly itself comes under targets of are your display ads suffering abuse? But I think that most of it tends to be on other networks that are perhaps less robust about caching unnatural clicks and things that happening there.

Ultimately, it comes down to the advertiser looking at the performance. If you're running these campaigns and suddenly these campaigns that you've not done anything different to are getting more interested in traffic but less conversion, there's something wrong. As an advertiser, you certainly moves you to examine it and raise the questions with the network itself about what's going on and where are my ads showing and how do I narrow it down to places that are actually working?

[0:42:11.4] JM: Here is the thing that alarms me, is that it's —One of these things, it's almost like the mortgage crisis where you have a system with really weird incentives, and the way that the incentives happen to be structured is such that the system isn't necessarily held accountable. If you have CMO's who they have an 18-month tour of duty in their job, then maybe they're not held super accountable. You have these account managers who are buying ads with other people's money and maybe they're not necessarily using the best measurement tactics because they're maybe not necessarily interested in how advertising affects the company. They're interested in how advertising will make their job look.

An account manager who'd purchasing display ads might buy it on a sketchy site where they can — As long as the account manager gets the right metrics, even if those metrics aren't correctly aligned with the business interests of the company where they're purchasing the ads for you, that's not going to lead to smart purchasing decisions. Then when you have all these by side incentives messed up, that just leads to the sell side, the ad tech businesses offering products that are just catering to the misaligned incentives of the buy side.

[0:43:33.1] DS: One thing I'd say is, first, for all the faults that I think happens with digital marketing, the accountability of the digital marketing seem to be far far better than the accountability with online marketing. Digital marketing platforms and purchases typically don't involve commission to begin with, so you're just not getting a cut of whatever you spend. You're having negotiate to get paid on top of what the actual media purchase is going to be.

Then with off-line, for example, if you look at those television campaigns that will run, you may get a sense that will this many people look at your ad, which really isn't taking it into the account that that many people just fast-forward it through your ad and didn't really see it and you don't really know who's watching it and you have the best guess of what the demographics. The accountability for off-line stuff is almost laughable. Really, I think it's of kind of sad in some ways that digital gets so accountable for the waste or the problems that has because it's so accountable that you can actually figure out what it is as opposed off-line where you can't tell that as well.

That having said, no one wants blazed and you do want to know that the things you're spending money on are actually effective. I think that even if the CMO's are saying, "Well, were just spending a lot of money and look at how great I am." Ultimately, more and more is coming back performance.

We've had CMO's, I think, it was the FTD one. I could be off. I'm trying to think off the top ahead a few years ago, but I think one of the reasons they were dismissed was that their search campaigns were not performing as well for them to sell ads in the way that they were expecting. That there was a real bottom line, like, "Why did we not make more money?" "Oh, we didn't do it because this didn't work." "Okay, we're going to get someone else who understands that."

I can certainly see that accountability problem happening, but I do think ultimately what you're going to get is more the performance marketers who start to move up into those roles. That I think you're going to have a change where you've had CMO's who come out of brand building and off-line purchasing and the traditional ways of advertising get replaced by people who grew up doing search ads and doing display ads who live and breathe the idea that everything should be measurable. If you can't back it up and if I'm not seeing that this is actually increasing things, we're going to make some changes.

[0:45:56.1] JM: The things that concern me the most are what — I saw this —Somebody I know showed me this scheme recently, and he's going to publish about it eventually. I don't want to say who he is, but it's a scheme where this well-funded company, basically, all that they — They do "marketing", but what they actually do is they pay celebrities to post content on their Facebook page, and because the celebrity is posting it, they get real clicks, and then they use those real clicks to launder fake clicks from bot traffic and then they just get this combination of real clicks and fake clicks pouring into their scrappy news sites with just salacious titles and there's ads all over these pages. This is the kind of scheme that ends up bilking advertisers for lots of money.

Certainly, a lot of these advertisers are naïve and, certainly, these sites are probably things that are not on — Hopefully, they're not on AdSense, or maybe they are, I don't know. Hopefully, these are served by less lower quality exchanges and whatnot. I just use it as an example of the sophistication — The sophistication of these add fraudsters is basically at the level of Wall

Street. It's like the Wall Street level of sophistication with zero oversight, like zero penalty if you get caught. I just look at that incentive structure and I have no way to comprehensively audit it. Nobody does. But I just look at the incentives. I look at the financial opportunity and it's a little alarming to me.

[0:47:37.7] DS: I think there aren't ways you can audit some of these stuff. Ultimately, if you're an advertiser, you should see every click that's coming back into your website and you should be able to know where that click came from.

If someone's doing this sort of thing where they're orchestrating bot traffic mixed in with, if you will, real traffic to drive visitors, you should be able to detect that and you should be — The challenge is depending on the size of the advertiser, you might not have the resources to be going through and doing that kind of an audit. If you're spending a sizable amount of money, it's probably that you need to spend some money to make sure that the people you're going with are actually doing the job as being claimed, and that includes, by the way, doing things like checking the references of the companies that you're employing and talking with people and how that they've been doing.

You do also have some checks and balances and that if you are trying to run a sizable agency that's going to be something other than fly-by-night, you really don't want to wreck your reputation with any of the major networks, because then your account gets hit and your account can't perform as well or things become more expensive for you. Potentially, you're not allowed to buy the ads as well.

There's a strong incentive, I think, actually to be doing it right. If you're going through a second, third tier kind of ad network and you're just purchasing stuff there because it's looking cheaper and they're promising you lots of stuff, then yeah, you're probably going to be getting less quality types of things and you get paid less for it in the first place.

I guess it kind of comes back to the adage of you kind of getting what you paid for. I do think that there are some checks and balances out there that's not just a complete free-for-all. Certainly, I think if advertisers themselves are not trying to make those efforts to review what's happening, then it's more easy for them to get caught on that type of stuff.

[0:49:35.4] JM: You wrote last month about what you consider to be Google's greatest search quality crisis. Explain what is going on with Google search quality.

[0:49:46.5] DS: Well, in the fake news era now, you've had greater attention drawn on to the kinds of content that people are seeing and how it's getting promoted. What was happening with Google was that you had a lot of attention focused on Facebook and then people notice that, like on Google, you had a fake election result story that was showing up in the top of its results promising that these were the real presidential election results, and they weren't.

Suddenly it was like, now Google has a fake news problem of its own. Then that got magnified, especially by a series of articles that the Guardian ran where you could search for things like did the Holocaust happen? Google came up with results that were dominated by denial sites telling you that it didn't happen where people are saying that this is all fake, or you had one search example that was done where if you said, "Are women evil?" It would come back and tell you, "Yes, they are evil. All women are a little evil. All women have a little prostitute in them," which is horrible.

[0:50:48.1] JM: I saw that. I saw you post that video.

[0:50:51.5] DS: Yeah, it's horrible enough when you just get that as a result. The other thing that's happened, of courses, is that Google does what's called featured snippets, where it will take one answer that they think really answers what you're looking for. Elevate it above all the others in the special box and if you like a Google home device, it will read it to you, and that's, of course, what the video was that Google home reading this sort of thing.

It caused people to, I think, especially focus on all these things that you could get Google to say or suggest or you know just a basic question, like, "How could you have an answer like did the Holocaust happen, when it clearly did happen, and you don't have these results that are coming up there."

When you talk to Google, there are good explanations for why they have these bad things that would happen. I'm not saying they're good explanations like that's excusing them, but I'm saying that, for example, in the case of Holocaust denial, you don't have a lot of sites that are authoritative about the Holocaust writing content explaining why the Holocaust didn't happen, because it's inconceivable to them that people would actually think that, right? They're explaining this well-documented history of this horrible thing that happened and they don't think we should write content to explain to somebody. "Yeah, it really did happen." It's sort of like people don't think they need to write contradicts explaining why the earth is not flat, which is a whole another search you can do on Google and you'll find people telling it is flat.

What would happen is you do these kinds of searches and then Google would give you back answers that had thought were the right answers because they addressed what you searched for, but Google itself isn't actually a truth machine. It can't actually tell you whether something is a factor not. It's making guesses and it's making these terrible guesses.

I think it got to the point where you just had people, especially questioning what was wrong with Google and why was it letting all the these stuff happen and why was it supporting climate denial and anything that they didn't like would be part of this big, huge Google failure. Which is why I said it was their biggest search challenge, because Google has built its reputation on being the perfect search engine and it has all the right answers, and then here it was just being demonstrably proven that if you search for is Obama planning a coup, that yes, President Obama is planning a coup, and he wasn't.

[0:53:15.0] JM: Yeah, I saw that one too. Google Home doesn't know how to pronounce coup d'état.

[0:53:19.6] DS: Yeah.

[0:53:20.5] JM: It's was like coup d'état I think.

[0:53:23.1] DS: Yeah, it has issues. This was really going to the core of their primary product and putting it into question. It's stuff to they're dressing now and they're trying to do a better job with it, but it was really something — This goes back to when you asked earlier, like how do they

decide what things to address and what things are issues? They could see that these things happened, but they could see that these things happened in very small amounts. Like not that many people would search for things like did the Holocaust happen? It's a tiny number compared to, say, to people who would search for Holocaust, and Holocaust would have very good results.

They didn't really view it as a priority of, "We better fix this," when we're trying to deal with something like we have sites that are hit by malware and we don't want the malware sites showing up for popular searches because that affects a lot of people. Now, they're kind of realized, "Well, even if the number of these searches is small, the impact is much broader than that, so we actually really do need to do something.

[0:54:22.5] JM: Do you feel that this is an outgrowth of having a person in power? The most important person in the world is somebody who contradicts himself on a regular basis, and basically has stirred up the world into questioning, basically, the foundation of what we consider to be truthful statements. Basically questions —He questions whether we can even know facts. That's the implication of a lot of what Trump does. Do you think this is an outgrowth of that? Is that why we've had this fake news problem, this Google search problem, this Facebook — Facebook and Google having to turn to human curation because — I mean, is this something new or is it just that it was a confluence of circumstances?

[0:55:13.4] DS: I probably lean more towards the confluence. I think you've continued to have this growth. I think you had people putting out their own versions of what the truth is in their own alternative facts and you've had people within their own filter bubbles that have been growing up more and more and more and that it's especially, I think, all come around the election where it's really caused a great deal of awareness and concern among people.

Now, you have these two great gateways, if you will, or these two great communication channels; Google and Facebook, trying to grapple with, "How do we deal with it?" They're tough issues too because in some cases you can look at stuff and say, "These are facts and they're indisputable." You want to say that, but then sometimes you're like, "I thought that was the fact and yet people are still disputing it."

Then you can also get into things where you do have discussions and things that really are debatable and you'll have both sides that will want to scream or multiple sides that want to scream fake news and you just get even more noise. Trying to find how you get the right balance from allowing people to have free, open, and healthy discussions as supposed to people getting lost in these rat holes of fake news and false claims, it's a big challenge.

[0:56:41.8] JM: The thing is we can't — This goes back to Google's basic profound realization, which is that understanding the internet is about understanding references to each other and learning who you can trust from references, and that is a recursive property. You can't go to the moon and confirm for yourself that there is an American flag on the moon and that we actually went there. Ultimately, you have to trust some people as being authoritative sources. You cannot go and scientifically confirm everything for yourself.

[0:57:16.2] DS: Yeah, correct. I ultimately bring it back to — This is just too big of a long-term problem, but our education process in general and the ability for people to exercise good solid critical thinking skills.

Yeah, we all can't go to the moon. If we want to doubt whether someone went there or not, we have to employ our own methods of evaluating all the information that's out there. I think most reasonable people would look at all the information that's out there and say, "Yeah, I think we went to the moon." That it would be this vast conspiracy of things going on to try to convince ourselves that we didn't go there.

I think the thing that's most disheartening to me is that one of the top resources that we've traditionally depended upon, which is been media and newspapers especially, that they just get attacked and dismissed and assumed that everything they write is lies, and it's not. It's terrible and it's disheartening.

[0:58:18.7] JM: Yeah, okay. On that note, I know we're up against time. I just want to ask you one more question, because Software Engineering Daily is a business I'm trying to build. You've had a lot of success building media companies or around technical topics. I just want to ask, do you have any advice around growing a media business, particularly one that is directed at — Not a narrow, but it's like a particular niche.

[0:58:44.1] DS: From my perspective, certainly, picking your niche and being the best that you possibly can, and it is one of the best things that you can do. The world probably doesn't need yet another general news site, but i think people greatly appreciate vertical sites that go deep on topics and do it with care and really try to educate and respect their readers and give them good information. That that's certainly something we've always tried to do and it's worked for us. I hope it would work for other people.

I think the other really important thing is understanding how you're earning off of that. I did a post about two years ago, and this was after the news site Gigaom went under and people were like, "That's it. The vertical site — It's all over." I was like, "No. It's not over for vertical sites. There are plenty of vertical sites that are doing fine and that continue to grow. You just don't tend to hear about them because they didn't take a lot of venture capital money."

My advice there was that if you're going to build something, understand how you're going to make money off of it if you're viewing it as an actual business. I think you've got some people who just view it as, "This is an exercise for me to get an exit down the line," and so they don't think about how they're going to fund it in the long-term. With our own aspect, we've always grown within our means. We never took VC and we add people as we can afford to add people and we grow out in the ways that make sense for us to grow. We potentially could have taken a lot of VC money, but then we might've built an operation that actually wasn't able to sustain itself to the amount of debt that it took on. That's sort of my own take on that.

[1:00:30.5] JM: Yeah, that's really instructive. Okay, I'll take those words of wisdom to heart. Danny, thanks a lot for coming on Software Engineering Daily and thanks for producing a lot of content than I've read over the years.

[1:00:40.3] DS: You're very welcome. Thanks for having me.

[END OF INTERVIEW]

[1:00:44.5] 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 to Symphono for being a sponsor of Software Engineering Daily for almost a year now. Your continued support allows us to deliver this content to the listeners on a regular basis.

[END]