Indexing (JavaScript) comments, and much more!

**Martin Splitt:** [00:00:00] Welcome everyone to the next episode of Search Off the Record, a podcast that we’re trying out. Our plan is to talk a bit about what's happening at Google Search, how things work behind the scenes, and maybe have some fun along the way. My name is Martin Splitt. I'm a Developer Advocate at the Search Relations team.

And I'm joined here by John and Gary.

**Gary Ilyes:** [00:00:30] You don't see this, but we are waving.

**Martin Splitt:** [00:00:32] That's fantastic. Everyone waves into the microphone, which is a great idea. Sweet. So, Gary, how is your website coming along?

**Gary Ilyes:** [00:00:41] Why do I always start this?

**Martin Splitt:** [00:00:43] Oh, I can start if you want to, but um, well, do you want to start?

**Gary Ilyes:** [00:00:47] Well, actually I’m funnier, so probably I should start.

**Martin Splitt:** [00:00:50] Yeah. So there you go.

**Gary Ilyes:** [00:00:52] Website. Websites are hard. It seems. And yeah, I'm on track, I guess, but I have reached a point where I would rather pay someone to actually create the things that I want to create rather than developing myself, because I’m running a WordPress installation and it's great as usual, but I need some custom plugins and I'm actually developing that. So the idea is that Lizzie, our tech writer, and I test out things on the website. Once we actually published the docs about those features or structured data or whatever, and for that, I need some custom plugins. Like for example, I cannot just use a normal recipe plugin because there might be a lag between our public release and how fast plugin developers implement those new features. So I want to be able to add custom structure data to every page. The second thing is that I do want to have a dropdown or an accordion or something where I say a few words about why I added certain things to the page, for example, why I chose this title, or why did I link to Johnmu.com/hacking and so on. And those take time. And I was wondering if I should just create, let's say an HTML page from scratch for every single recipe that we want to publish. And then that would be relatively easy to add whatever we want to, but I thought that is perhaps a bad idea. The second idea was that I actually create a CMS from scratch, which is not a foreign concept for me, because I did create a CMS, I think right before I joined Google.

And that was actually quite great. And it was working pretty well. It was based on the Google Docs API, and basically you are writing your content in Google Docs. And then there was a container running on a standard web server, an Apache web server, and that was pulling in content from Google Docs. And back in the days, that was quite cool. And it worked quite fast, like serving time was under one second, but then it got more slow - as I imagine more users started to use Google Docs API - and at one point, I think the load time for the content itself, the container loaded perfectly and snappy - in a snappy way - but then the content appeared only like with a five second lag, I would say. And that was a no from me.
And the second problem was that the websites got fairly popular. And I got an email from
the, I think, API team at Google that maybe I should scale back or upgrade my plan. And at
that point, I think I already signed my contract with Google and I was like, 'hmm...maybe
this is a good time to stop.'

**Martin Splitt:** [00:04:19] Oh my goodness.

**Gary Ilyes:** [00:04:20] Yeah, but it was fun. So yeah, back to the website or my new
website, it's probably on track, but I like procrastinating. And this is a good time to
procrastinate as well. And I'm baking cookies for Martin

**John Mueller:** [00:04:34] Third-party cookies.

**Gary Ilyes:** [00:04:35] So that's more important than the website itself.

**Martin Splitt:** [00:04:38] Absolutely agree.

**Gary Ilyes:** [00:04:39] Yeah, the CMS that was right before I joined Google and it was one of
my favorite projects.

**Martin Splitt:** [00:04:45] It kind of makes me happy to see that other people also build their
own CMS. I think I built multiple CMSes myself because the first one sucked. And then
the second one was quite nice, but also relied on some third-party service that then got
discontinued or something.

**Gary Ilyes:** [00:05:01] Cause they do.

**Martin Splitt:** [00:05:02] Fun times.

**Gary Ilyes:** [00:05:03] How did you join, John? Or when and why?

**John Mueller:** [00:05:06] When did I join? Ah, that was a long time ago. I think. What was it?
Maybe 2007. So I don't know, 12, 13 years ago? I forgot how to count. Something like that.
I think at the time I was also making my one CMS. Maybe this is like a requirement. I don't
know, but basically, yeah, I had a small software company and we started doing more and
more things on the web, as you do.

And I started working on some of these websites because it was interesting, something
kind of new. And at some point I started helping out in the webmaster forums as they were just
started. So I think sitemaps just launched around that time. And I was active, kind of helping
people with the sitemap stuff and trying to give them my opinion on things they're all doing
wrong. And at some point, I received an email from someone at Google saying it would be
nice if I would drop by their office in Zurich, since I'm already in Switzerland, it was kind of
weird because the email came to a domain that I don't actively check for emails. So it was
almost accidental that I noticed.

But it was pretty neat. So I went to Zurich. Yeah. I met some of the people from the
Sitemaps team who were active in Zurich and it was pretty interesting. It was like a small
office, not a ton of people and that's kind of the first connection I had there. And then we
set up interviews and I learned all of the Webmaster Guidelines by heart because that's what I assumed I would be asked about.

I flew over to Mountain View, met a bunch of the people there and we had good conversations. It was really interesting. And then essentially I had to make the tough decision of, should I keep my own company or should I join this big corporation, which at the time was not as big as today, but it still felt quite a bit bigger than kind of running your own company.

And I decided, well, might as well try something new out for a change. So I ended up going over to Google and doing stuff there. I think at the time there were two other people on the Webmaster Trends Analyst team - Jonathan and Susan.

Gary Ilyes: [00:07:36] Oh... I miss them!

John Mueller: [00:07:37] Oh yeah, no, they were great. And I don't know, for a while we were pretty much kind of a small team working together with the Webmaster Tools team back then.

So that was like, It's my one chance to say 'Webmaster Tools' and be correct. I guess. And we started trying to expand and trying to find good matches for the team and for Google, which is kind of hard because it's a tricky role in that it's not like a software engineer where you can test on whether they can do coding properly, but rather there's some amount of communication, some amount of guessing and understanding involved. And I think you, Gary, you were also active in some of the forums at the time, especially at the News Publisher forums. And for whatever reason, I don't know we were in touch. So at some point we thought maybe we could hire this guy.

Gary Ilyes: [00:08:34] Yeah. That was a mistake.

John Mueller: [00:08:35] No, that was great. That was great. That was great.

Gary Ilyes: [00:08:38] Yeah. The way we met actually, was... you set up a chat, I think, and you were trying out things on that chat and some of the bionic posters, some of the contributors, joined the chat and then we were just chatting there.

John Mueller: [00:08:54] Oh, and one thing that I thought was very cool - sorry to interrupt, Gary - so sorry - was you set up this cool system on how Google could index JavaScript pages.

Gary Ilyes: [00:09:07] Oh!

John Mueller: [00:09:10] Remember?

Gary Ilyes: [00:09:11] I remember! I hated JavaScript back then as well. Surprise. And, and I set up a headless Firefox that was doing server side rendering. Oh, that was a long time ago. That sucked so much.

Martin Splitt: [00:09:26] I'm shocked. Shocked! I tell you.
John Mueller: [00:09:29] Yeah, it's so weird. It's like, and then we found someone else to also help us with JavaScript at some point...Martin!

Martin Splitt: [00:09:40] That was such a weird, weird thing. I got contacted by a recruiter at some point. And then Ilya reached out to me like, 'Hey, I saw that you're in touch with a recruiter on this position.' And I'm like, yeah, but I still don't fully understand the position because it was like web content ecosystem. And it was really vague.

And I'm like, [00:10:00] uh, sounds interesting, but I'm not exactly sure what's happening there. And then I asked Ilya and Ilya is like, 'Oh, you know, this is not so much about the web platform and JavaScript. This is more about Search and SEO and maybe news and maybe a little bit of Assistant.' And I'm like, Hmm. Okay. It sounds less interesting. And then I had a conversation with you, John, and I think we went for lunch because I was in Zurich at the time. And you all have also been in Zurich. And then I asked about the position and it's like, yeah, it's a lot about JavaScript and the web and Chrome. And I'm like, Oh, okay.

This is misleading. This is weird. And then I asked all of you what a typical day looks like. And I got the least satisfying response. Which was like, there is no such thing as a typical day, it can be whatever. And I'm like, wow, thanks. That's helpful. And then I was going through the interviews, nonetheless, and then once I got the offer, I was like, I'm still not sure what the hell I would be doing.

And I, I figured out that you all were like, yeah, basically super pros in how Google Search works internally. And I had no idea and I couldn't see myself being like super helpful in that regard. So I remember this moment that sealed the deal for me was when I asked you. I think either during our second lunch or, well, after our second bunch, like, okay. But why me? And then you're like, because you are a web developer and we need to figure out this JavaScript thing and I'm like, Oh, Oh yeah, that makes sense. Oh yeah. That's actually pretty cool. So, yeah, that's how I got my mission and how I signed the contract in the end. Probably also a big mistake for you all.

I'm sorry.

Gary Ilyes: [00:11:36] You are way too optimistic and way too happy all the time. It's kind of annoying.

Martin Splitt: [00:11:41] I'm kind of balancing someone else, I guess.

John Mueller: [00:11:44] It's a fantastic team. Don't worry. Yeah. It's like everyone fits in.

Gary Ilyes: [00:11:49] That's just like your opinion.

Martin Splitt: [00:11:50] I'm super happy to be part of this team. I like you all.

Gary Ilyes: [00:11:53] Of course, you are, Martin.
Martin Splitt: [00:11:55] Oh, that's wonderful. And it's quite cool that we get to help so many webmasters all around the [00:12:00] world. That's really interesting. At least to me, that's interesting.

John Mueller: [00:12:03] Yeah.

Martin Splitt: [00:12:03] Maybe I help you with the plugins, Gary.

John Mueller: [00:12:08] So how is the JavaScript stuff working out now? Do you think we have it all solved and it's like no more bugs?

Martin Splitt: [00:12:14] Oh my! Definitely not. The biggest thing I learned in the last two years at Google is that developers and SEOs are both extremely creative in building things that don’t work - in interesting ways. Like it's quite fascinating. Most of the things are fantastic ideas on paper, and then you implement them and then you find out, Oh, it doesn’t work the way that you would expect it. Sometimes that's because the creativity is misplaced. And sometimes it's because we’re just having glitches. Like recently, I think it was on Twitter. I think Gary made me aware of it. So someone was asking about Disqus comments, I think. And that's not something that is very specific to Disqus, by the way, it's basically websites using a third-party to pull in content that they care about. This can go wrong for many different reasons. In this case, it was a glitch on our side.

We found out what the glitch was, fixed it, and basically within the day we had it back working. But I think this kind of started a larger discussion as to how you should do things when it comes to JavaScript and critical content from third parties. Because the challenge is that you, as a website owner, don’t really have control over a third party.

And if you are using client side JavaScript to pull in content from the third party in the browser, things can go wrong. They could robot their JavaScript API, and then we can’t make the request or maybe their servers are really under load. And then we decide not to make these requests to the third party because they are already experiencing high load situations.

And there are usually ways of doing this on the server side. So if the third party exposes an API that you can interact with from the client's side, [00:14:00] from the browser with JavaScript, you can very likely also do that on the server side and then basically avoid these problems because then your server controls what happens when, in terms of when the data comes in from the third party.

But I think not as many people do that and I would hope that people are kind of warming up to the idea of doing that instead of doing everything in the client side.

John Mueller: [00:14:23] So is it a bad idea to rely on third parties or is it just, you have to be careful?

Martin Splitt: [00:14:28] It's an okay idea to rely on third parties. You just have to be careful and you have to understand that in the browser, you have very little control over what happens and how it happens. And if you are relying on Googlebot to do the heavy lifting and figure out how to get the data from the third party, then you are less in control than when
your server does that work because your server is an environment that you have control over, hopefully.

**John Mueller:** [00:14:52] That sounds a bit like Gary's problem with the plugins where if the plugins are doing one thing and you're trying to do something slightly different than kind of balancing the offloading of writing code to whoever's making the plugin versus being able to choose exactly what you want to do.

**Martin Splitt:** [00:15:12] Yeah. Yeah. And I guess for many people, it's fine. I think if you're pulling certain pieces of content, I dunno, like comments or something, and you don't really think that they are your main content. Then I think it's fine to just go with an off-the-shelf solution with a third party that may or may not work well in Googlebot and other browsers.

But if you want to have control over things like Gary wants to have control over the structure data on his pages, then I guess you want to run that on your server in a controlled environment.

**Gary Ilyes:** [00:15:43] So, Martin, I was taking notes while you were speaking, and I just wrote down one thing and that's 'Do not use JavaScript!'

**Martin Splitt:** [00:15:54] That's a...

**Gary Ilyes:** [00:15:54] Did I summarize it right?

**Martin Splitt:** [00:15:56] No, not really, but I mean, for your point of view, JavaScript is evil. [00:16:00] So I kind of see where that's coming from. But then I think you can use JavaScript. It's kind of fine, but you know, if you don't want to, then don't, by the way, JavaScript can run on the service side, just FYI.

**Gary Ilyes:** [00:16:11] Because it's such a great idea. Yes.

**Martin Splitt:** [00:16:13] It's a fantastic idea. You could hypothetically, so there's a thing that lets you run PHP in JavaScript. So you could do that.

**Gary Ilyes:** [00:16:22] Brilliant.

**Martin Splitt:** [00:16:22] Yeah. And then there's a thing that lets you run JavaScript and PHP. So you could like build a fantastic stack.

**John Mueller:** [00:16:28] Don't give him any better ideas. I mean...

**Gary Ilyes:** [00:16:30] I am rolling my eyes.

**John Mueller:** [00:16:32] You already started with a CMS that was built on top of Google Docs, which has built on JavaScript. So I don't know.

**Martin Splitt:** [00:16:39] It's a great thing.

**Gary Ilyes:** [00:16:40] No, no, no. I was actually doing with that what Martin was saying: pull the content on the server side.
Martin Splitt: [00:16:47] Good boy.

Gary Ilyes: [00:16:47] And that actually worked excellently.

Martin Splitt: [00:16:50] Good boy.

Gary Ilyes: [00:16:50] Until things started to slow down. But that was out of my control.

John Mueller: [00:16:54] That was probably before speed was a ranking factor. So you were probably okay.

Martin Splitt: [00:16:58] Probably.

Gary Ilyes: [00:16:59] Oh no. You said it. You said the bad...

Martin Splitt: [00:17:03] Bad word?

John Mueller: [00:17:04] Oh my God!

Martin Splitt: [00:17:04] Because of ranking? Is ranking a bad word now?

Gary Ilyes: [00:17:08] Well, this killed the discussion. alright.

John Mueller: [00:17:09] No, no, no.

Martin Splitt: [00:17:12] Well done!

John Mueller: [00:17:13] So one last thing, Gary, I guess, with regards to indexing the comments, because Disqus is basically comments from other people. Is that generally a good idea or a bad idea? Should sites kind of block the comments from being shown or should they get them indexed?

Gary Ilyes: [00:17:28] I mean, it depends on the comments, right? Like you can have comments that are like super useful and worth having them in the index and having them help you rank for certain things a little at least.

But then you go to certain blogs that are, I don't want to say. That will be R-rated then, you go to certain blogs and you read the comments and you just want to take a [00:18:00] spoon and...

Martin Splitt: [00:18:01] Just eat ice cream?

Gary Ilyes: [00:18:02] I don't know, just pop out your eyes or something, like, it's just really bad. In those cases, you probably don't want the comments.

Also, I think, I think we said this before, but depending on where the content is on the page, it might weigh less than for example, content that's in the centerpiece. Like for example, if we can detect that the meaty part of the content is in the middle and between these sections, then very likely that that will be the pooling content for your ranking. Meaning that that will help you the most with topicality, with relevancy, I guess. And then if we detect that something is outside that boundary, then that will weigh less and it will help you less with
ranking. And then you have, for example, footers where generally people just put a bunch of links. We kind of detect that that’s not that useful for users. And so it doesn’t help you all that much with ranking.

**John Mueller:** [00:19:05] Okay. So it sounds like if suddenly the comments of a site started to get indexed, it is something worth kind of looking at, but it’s not probably not your highest priority to figure out like, do I need to, I don’t know, clean out all of the comments and all of my old posts, it is like, it’s okay to look at it, but it’s not critical.

**Gary Ilyes:** [00:19:27] Yeah, sure. I mean, there is one corner case where you probably want to look at it. And that’s how many, for example, f-bombs, people drop in the comments and how R-rated the comments are because if you are writing about cookies, for example, lemon cookies specifically, and there are a few thousand comments below your recipe that are all of them R-rated, then that might easily confuse our SafeSearch algorithms, for example, and then your [00:20:00] content wouldn’t do so well when SafeSearch is on in Google Search.

**Martin Splitt:** [00:20:03] Fun!

**John Mueller:** [00:20:04] And links, just use rel="ugc"?

**Gary Ilyes:** [00:20:07] rel="nofollow", "ugc" if you want to help. I would definitely use nofollow, just to make sure that you are following the Guidelines that you learned by heart before you got hired.

**John Mueller:** [00:20:19] Thank you. Okay, cool.

**Martin Splitt:** [00:20:22] Oh, my!

**John Mueller:** [00:20:22] All right. Well, this was really interesting. I kind of liked the excursion into understanding what to do with comments. I think we can take a break here. It’s been great doing this session with you all! Looking forward to seeing how your cookies turn out, Gary, and maybe hearing a review from Martin in the form of a comment on a blog post at some point.

**Martin Splitt:** [00:20:44] I’ll comment on Gary’s blog.

**Gary Ilyes:** [00:20:45] They will be epic!

**Martin Splitt:** [00:20:47] They’ll be epic.

**Gary Ilyes:** [00:20:47] I will not have comments on my blog.

**Martin Splitt:** [00:20:49] Then I will comment on Twitter.

**Gary Ilyes:** [00:20:51] That you already do.

**John Mueller:** [00:20:52] Does Twitter get indexed?

**Martin Splitt:** [00:20:53] I think so.

**Gary Ilyes:** [00:20:54] Let’s not go there.
John Mueller: [00:20:55] Okay. All right. Well, thanks everyone for listening in. I hope you found this useful and insightful and a little bit entertaining, perhaps.

Hopefully I'll see you all again, or rather you'll hear us again in one of the future episodes. And until then, bye everyone!

Martin Splitt: [00:21:17] Bye!

Gary Ilyes: [00:21:18] Have a day.