EPISODE 23

[INTRODUCTION]

[0:00:00.4] JM: Welcome to Software Engineering Daily. Today's show is guest hosted by Edaena Salinas who is the host of The Women in Tech Show. If you're looking to guest host an episode of Software Engineering Daily, send me an email, jeff@softwareengineeringdaily.com. You can also find more information at softwareengineeringdaily.com/host. We're looking for outside voices who have interesting subjects around software engineering that they want to report on. Again, send me an email, jeff@softwareengineeringdaily.com.

Now, I'm going to hand it off to Edaena Salinas of The Women In Tech Show.

[INTRO]

[0:00:44.3] ES: A data-driven organization is more efficient because the company can learn what to focus on. Rya Sciban, given product manager at Periscope Data explains the needs of data teams in an organization. We talked about what data analysis is and how this changes as the amount of data goes. Periscope data has been successful in retaining women in product and development teams. We talked about effective strategies for this and for having more women in leadership positions.

[SPONSOR MESSAGE]

[0:01:22.4] JM: Spring is a season of growth and change. Have you been thinking you'd be happier at a new job? If you're dreaming about a new job and have been waiting for the right time to make a move, go to hire.com/sedaily today. Hired makes finding work enjoyable. Hired uses an algorithmic job-matching tool in combination with a talent advocate who will walk you through the process of finding a better job.

Maybe you want more flexible hours, or more money, or remote work. Maybe you work at Zillow, or Squarespace, or Postmates, or some of the other top technology companies that are desperately looking for engineers on Hired. You and your skills are in high demand. You listen to

a software engineering podcast in your spare time, so you're clearly passionate about technology. Check out hired.com/sedaily to get a special offer for Software Engineering Daily listeners. A $600 signing bonus from Hired when you find that great job that gives you the respect and the salary that you deserve as a talented engineer. I love Hired because it puts you in charge.

Go to hired.com/sedaily, and thanks to Hired for being a continued long-running sponsor of Software Engineering Daily.

[INTERVIEW]

[0:02:49.0] ES: Rya Sciba, product manager at Periscope Data is joining us today. Rya, welcome to Software Engineering Daily.

[0:02:55.8] RS: Thanks, Edaena.

[0:02:57.6] ES: Periscope Data is a tool build for data teams. What are the needs of data teams?

[0:03:06.8] RS: Great question. Actually, a little bit of history about myself, before I became a product manager I was actually a data analyst for a video game company. I worked with all the games across 22 [inaudible 0:03:20.0] of 22 games, and so I have a lot of experience on that front. I'm really excited to be able to bring that experience towards building tools for data analyst, like myself in my past life.

Obviously, there are many needs that data teams have. It may vary by team, by structure, and by the particular organization. However, there are a few key things that are pretty consistent across all of them, and the first one I think is pretty clear is easy access to data. Being able to access the data in a reliable and easy manner is pretty key to doing a job as a data analyst.

The second is a way to analyze and visualize the data. It could be numbers and comparing numbers in a table for ROI of marketing campaigns. It could be a heat map frame where the

biggest spenders live, but some way to look at the results of your analysis as well as to visualize and compare and analyze for what you're looking for to detect trends.

This kind of ties in to repeatability and reproducibility, so being able to repeat your analysis in an easy way is a very be efficiency game for data analysts. Finally, a way to share your results. You can spend eight hours of a day digging into one question, but at the end of the day you're measured and your impact is determined by how you share those results and those findings to the stakeholders.

Periscope targets all of those things. It focuses on ease of use which increases the analyst time to value. It focuses a lot on automation and workflow, so it makes all of those repeat tasks that an analyst does, it automates those into a way that saves the analyst tons of time. If the same stuff needs to be done, every week you're pulling a report for your boss. Now, you can just animate that and instead of doing it every Monday for an hour, you read it once and you've saved all of your Monday's time for everyone for it.

[0:05:25.7] ES: You mentioned that before you worked as a data analyst. What is the difference between data analyst and data scientist?

[0:05:33.9] RS: In industry, these terms can be used to pre- interchangeably. However, for the most part, the data analyst looks at — They look at reporting on the existing data, whereas data scientist focus more on modeling and forecasting future trends. Data analyst will work more with Excel and they'll work more with SQL, and they'll pull existing data that you have tracked and you have recorded and a data scientist will take those, take that date, they'll use a complicated, sophisticated machine learning models to predict and forecast how this will look in the future.

[0:06:15.2] ES: There are different ways of having that existing data that you talked about. I've log files, XML files, TXT files, databases. How is this existing data stored when a data analyst is working on it?

[0:06:33.3] RS: The most important thing is to have all of these in a consistent format. One big part that usually isn't done by the data analyst but is typically done by engineers is a process called ETL where we extract, transform, and load the data into a system where it is consistent,

where the data analyst can query and can pull data in a fast and efficient manner from that. All of these different sources of data eventually will need to be put into the same system if you want to do comparison and pull data from all sources and join that together to really get that 360 view of your business.

[0:07:14.9] ES: Is this end result particularly in a database, for example, a relational database or a non-relational database?

[0:07:22.7] RS: It typically is in a relational database.

[0:07:25.7] ES: How is the relational one different from the non-relational?

[0:07:29.5] RS: The main difference between relational databases and non-relational databases is the structure. Data in relational databases is organized based on relationships between the data that's stored. For example, if you're tracking purchases at a store, you might have orders, and so would have an orders table and each row in the orders table would indicate one order that was completed by customer, but then you have another table for the user where you would say — and it would indicate all of the users, and one user might make many purchases and one user might make one purchase, so there's that relationship there between that users table and that purchases' table. Whereas non-relational databases, they pretty much —the definitions is they store data in other forms other than an organization where it's based on the relationship.

[0:08:23.0] ES: Does the fact that it's relational make it easier to perform data analysis?

[0:08:28.4] RS: Certainly, in many cases, it does. The non-relational databases, NoSQL databases are typically more flexible in structure, but at the same time that flexibility adds ambiguity. It allows for variances in the data. As a data analyst, it can be harder to decipher and to analyze non-structured data because there is not that consistency in data that's guaranteed by the more structured form of storing in.

[0:08:58.8] ES: In the last decade, we've seen companies switch to this data-driven mindset where they're gathering data that they wouldn't normally save. These databases can keep

growing and get much bigger. What are the challenges of working with databases that have so much data?

[0:09:20.0] RS: There are some challenges with having a lot of data, but I would start by saying while the amount of data has absolutely a scale than grown considerably, so has the processing power of databases, but you can obviously have tons of data.

The main issue with really large quantities of data is the speed. There's two parts, one is if you're processing the data to be in a more structured format, you have to process a lot more data so that's going to take more time. Also, with your querying or just trying to figure out result from data and you want to look at all of it, the accrues will take longer, the analysis will take longer, and so this will just all add time towards how much you're spending on your analysis if you're not also at the same time scaling up the computational power that you're throwing at these problems.

[0:10:12.8] ES: One of the things that I saw at Periscope Data is that they have analytics clusters. What are analytics clusters?

[0:10:22.0] RS: I think there are two parts there. One is just an analytics database. It's a database that is better suited to run analytics queries just from the way that it set up. Most databases are what we would call row storage, so those stored data as one row. If you have a user, you'll get the user and you'll get the address, and you'll get their age, and maybe all get their email, and so on and so forth. Whereas analytics clusters, typically, will store data by the column, so you will have a column and you'll look at the first name of all of the users, and then you look at the last name of all of the users. You store it just differently.

What this does is this allows you a lot of efficiency gains when you're running queries that are typical to analysis, because for analysis, you're looking across one attribute for a large number of users or a large number of rows instead of looking at the entire row relating to one specific record. Just that setup itself makes it a lot faster to run analysis on these different types of databases.

The second part you mentioned was clusters. Clusters is simply a way to run multiple parallel computations over multiple sets of resources. We use Amazon Redshift. There are other databases that also allow for clustering. Basically, what they do is they break up your computation and run it in parallel across many nodes in Redshift's case.

I think a good example of that is if you're summing 10,000 numbers. Typically, if you want to add up 10,000 numbers, you have to look at each one and you add the next one to the sum and you get to the end and you're done but you've now looked at 10,000 numbers. What the parallelization or what the clusters do is it lets you split those 10,000 numbers into 10 batches of 1,000 numbers and you'll send each of these 1,000 numbers to a different server and it will add up those 1,000 numbers and they'll send it back to a central place where you'll end up where you'll sum up the 10 numbers. Because you can do this in parallel, it results in you saving a lot of time.

[0:12:40.7] ES: You mentioned the column-based approach for a database leads to much faster ways of doing analytics. I visualize this as an Excel spreadsheet where you start of by putting the column names and then fill in the data below. If my data is in the rows approach, not the columns, to perform analytics, does this mean I have to create a new database and that's column base and pre-populate it?

[0:13:10.2] RS: I think what you can do is you can copy the data to a different type of data stores, so that does store by columns. Redshift does it. There's a few others that do PostgreS, it does have some options to turn it into a columns store as well I believe.

Yes, if you wanted to make us of and take advantage of the speed increases of columns stores, you would have to migrate your data to one of those data sources.

[0:13:39.1] ES: I've seen that still a lot of people are using Excel, and we have these other tools that are better suited for data analysis, certain types of operations. Why do you think Excel is still widely used?

[0:13:56.0] RS: Excel is an extremely powerful tool. It is very flexible. It will allow you to do some things that are really hard to do in SQL really easily in a spreadsheet and. One of the

beautiful things of Excel is pretty much anyone can do it. It's just using math formulas. It's very intuitive, so it's very easy to gain a wide user base.

On top of that, it's pretty easily accessible to folks. A lot of people will use Excel in school, in their jaws, and they'll be very used to spreadsheets. It's something that they're very comfortable with. Excel does some things very very very well and much better than a database. For example, they do a lot of post-processing and a lot of ad hoc analysis and you can copy-paste things from different sheets.

I think Excel will be around for a long long time. I think the use cases for Excel versus large structured data analysis from data in a database are a little bit different, but Excel will certainly still hold its market share in terms of flexibility and a lot of the exploratory and ad hoc analysis that pretty much everyone does on a day-to-day basis at their jobs.

[0:15:10.1] ES: Tools like Periscope Data involve using the customer data because you want to get analytics in your data. What are strategies for pricing a product where the data varies in size and shape?

[0:15:25.2] RS: The pricing is definitely one of those interesting conversations, and there's a wide range of strategies throughout the industry. Here at Periscope, we believe that the best strategy is one that fits the customer's requirements and needs and scaling the pricing based on value metrics, like rows or database size is easy to understand for people, but only if we also scale it alongside a set of features and functionality that actually helps them get more out of their deployment and their data. It's a little bit of a two-part system.

This is a little bit different than most of our competitors who will do per-seat or per-user pricing which we find runs counterintuitive to having a more data-driven organization. We view per-set pricing as more of a blocker to having more data-driven organizations, and Periscope, I think, shines the most when you have all of your data in Periscope and you're really driving all of the insights from there.

[0:16:32.6] ES: Customer insights are an important component, like you mentioned, for determining the price of a product to do it based on their requirements, but it's also important to

collaborate with customers for product development. You, as a product manager, how do you collaborate with customers?

[0:16:52.5] RS: Absolutely, as much as I can. In many different avenues as I can as well. We get customer feedback from many different sources. We have account managers who are in charge of particular accounts. We have our customer success team as well as our customer support team. They have a lot of calls with customers and I will join as many as I can just to always hear what customers are working on, what they're doing, what they're not doing and why.

More specifically, we also do work with key partner customers who'll help us if we're doing a beta of a feature or they'll help us out with some user testing just so we can identify how well the solution actually fits their needs and identify any areas of improving the user experience or identifying next steps in the future better.

[SPONSOR MESSAGE]

[0:17:44.4] JM: For more than 30 years, DNS has been one of the fundamental protocols of the internet. Yet, despite its accepted importance, it has never quite gotten the due that it deserves. Today's dynamic applications, hybrid clouds and volatile internet, demand that you rethink the strategic value and importance of your DNS choices.

Oracle Dyn provides DNS that is as dynamic and intelligent as your applications. Dyn DNS gets your users to the right cloud service, the right CDN, or the right datacenter using intelligent response to steer traffic based on business policies as well as real time internet conditions, like the security and the performance of the network path.

Dyn maps all internet pathways every 24 seconds via more than 500 million traceroutes. This is the equivalent of seven light years of distance, or 1.7 billion times around the circumference of the earth. With over 10 years of experience supporting the likes of Netflix, Twitter, Zappos, Etsy, and Salesforce, Dyn can scale to meet the demand of the largest web applications.

Get started with a free 30-day trial for your application by going to dyn.com/sedaily. After the free trial, Dyn's developer plans start at just $7 a month for world-class DNS. Rethink DNS, go to dyn.com/sedaily to learn more and get your free trial of Dyn DNS.

[INTERVIEW CONTINUED]

[0:19:44.1] ES: How do you evaluate what customer feedback is incorporated? I assume not every piece of feedback makes it as a feature.

[0:19:53.2] RS: Absolutely. I think we have a list of around 2,000 feature requests. For sure, we're unable to keep up with that pace and not all of them are aligned with our product vision. I think it's important to listen to all of the customer feedback, but the really important part is understanding why they have that feedback and where it's coming from. Often times, you'll hear a request for feature and you might think, "That doesn't make sense."

The follow up question there is; why do they want it, and maybe they're trying to accomplish something else and there's a better solution that fits more cohesively with their product and also to get more value to the rest of the customers. We certainly don't deliver on every single customer feature request, but we do listen to them and we do look at the high-level trends as well as the individual request and will talk to customers at individual request level to get more detail and to really drive what the next feature development initiatives.

[0:20:57.4] ES: How do these processes for taking feedback from customers and meeting with them change as the number of customers keeps growing?

[0:21:07.7] RS: For sure. I remember when I first joined Periscope, I think I was employee number nine. Obviously, the number of customers we had was a lot smaller than the number we had then. Back then, I knew all the customers and I knew everybody who was on chat-support and emailing us. I was chat-support. I was emailing with customers and doing customer success, and so I was able to keep a really good heartbeat on all of our customers.

As we grew, bandwidth of one gets pretty tough and we've had to hire teams so that we have a great team of support staff and a they knock it out of the park. Really, they're rock stars with

their customers. We have a group of account managers. It's more about building processes with these partner teams as well as looking at the raw data itself to identify how our customer is feeling, what are some of the trends that we're seeing customers and what are trends and feature requests and what customers want to do and where they want to go with the platforms. It's more about processes to bring all of these different sources of feedback together.

[0:22:11.7] ES: You're a product manager currently. What is the role of a product manager?

[0:22:18.0] RS: The role of the product manager is really interesting one. When I first started as a product manager at Periscope Data, I kind of had to ask the same thing myself. One thing I've learned over the years is that there really is no one definition. It really depends on what the organization is, what their product is, what the actual needs are. At Periscope, I can tell my role at Periscope. I own NPS for our visualization product, which is our core product that we have today.

[0:22:47.8] ES: What's NPS?

[0:22:48.6] RS: It's net promoter score, so driving customer happiness and really being able to deliver on increasing value for our customers. We have two. I work with two engineering teams, and I work with them on setting the roadmap as well as the overall strategy for those teams on what we want to accomplish.

On a daily basis, I will do anything from collect feedback from customers, do customer research. I'll marry that with feedback from our sales team, our account management, our support team, as well as the strategy and goals for the company. I'll also be setting goals and developing product specifications and requirements based on the feedback from our customers. I'll work on product design. I'll work pretty closely on iterating on product design with our design team to figure out what is the solution space, what is the best solution for our customers in their specific set of needs, to product implementation, coordinating with the engineering teams, coordinating talking about technical versus product trade-offs.

We really could build that extra button. It would take two more months. Do we want to? Proposing features, making sure purpose of features, so talking to the engineering team about

why we're building what we're building, what the motivation, what the customer need is. Behind that, it's very important to make that component that make sure all the teams are aligned behind the same goal, and then just making sure that the engineering team is running as efficiently as they need to be, and then partnering with other teams.

I talk a lot with our marketing team, with our customer success team, with our support team, with our sales team. We're in charge of selling the product, what's the positioning that we're going to go to our customers and potential customers with. What kind of training do we need to educate customers on new features? Where does it fit in with the rest of the product? How does it add value to our customers?

Then closing the loop, so after a new feature is developed, gathering the feedback, taking a look at the metrics. Did we actually hit our goals? What was the impact of this feature and what do we want to do with it next?

[0:25:01.6] ES: You mentioned that you communicate with people from other teams. We were talking earlier about companies, gathering data, and getting insights of these data. Another important part of this is sharing the data between the organization, and the important distinction here is not everybody is an engineer or a data analyst. What are effective ways that data can be shared between the organization, for example, between marketing, and recruiting, and other areas?

[0:25:35.1] RS: Yeah. At Periscope, a lot of our customers, pretty much all of the all the departments, will use data with respect to marketing and with respect to sales. Often, you'll see either a centralized analytics team that's in charge of creating analyses and digging into questions for their partner teams, or you'll see an analyst kind of embedded with those teams. You'll seek a marketing analyst who works with the marketing team and they have access to the data and they're in charge of all of the analytics for the marketing team.

There will be someone serving that data to these and users. There's also — You can build out tools and flexibility for themselves. You'll see a lot of her competitors use drag-and-drop tools to enable their end-users marketing, recruiting sales to pull their own data. We've found that there is a lot of risk in that because if you don't define the data models properly, it's really easy to

make a mistake. I think this is a lesson that some of our internal folks have learned as well. They've pull data assuming it was correct, and because they didn't have the full context and full understanding of the data, they've made it come to an incorrect conclusion.

At Periscope though, we do have a lot of tools for our business users, so we have a lot of filters so you can really filter down your data. You might have one dashboard that shows you the customer usage and you control through it by day or you can filter it by user so you can drill into some of the data to get your insights there. We also have some business user features such as pivot tables that you can do directly in Periscope. You can request kind of a dataset from one of the data analysts and they'll put together — Put the pulled out together for you and you can pivot that directly in Periscope yourself without even having to leave the system. Then you know that that data is correct and that the logic behind it is sound.

[0:27:35.3] ES: Another thing that you mentioned earlier was getting the feedback and evaluating the effectiveness of the product or the product satisfaction. What metrics do you look at for this?

[0:27:47.4] RS: Yeah. There are a lot of metrics that we can look at and it really depends. I think I'll give you an example and I'll give you a pretty clear idea. One thing I recently did was I did an analysis on how our customers are using the filters that I mentioned. They're one of the most popular features in Periscope, and so I was digging into how exactly they're used to the data.

Obviously, I looked at things what percentage of customers are using our filters, break this down by customer profile; are they data analysts, are they business users, are they administrators. How often do they use them in any given day? How many dashboards might these filters exist on? How people are actually changing the filters versus just setting it once and never changing it again? What interesting thing that we found was, "Well, almost all of her customers use filters. Only 20% of dashboards used filters."

Based on this, we actually went ahead and we modified the product to hide the filter layer, the filter feature and it can now be enabled for each dashboard, but by hiding it, it present a cleaner surface for people to present their data and it just reduces that clutter.

[0:29:02.1] ES: Did you notice this insight by gathering data programmatically?

[0:29:08.0] RS: Yes.

[0:29:08.6] ES: Another way that I know of gathering data is getting customers in a focus room and just watching them. What things are better to notice in a programmatic way versus an in-person way?

[0:29:23.1] RS: Yeah. I think the type of insights you gain from talking to people is more about exploring their workflows as well as looking at how they are using your product and their workflows within your product, it's more data around workflows. How do they interact with the workflows that you have developed? Are they efficient? Do they work? Do they make sense? What workflows aren't addressed by our product? These things are — is data that you gain really well from just talking to customers or looking at them using their product.

There are other things like how many queries do you run a day? I'm pretty sure if you ask one of our users, they would have no idea how many queries they ran in a day, but that something that's really important for the technical stability side of the system. That's something that we can pull really easily from data and we'll get an accurate number, but if you ask the user, they'll get something way off.

[0:30:21.2] ES: I want to talk about company culture next. Google is 19% female in technical roles. I saw Periscope Data is 60% women overall, and 40% women in the product and development team, so in technical roles. What are effective strategies for retaining women in tech?

[0:30:45.4] RS: I think since the last time you checked, probably, we've hired a ton and I think we crossed 100 people last week or the week after. It's really hard to keep track, but right now we're sitting at closer to 43% women and then 57% men. I think the product and engineering is pretty spot on, so we're about 40% women and 60% men as well in the product and engineering side of things.

I think most it is just about the culture. It's about being accepting and supportive of everyone. I haven't seen any issues, but I have a feeling and I have full confidence in the team here that if I or anyone else at the organization were to raise any issues, they would be taken seriously and they would be addressed promptly. I think just that accepting feeling and as well that supportive of feeling and that confidence that things will be handled in a manner that's appropriate really goes a long way to good attention across the board.

[0:31:52.6] ES: Are there internal groups for connecting with other women?

[0:31:57.9] RS: Not at Periscope. No. What we do is we do have catered lunches every day, so we started this process when we are at our last office, in a dog patch, and we're a little bit far away from all restaurant, so we wanted to provide a convenience for employees. What we found is it's lent itself towards a very inclusive culture where that's where I'll go to chat with anybody else.

Most of my day, I'm talking with probably product engineering and maybe some marketing and sales. Now, I can go talk to some of the roles that I don't have any chance to interact with. On a day-to-day, I'll sit down and have deep discussions with our front-end salespeople or our sales development representatives, or our account executives, or some of our data analysis team. I can really chat with everyone, and it's very friendly and it builds a very warm and welcoming culture and that's where I'll be making a lot of friends.

[0:32:52.5] ES: Once there's women in technical roles, we've seen that there still a small number of them in leadership roles throughout different companies. What do you think are effective strategies to have more women in leadership roles?

[0:33:08.1] RS: That's a great question. I think the first is to have women in leadership roles. Our VP of engineering, our head of people ops are both women. I think, across the board, 29% of our leadership roles are held by women. Periscope don't have any bias in hiring, but I think we always want to hire the best person for the job.

If you look externally for a candidate, because there are fewer women in leadership roles and if you're looking to hire into a leadership role, it's possible that you'll just have more male

applicants. One strategy that you can do to ensure that there are more women in a candidate pool is to actually nurture and grow people, like women and men, equally in the organization so that they will be in positions to be promoted into the leadership roles.

[0:34:00.7] ES: What I liked was what you said earlier, where you're giving a chance to network with people from other areas in the company. They're not necessarily only women, which I think helps even more, because you'll just network and talk about the product and ideas and have more visibility that way.

[0:34:19.0] RS: Mm-hmm.

[0:34:19.7] ES: Rya, thank you for coming on the show. It was great talking to you today.

[0:34:25.3] RS: Thank you very much, Edaena. It was my pleasure.

[END OF INTERVIEW]

[0:34:31.4] JM: You have a full time engineering job. You work on back-end systems of front-end web development, but the device that you interact with the most is your smartphone and you want to know how to program it. You could wade through online resources and create your own curriculum from the tutorials and the code snippets that you find online, but there is a more efficient option than teaching yourself.

If you want to learn mobile development from great instructors for free, check out CodePath. CodePath is an 8-week iOS and android development class for professional engineers who are looking to build a new skill. CodePath has free evening classes for dedicated experienced engineers and designers. I could personally vouch for the effectiveness of the CodePath program because I just hired someone full-time from CodePath to work on my company Adforprize. He was a talented engineer before he joined CodePath, but the free classes that CodePath offered him allowed him to develop a new skill, which was mobile development.

With that in mind, if you're looking for talented mobile developers for your company, CodePath is also something you should check out. Whether you're an engineer who's looking to retrain as

a mobile developer or if you're looking to hire mobile engineers, go to codepath.com to learn more. You can also listen to my interview with Nathan Esquenazi of CodePath to learn more, and thanks to the team at CodePath for sponsoring Software Engineering Daily and for providing a platform that is useful to the software community.

[END]