EPISODE 1202

[INTRODUCTION]

[00:00:00] JM: A data-driven organization collects a wide variety of data to help in strategic decision-making. The cost of storing large amounts and variety of data has dropped dramatically in the last two decades, but too much unstructured data may not improve decision-making and can even lead to analysis paralysis. Organizations react by extracting the most important, actionable data and placing it into a data warehouse, which has a pre-designed structure meant to streamline the data in preparation for analysis. The key challenge with this approach is identifying what should be streamlined and how to structure the data warehouse to focus on the most important actionable items. This is especially important for organizations seeking to scale as the necessary structure to generate the most relevant insights may change as the organization grows. Narrator.ai is building a data intelligence that uses a simple, proprietary, universal data model to help organizations streamline their data warehousing. Narrator is built on the belief that data tells the story of a system and its platform empowers organizations to use these stories to make better decisions.

Ahmed Elsamadisi is the founder and CEO of Narrator. Before founding Narrator, he spent several years working in data analysis and algorithm design for WeWork, Raytheon and Cornell's Autonomous Systems Laboratory. He joins the show today to talk about how Narrator generates the most actionable insights from a data warehouse and why a universal data model is so important when scaling as well as what makes Narrator's approach to data analysis different.

[INTERVIEW]

[00:01:33] JM: Ahmed, welcome to the show.

[00:01:35] AE: Hey, Jeff. How's everything going? I'm excited to be here.

© 2021 Software Engineering Daily

[00:01:38] AE: Yeah, it's great. So let's say I've got a ton of data in my company. I've got marketing data. I've got data on how my customers are interacting with my website. I want to do useful things with that data. It's 2021. Can I do useful things with that data?

[00:01:58] AE: Yeah, you can do a lot of really incredible things with that data. You can understand how – You can actually start changing customer behavior just by understanding how customers behave and figuring out what makes some customers behave better than others and then pushing all your customers to do that behavior. And that is done by an analysis. The process to get there is a little bit of work, but the goal is always the same. It's for you to make better decisions that cause your customers to behave differently, which leads them to your company to kind of increase the revenue and decrease their spend. The big question is how we do it, right?

[00:02:37] JM: Yeah, exactly. I mean I would be tempted to just hire a data scientist and say, "Hey, data scientist. Start to figure out these kinds of problems." Is there anything wrong with that approach?

[00:02:49] AE: Well, the data scientist is going to have to do a lot of things before they can use the data. So your audience probably knows a lot about data, but for those who don't, data is captured in all these different systems as independent pieces. So we often put them all together into what's called a data warehouse. And to give you an idea, if you're like a small startup like us, you have an internal database, you have your segment. You have your ads. You have your email client. You have your CRM. That's already around like 10,000 raw tables of data, and your data scientist has to like figure out ways to combine that data and make useful sense out of it and make sure that he or she is able to actually make decisions off of it. That process is what's common known as ELT. EL is when you take the data out of the systems, dump it somewhere. And then transformation is when you take that data and structure it to make it useful. And most companies spend like a lot of their resources, and what majority of the work is done in is in this thing that we call transformations. It's taking this data and making it useful so you can begin to analyze it.

A good example I would like to say is like imagine words. If you're trying to understand something and you have all the words in the dictionary, it's not really useful. There's somebody who has to take it and structure it into stories that that story can be consumed. And with all these different topics and different questions, you often need to create multiple stories so that someone can actually consume that data and make meaningful decisions out of it.

[00:04:27] JM: Gotcha. So you've touched on some of the technical difficulties of working with data. It's worth now getting into a little bit about what you do at Narrator. Explain what Narrator does.

[00:04:41] AE: So Narrator is about giving customers expert handwritten analyses in minutes. So we realized that the questions that everybody's asking like, “How do I optimize my LTV? How do I decrease my CAC? What attribution model should I use?” They're all complex questions that require a lot of analyses. And like I, said every analysis requires different data to be transformed. So Narrator came at the world with a very different angle, and we said what if all the data looked the same? What if every single company, all their data was transformed to be the exact same data model? And what if it was just one single table that was 11 columns?

If we can get all the companies to have the exact same data model, then we can start reusing analyses, and that whole process of transforming data and trying to get the data in the way you need it to answer question goes away, because we can actually build analyses on top of this standard data model. And what Narrator does is it provides that entire end-to-end experience. We help companies take the raw data and make it into our – What we call the activity stream, which is the standardization of data that is the same table for every single company independent of your industry. And then because of that structure, you can instantly run any of these analyses that we have handwritten by data scientists who are experts in the field. And these experts get to write these analyses and test them with multiple companies. And once they make that they're the best analysis that they can produce, they make it available in Narrator. And then any company that uses the Narrator standardized data model can run that exact analysis in a minute. It's like skipping the line of all that hard work that you would have to

do. And it took us about three years to do this thing, which is standardize all of data so that you can answer any question.

[00:06:41] JM: So if i standardize all the data in my company into one schema, like what are those columns? What are the columns that you focus on?

[00:06:51] AE: So what it turns out is remember when we talked about the point of analysis? We mentioned this idea of changing customer behavior. So it turns out like from the old time, and this is very common in robotics, this universal data model is very simple. It's an entity, often a customer, doing an action in-time with a couple of metadata columns about that action. So the 11 columns are pretty much customer timestamp. Some features on the activity, usually three, and the activity name. And with this simple structure, you kind of can imagine a timeline of a customer. And what we've developed is a very unique way of querying this time series data to enable you to answer any question.

So think of example, any question like someone who came to the website called us. Did someone who come to – Like what percent of customers who came to our website called us? If you think about it from like a timeline perspective, you can look for every person who came to the site and see if they called us a couple minutes after they came to the site and we can count how many people did that. Traditional data modeling would actually put – This would be a very hard problem. It took me like three weeks to answer this exact question at WeWork where I ran the data team there.

So this unique structure of putting everything in this time series way and asking questions about a customer in-time allows us to kind of create this new language of communicating with data. And with this new structure, which we've proven, can answer any question, allows us to kind of copy and paste these analyses across companies. Creating a world where things are shared. And in the future you'll see more and more of that where companies are no longer hiring more and more data engineers to build data models so that they can answer more questions. They just do the upfront work, which takes our customers an average of a week to fully set up Narrator with 20 activities. And then they can unlock thousands of questions they

can answer and any ad hoc question they can do in minutes. And that's kind of like the power of standardization. And this has been seen in history with many, many companies before us. Like Salesforce did it with standardizing sales and then opening up like the stack exchange, this idea of standardization to share like from the industrial revolution with interchangeable parts. It's been like a historical thing. And now we're bringing that to data analysis.

[00:09:21] JM: Can you take me through a use case to illustrate this in a little more detail? Because it's still hard to – I think it's a little hard to believe upfront that you could standardize the data schema for so many different companies.

[00:09:36] AE: Yeah, I know. I've spent the last three years literally trying to get people to see it and try it and slowly, and more and more customers have been like seeing that it works. So let's take a company of your choice. Name any company you want and I'll do this live. Like name any company.

[00:09:56] JM: Okay. Let's say – Can I say Software Engineering Daily?

[00:10:02] AE: Great. Great. Great. Great. Software Engineering Daily. Great. Great. Still, you have customers. Your customers visit your website. Your customers maybe download a podcast. They might start listening to a podcast. They might stop listening to a podcast. You might have all this podcast information in your internal database. You have your web visits from your segment. You send them email reminders to listen to a podcast. And they interact with you many, many more ways. You might spend money to acquire new listeners. What are some things that you can imagine if all these different things that the customer does?

So let's take these simple examples, these five examples that we have and imagine we write them in a single table. So a customer comes in, starts a podcast. Customer receives an email. Customer opens that email, and you have dealt with three different systems, and we just inserted all these activities into one table. Then you ask a question. For example, one of our narratives that people really like, is how does emails impact returned listeners? And then you press run. And what Narrator will do is through our approach is it goes through every single –

Everyone who listens to a podcast and sees how likely they are to listen to the next podcast. Then it checks for how likely they are to do it if they had an email in between. Then it checks for consistency. Then it actually adds more layers to it. Does the number of emails they open between each podcast impact how likely they are to listen to a second podcast? And again and again and again. Does the time they takes for them to get their first email impact how likely they are to listen to a podcast again? And this whole thing that it's doing right now is just, again, looking at a single structure, and we're just talking about a customer doing something in-time, and Narrator writes this entire story and you can read it in plain English. It will appear to you like a McKinsey presentation and it kind of looks at the natural behavior of the customers. Groups them to see the ones who open the emails versus the ones who don't. Does the number of emails matter? And it goes through and checks it. Is that behavior consistent in-time? Maybe it used to be the case a month ago, but now it's no longer the case. And it slowly tells you – And it ends up making a clear recommendation. It goes, "If you got like your customers to open an email after their third podcast and you got them to open their first email within one hour of listening to a podcast and you only send them three emails, you're going to increase your listeners by 20%."

And what we watch is customers see that recommendation, follow their data, understand that recommendation and really, really believe it, which is one of the hardest things to do is get people to act. Then they act upon our recommendation and they monitor the results as they see why this analysis is right. And that rigor of doing that analysis has made our customers very happy. And this exact analysis, somebody asked it about how does email impact getting customers to return to the website? And they followed a recommendation. And they increased their sessions by 17%, which is like a huge number in like returning visitors. Does this kind of make sense?

[00:13:17] JM: It does. Definitely. So who is making this analysis?

[00:13:22] AE: Right now it's us. So Narrator has its own team of data scientists. We hire experts in each specific field. And each one of them takes – Because we get to reuse our analyses, each one gets to focus on one question for a couple weeks and then we get to test

that analysis on a couple of our customers. So when you look up our analyses in our library, it'll tell you this analysis was written by like Brittany, the head of data of Narrator. She ran this analysis for 25 companies. And we know this analysis has been like kind of been battle tested, which is something that a lot of data analysts don't get to do. Like you don't get to build your LTV model and/or an LTV optimization analysis and test it with 50 companies to see to make sure that it's reliable for all these different companies in different sectors.

So right now we have an internal team that is building these analyses. But pretty soon we're going to open that up. So like in the kind of mid this year to end this year we're going to launch what we call the Narrator Store. What that's going to allow you to do is anybody who is using Narrator and is building their analysis in Narrator, they're going to be able to actually kind of like open source it and share their analysis to the whole Narrator community. So then you'll imagine for every question you have, you'll just look and build upon what people have already done. You'll see an analysis available by another company who's done it and you get to just simply click run and run it with your own data.

All this thing is done without data ever leaving your system, no data transfer. What Narrator has done is captured the way you analyze data and make that reusable. And because of the standard layer, it works for every company. Like it's still like kind of like the magical thing of Narrator. And it's really hard because normally I would show it in a demo, and I think that's like what gets people to really get excited. But in a podcast, you have to kind of imagine a little bit more. But once you standardize data, it opens up so much more of this shareability aspect. And that means that the entire data community gets to move forward as a unit together. Like everyone right now is starting kind of from scratch and there's tools to help you, but everybody's building their own models and trying to figure out their own way of structuring data and that we're all learning the same mistakes and it makes analysis take so long. I would much rather enable everyone to kind of get the data they want instantly. Thanks to Narrator standardization. Get the analysis that they need right away and have your data scientists focus on the little bit of the unique stuff that's really, really requires a deep dive in your business. And that's kind of what our customers really appreciate.

[00:16:11] JM: So from a business perspective, it's kind of a dangerous position to be in, right? Because right now it feels like you're a little bit of a consultancy, right?

[00:16:20] AE: When you receive our analyses, it feels like consultancy. But we don't write them for you. Like we have a library, you look them up you search them and you run them and you answer a couple of questions using your data. So you're doing the whole thing relatively self-serve. But when you receive our analyses, they look like they're handwritten for you. Like most of our customers will tell you that they can't believe that this was not handwritten for them.

[00:16:45] JM: Got it. So I understand. So first, before the type of analysis is viable for me to use as a customer in a self-service fashion, it first has to be vetted by the manual process.

[00:17:02] AE: Exactly. Before we add an analysis to our library that you can use, we actually build them and test them with a couple customers in our beta and then we add it to our library. And we have a decent amount of narratives right now in our library available. So like for the first like 50 questions you would ask, you would get them off the library. And if you don't find something, you submit it. And one of our data scientists will pick it up. Do it by hand. Test it with a couple of companies. Test it with you two too make sure that the analysis makes sense and there's no gaps, there's no conditions that we didn't think about. The recommendations are valuable. And if you take action upon them, they produce the results that we expect, and then we add it to our library, and this is a cycle. So the library continues to grow. And that's what I mean. They're literally handwritten analyses by experts. The magic of it is that it's delivered to you in a minute. If it's in the library, it's instantly available.

[00:18:01] JM: Got it. So can you talk a little bit more about your infrastructure? I mean I guess in order to develop these kinds of narratives, the data science around them, you have to have access to some customer datasets, right? Like the customers have to give you access to their data warehouse? Is that right?

[00:18:23] AE: Yup. Narrator sits on top of your warehouse. So we just connect to Snowflake Redshift and you follow a very simple tutorial. It averages around 12 minutes per activity where you kind of map your data to the building blocks that we call activities in Narrator. So you're like, “Okay, this is what a visited website is. This is what a call is. This is what a received email is. This is what an open email is.” By just simply writing SQL, tiny little SQL snippets that average around 25 lines. So like kind of a little bit more like a very basic version of transformations. And what Narrator does is Narrator – Actually everything is done on the warehouse. So we take those little snippets, insert them into this table, do a bunch of cleanup on it and make it to make sure it's really fast and efficient. Handle like cookie mapping and identity resolution. We do like a multi-user, multi-device in-time identity resolution. Probably one of the world-class ways of connecting a user across multiple systems. And then you have the single table that we call the activity stream in your warehouse. And when you're using Narrator, what we're doing is we're actually just compiling very special queries on your warehouse that run with that using that single table. Kind of taking that table and self-joining it in very clever ways. And that's kind of what gives you the data that you need for the analysis.

So the input to starting Narrator, often people do it in the first couple days. It's really easy. You're just defining the building blocks. And then in our documentation we actually have like a lot of the common websites like Shopify, Segment, Sendgrid, all these platforms. If you're using like Fivetran, we have like SQL snippets to kind of build those building blocks for you automatically. And then once you have your snippet that defines your business logic, then you're ready to start using Narrator. And you can assemble any table you need in our tool or you can run our analyses instantly and it will assemble the data it needs, because Narrator is aware of the single table structure that it uses. So it knows how to actually get the data it needs out.

[00:20:35] AE: Yeah. So in terms of actually developing these narratives and the engineering around them before you actually make them available to self-serve as the customers, do you have a bunch of like sample datasets that you develop that you work with internally to build these recipes?

[00:20:55] AE: Yeah. So we have a whole demo account where we actually wrote our own like auto-generated data that we often build our analyses on this auto-generated data. The thing about like analyses done very well, I came from a robotics background, like what you're capturing in a narrative is a thought process. It's a systematic way to understand and analyze data to make a decision. That's why we can actually interchange parts. So like that question I used earlier, which is how does emails impact people returning to – Start listening to another podcast? It's the same question as how does email get someone to come back to site, which is the same analysis as how does calls get someone to come back to a site, which is the same thing as how does tickets get someone to like delay payments? Like it's the same structure of how you would answer that question is actually a very systematic flow.

So we often like to start by doing the analysis on fully demo data. Then we would actually reach out to a couple of our customers. Let them know that we're doing this new analysis and see if they're interested in being our beta. And then we run that analysis for them for free. We hear their feedback. Often the feedback is about like what's confusing in our text, because if you saw on our website and in our docs, our narratives are stories, like you read them as a story. So they like tell you the takeaways and they give you details about why we make those takeaways and they give you plots as supporting evidence to the story. It's not plot first. It's actually text and takeaway first. And then plots are supporting evidence. So often, we spend time with customers making sure that all our takeaways are very clear and crisp and that they understand the narrative fully. And that's kind of what this iteration process of running the analysis from multiple customers and seeing if there's any edge cases that we didn't consider. And at the end of it ends up being the story that's like templated out that we can reuse.

Think about it – Like the way I like to think about it if you want to wrap your head around it is if you're McKinsey or any consulting firm, you have a systematic way of doing analyses and it's based on like kind of repetition. You have like a method that you go through and you're going to check some things and see if they impact it and you're going to do your analysis. If all your inputs were the same, you can kind of systematically kind of automate that analysis. And that's what Narrator does. It captures that mental logic and the systematism of the way you analyze

data for a very specific question with very specific assumptions. And that's what we make available in our library.

[00:23:37] JM: So it's taken you more than three years to get to where you are today. What have been the toughest technical challenges?

[00:23:44] AE: Yeah. So figuring out a single data model that works for every company, people often think that's the challenge, but that was actually the easiest thing. However, getting a single table to answer any question is really hard. And the nuance of why it's hard is because of this thing called foreign keys. Like the way that you query data is you have to join it with keys. If you're trying to use a single table, the relationship relationships no longer exist as keys. Like there's nothing that ties a session to a call. The only thing you really have is customer in-time. So we had to actually reinvent how you query data and create new ways to assemble the data that you need. Ways that can be comprehensive enough to answer any question and simple enough that people can pick them up. And like I wish one day we'll release our first versions that some of our customers had to suffer through. Like it used to look like an airplane dashboard with a thousand buttons to try to use our new way of querying data. And today, it's a single 11 option drop down and a couple of sub-filters if you need it.

So that iteration about – We iterated on the system probably seven times to get it to be the simple way of asking questions where you get to use words like give me the first ever web visit they have and give me their ad source. And create that system that allows you to answer any question with a single table. This has like been the thing that I still am amazed that it works is that. Like we've been doing this for three years. We used to battle test our way of querying data. We used to offer customers unlimited questions. So they would actually ask us, Slack us any question and we would set up Narrator for them, kind of like a consultancy early on, to make sure that our technology worked. And we would get like hundreds of questions that we would have to answer in five minutes. And just seeing that we can consistently answer any question that somebody asks using the same single table data model and our really simple language that we created, that like was the hardest thing. And even today, explaining it – I wrote it and I'm like amazed that it works. I'm always like, "Wow!" Like, "I have no idea how

this came to be through years of blood sweat and tears." Somehow this came together. But that is probably the true innovation that makes everything we do possible, is this way of relating data without using foreign keys, which is like the most counter-intuitive thing, because that is the entire concept of data analysis and like the entire world that exists today from like Looker and dimensional modeling and every single thing you see. The assumption that they have is that you have a foreign kid or late data. And when you break that core assumption, like you have to kind of reinvent data analysis from scratch.

[00:26:55] JM: So can you just go over that again? I'm a little bit confused. So how do you join data against each other when you don't have a foreign key?

[00:27:06] AE: Yeah. So we use customer and relative time. So think about if you have one customer and I told you like the same example. If somebody came to the website and called you, okay? Traditionally, that call should have a session ID that you can join two sessions, right? But in reality, your call is in Salesforce and your session data is coming from a page view in Segment and there's never that foreign key. But if it was all in one single table and you only had one customer, you could imagine you would just go down the customer's activities and you would see they had a session and just look underneath it in order to see if they had a call right afterward or not. If they did, then they did have a call. If they didn't, then you just put no. That exact kind of methodology, we call it the first in between. So give me every session and give me the first in between call. So it's give me the call that happened after that session before the next one. So that first in between is how we relate data. It's an example of one of them. We call that a relationship. So it's relating the objects or the activities using customer and relative time. And it turns out you can make that really do anything.

So any question you have can always be reassembled to describing something happening in time and you're pulling something out of time. Another easy way to really wrap your head around it is think about how you debug a join. If your thing isn't working, what you do is you follow a customer and you make sure that things happen in the right order so that it made sense. And you use like you look up a customer in every table and look at what the time stamps and see if the table that you're creating makes sense. Instead of using it to debug, we

make that the essence. So customer and time become the unique – Like those two together and relative time of the activity is what we use a replacement for all of foreign keys.

[00:29:14] JM: Very interesting. Again, this is on tables where you're working with a customer and you've built this complete customer table for one of your users like on top of Snowflake for example.

[00:29:33] AE: Yeah. Customers built it themselves, and they find it to be the easiest thing. So customers are mapping this data by just like writing basic SQL queries as building blocks and then our system takes that little building blocks and generates that one table. And then once you have that table, you get to use our data set tool, which uses our unique language of relating data that's foreign keyless to assemble any table you want. It's kind of like an end-to-end system that we had to build because we reinvented one of the smaller parts of one of like the core assumptions of all of SQL.

[00:30:11] JM: So if I want this like whatever question I'm having answered. Like if I'm trying to answer some question about improving customer lifetime value and I want this to be periodically updated, I don't want to just run it once. I want to update every single day. What do I need to do to have that question be answered every single day?

[00:30:33] AE: That's already built in. When you actually run the analysis, it actually asks you how often you want to run it and you pretty much choose a cron schedule. And Narrator actually keeps track of every single snapshot when it ranks. You can actually go back in time and say, “What did this analysis look like six months ago? And let's see that exact analysis.” So it's running those analysis and it actually will update the recommendation. So like we're about to actually release the feature where it will email you if the recommendation changes so that you can go and say, “Oh, the recommendation changed. Relook at the analysis.” And it's all kept in the same place. Because as a company, you do a lot of great analyses and you end up redoing a lot of great analyses every six months because you forgot you did them. But Narrator doesn't let you do that. Narrator keeps track of it. It's rerunning it built-in. You can go back in time and see what that analysis looks like and the recommendations and decisions that

it makes are constantly updating as your data changes. Like we see this all the time especially with CAC. And I'm going to keep using these like high-level examples just because I think everyone understands them. But like your CAC changes. Sometimes AdWords is better consistently for three months and then it's Facebook is better, and then there's all these different decisions and behavior of your customer changes. And having that analysis constantly run and constantly update the recommendation to help you make sure that you are really tied to the pulse of your customer and always making the best decision is what Narrator is built for. So we make sure that it's constantly running, constantly letting you know and constantly making updated recommendations so you're always making the best decision.

[00:32:22] JM: So I really want to drill a little bit deeper on this kind of data normalization that you do, because the normalization for the data to get put into the activity stream. So there're all kinds of different platforms that people are using to generate customer data. You've got Segment, Salesforce, Zendesk. Name your platform, Shopify. And again, you are trying to normalize this data into 11 different columns, right? Can you tell me a little bit more about how you have that holistic system built?

[00:33:01] AE: Yeah. So like we have snippets online that actually shows you exactly what those transformations look like for every single data source you just mentioned. But for example, let's take Zendesk. You would transform your data into these – I think it's called activity. So there might be an activity called submitted ticket where a customer submitted a ticket at a time. There is customer commented on a ticket. A customer closed a ticket. A customer submitted satisfaction rating. So now you have these four activities and you can reuse those activities with the rest of the system any way you want. Then you have your Salesforce. And a customer might become a lead. A customer might have a conversation. A customer might go into negotiation. And now with these – Now you have these additional three building blocks and it's each – Like the customer goes into a negotiation. You have the timestamp. You have like what the goal of the negotiation was as a feature and you might have like a link to the Salesforce object, for example.

Now you can answer really a lot of questions. First of all, do customers who submit a ticket before negotiation actually end up canceling the negotiation? Or does actually submitting a ticket before the negotiation within a day affect how likely you are to go from starting a negotiation to a closed one? And those are questions that bridge systems very easily, but still maintaining that single relationship between customer time and activity even though I just bridged two systems. Does that kind of make sense?

[00:34:38] AE: It does. Yes.

[00:34:40] AE: So you can see that any single data source, what every data source is designed to capture is a customer doing things in-time. Now each data source has its own perspective. From Zendesk, it's from the ticket perspective. All we're doing is denormalizing it back to the customer perspective. When it comes to Salesforce, it's from the opportunity perspective. We just bring that back to the customer perspective. For email, like Sendgrid, it's from an email perspective. Or from Mailchimp, it's a campaign perspective. We just convert all that back to the customer because that data is affecting customers. That is what your customer is seeing. Your customer doesn't know that you're using Sendgrid. Your customer knows that they receive an email. Your customer opens that email. Your customer unsubscribes. And when you're trying to change customer behavior, you need to be asking questions from the customer's perspective. And by putting all this data into a single simple table that is centered around your customer, you're able to ask and answer these questions very, very fast and keep it very accurate, because it's always from the customer perspective.

And we can talk a little bit more about other benefits of this structure, which is like speed, simplicity, standardization, single source of truth. Like an actual single source of truth, because it's one table, consistency, dependencies. You get a lot of other benefits out of this simple single table approach to a data system. But the biggest thing that I think everyone who's listening here, and you are, is can this really work? Can you really – Like did you guys really take all the complexity of data analysis and make it into a single table? And can you really answer any question? And like I can give you hundreds of examples, and I think what I always tell people is try it. Or email me any question you have, and if I can't send you an example of

how – Show you how you would answer that in Narrator, I'll give you a Narrator for free. Like that's like my promise to the world. So like this is the power of customer time and action and why we are so excited about this new world that we're doing it. But we understand that this is super different. Like for all of time we've always thought about like building dimensional models and building these fact tables and creating tables that we can plot. Now we're actually creating structures that help us answer questions.

[00:37:03] JM: Very interesting. Can we dive a little bit deeper on what you're working on today? What are the core challenges you're faced with today?

[00:37:14] AE: Yeah. So today, really, we're trying to grow our library and make it as intuitive for people to create their own templated analyses the same way we do it. So we built a system to standardize your data. Most of our customers really, really love it and they find it very easy. We built our dataset tool, which allows you to create any table. We have a lot of customers that replace their entire BI layer just on top of that single table and dataset. And then we have our narratives that we're building. And we're trying to really grow our library. We want to make sure that we have enough questions in there that customers can go a year or two with almost every question they have it be available. We think that like that is the competitive advantage that companies like Amazon and Netflix can afford to hire experts to really fine-tune their systems and optimize and smaller startups cannot. So creating these like really valuable high-quality analyses and making them available for any company at a very reasonable price I think just puts everybody on the same playing field.

So right now we're really focused on growing our library and making sure that creating our analyses is as intuitive and powerful as it needs to be. So like, for example, recently we wanted to make sure that we can actually – An analysis can understand if your company's data starts like massively changing its behavior. So your conversion rate was moving at a steady pace and then like it started spiking. So we had to kind of create an algorithm that helps interpret that so that we can write an analysis that says, “Oh, I noticed that like you were steady moving and then it exploded.” And we want to make sure that that analysis is built into our storytelling tool, which is called Narratives, in a way that any customer can use it, and it's like relatively invisible.

And we want to make sure that we have all those features in place so that when we create our store and anyone can create any analysis they have limitless power to really go and really show their creativity. So that's really our focus right now is making that narrative creation and templatization process as seamless and powerful as it needs to be to get more enough people on board with creating those narratives.

[00:39:49] JM: So the narratives that get delivered, they're basically suggesting action items for these companies, like put more money into marketing or get more sales reps or get back to these customers more quickly, something like that. How often do they get acted upon? I mean you don't want this to turn into kind of like alert fatigue where it's just you get these notifications, you get these narratives but you don't do anything about them.

[00:40:14] AE: Yeah. We actually talk about with the team a little bit. So to build a new analysis, if you have built it before, takes us about a day. To get an analysis written in a way that customers take action takes two weeks. Like we often just get the answer really quickly very fast in Narrator. Usually a day, we know the answer that we're trying to communicate. And we spend about two weeks thinking through how do we convince someone to take ownership and act? Narrator doesn't measure success by delivering analyses. That is not our goal. We measure our success based on our customers taking the actions that we recommend and is that action giving the positive response?

So we see an actually extremely high rate. I'm trying to think of examples where customers ignored our recommendation and like I can't think of one on top of my head. Most customers actually end up taking the recommendation. And the reason why is that we don't give them fun facts. We don't tell them like, “Oh, just decrease this number.” We tell you a story and you follow along to understand how we made that decision and we make sure that it's done in a way that you are believing and you trust it and it makes sense and it's what you would do and it's intuitive and it's logical and it gets you to reach the same conclusion that we end up recommending at the end.

So most of our customers really feel like they are making that decision very informed. And I think that's what separates us from a lot of the AI and machine learning tools that exist where they're like, “Ah! Women convert better than men.” And the person goes, “Cool!” And no one does anything with it. We are taking our time to like tell that story to make sure that our customers are fully understand the nuance and the detail of what's happening in their system and feel like they are making that decision and they have ownership and trust and have confidence in the actions they're going to take. And I think that's the future of data analysis. So much of data analysis goes unused because no one acts upon it. And it's really hard. Like the numbers I'm not telling you are really honest. We schedule in one day to do the analysis and two weeks to figure out how to get someone to act upon that analysis. What is the things you have to say? What are the steps you want to go through? What is the story you're telling to get someone to believe you? And that's I think what makes people really fall in love with Narrator, is that they're getting a story that they can believe. Not a fun fact.

[00:42:58] JM: I'd like to zoom out and talk a little bit about data science and data infrastructure with you. You're coming from a perspective where you're mostly thinking of data infrastructure as a big data warehouse for the typical customer that you're working with. Of course, modern data infrastructure is a mess of HDFS and Hadoop and Spark and Presto and all kinds of things going on under the covers, AWS infrastructure and Redshift and Snowflake of course. Do you have any predictions for where this goes or are there any simplifying pressures being put on data infrastructure or does it just continue to get more complicated?

[00:43:50] AE: So I think that if you try to build an analytics layer that is your production layer, you're going to hate your life. So i think the future of the world is that product or production engineering are going to continue to use more and more tools that are perfect for their specific problem. So like graph databases, versus Postgres, versus Hadoop. Whatever you need to use based on your data and your structure you're going to use it to kind of keep your application running.

I think you're going to see a lot of customers have warehouses much, much sooner. And the reason is because SaaS is growing really fast. Unlike when we were a startup of like

three-months-old, we already had like three, four data sources. We had our internal database and like our email client and our ticketing system and like our marketing website. Like at that point, that data needs to live somewhere. And my prediction is that you're going to see a lot more people move to having a warehouse much, much sooner. And doing a lot more of their work in their warehouse and pushing that data back to their product. I think that's like where the world is really going, because warehouses are so fast and they're mostly cloud-based now. Like we're an entire company on top of a cloud warehouse. So you're going to see warehouses continue to be like a place where people are comfortably putting their data inside and not trying to like build bigger systems. You're going to see a lot of small tiny systems and all that data is dumped into a warehouse that's kind of like a master brain, and that does all the processing it needs and sends that data out. And I see that infrastructure growing to be much, much more popular.

We actually use that infrastructure in – We were having this problem in Narrator early on where we wanted to send very, very personalized emails to our customers who were onboarding. So we realized that we want to send a different email if someone viewed the doc site or if they created their first activity. Or based on all these different behaviors, we wanted to create a different email. And we knew we would change our mind. So we don't want to build that into code and have all these triggers in our code because we would change our mind like 30 times before we like it. So what we ended up doing was – And to give a note that all the data that we wanted to use for the emails were in all these different systems. And instead what we did was we ended up adding a webhook feature into Narrator so we would take all that data from all these different places and it would go into the warehouse. It'll get structured in Narrator. And we create these data sets of the customers we wanted to send a very specific email to and we would webhook that data into a Lambda function that would just send that specific email. And it ended up allowing us to kind of make a lot more changes to the email and really have this product that was really powered by a warehouse. And it saved us like tons of engineering time and kept our production engineering systems very simple and very localized. And that's kind of my prediction for the future of a lot of this infrastructure is simple localized production systems that use the technologies that they need. A warehouse very early on where everything is doing,

and the warehouse does kind of like the brain and they send the data and command actions as needed.

[00:47:17] JM: Interesting. Interesting thesis.

[00:47:20] AE: If you're familiar with the Kiva System in Amazon.

[00:47:24] JM: Yeah, sure. The robots?

[00:47:27] AE: Yeah. So they were actually built by a woman named Hadas Kress-Grazit. She was a Cornell research. I did research in AI with her and Mark Campbell. And the idea of that was the same thing. It was instead of having like localized robots make decisions. What you had were these very simple robots. I knew how to do a couple things. And they would just send their location to this master computer that would relocate and readjust and redistribute based on demand and what they needed to do. And it would then send information once they made the decisions back to each of these robots to do it. It's like a central unit. And I think that like most complex systems end up going and creating this like central node that has all the information that distributes. Just kind of like how almost all systems like from military to robotics ends up doing.

And right now with data and with production systems, you have a lot of async communication and like sockets and like notifications and every system is talking to another system and every system is responsible for making small decisions. And I would say that that systematization of creating a central unit, you'll probably start seeing that in software and data in the next one or two years.

[00:48:46] JM: Awesome. Well, anything else to add, Ahmed?

[00:48:49] AE: Well, I have to do a plug for anyone who's here. If you guys – If anybody wants or is excited about what we're doing, we know it's different, we know some of it sounds unbelievable, but we're always willing to take the time to show you all these things in action

and really, really try it is what I tell people. And if it works, then you'll be one of the believers who really loves our approach and creating that community where the future of data is shared.

[00:49:17] JM: Okay. Well, thanks for coming on the show. It's been great talking.

[00:49:20] AE: Thank you too.

[END]