SDS PODCAST

EPISODE 891:

CONVERSATIONAL AI IS OVERHAULING DATA ANALYTICS, WITH

MARTIN BRUNTHALER

Jon Krohn:	00:00	This is episode number 891 with Martin Brunthaler, co-founder and CTO of Adverity. Today's episode is brought to you by Trainium2, the latest AI chip from AWS and by the Dell AI Factory with NVIDIA.
	00:21	Welcome to the SuperDataScience Podcast, the most listened to podcast in the data science industry. Each week we bring you fun and inspiring people and ideas, exploring the cutting edge of machine learning, AI, and related technologies that are transforming our world for the better. I'm your host, Jon Krohn. Thanks for joining me today. Now let's make the complex simple.
	00:54	Welcome back to the SuperDataScience Podcast. Today, we've got an interesting one for you on how generative and agentic AI are transforming data analytics.
	01:02	Our guest for this quest is Martin Brunthaler, who is CTO of Adverity, an Austrian data analytics platform he co-founded a decade ago, and that has since raised over $160 million in venture capital. Before Adverity, Martin was co-founder and CTO at two other European tech startups giving him over 20 years of combined experience in starting, scaling, and exiting companies across multiple industries, including eCommerce and mobile. He holds an engineering diploma from the Salzburg University of Applied Sciences in Austria.
	01:37	Today's episode should be of interest to just about anyone who'd be interested in this podcast because it touches on data analytics, transforming user experiences with modern AI capabilities, and growing tech businesses. In today's episode, Martin details how a childhood fascination with computer programming evolved into founding a globally leading platform for marketing data analytics, what data democratization really means, and how the traditional dashboard-based approach to data reporting is failing businesses, why data analysts are

		spending too much time on busy work instead of delivering business value, how conversational AI is overhauling how data insights are gleaned for hands-on data practitioners and business users alike.
	02:15	Finally, he provides his no-nonsense tips for tech startup success. All right. You ready for this insightful episode? Let's go. Martin, welcome to the SuperDataScience Podcast. It's a treat to have you on the show. Where are you joining us from today?
Martin B.:	02:36	Hey, pleasure to meet you. I'm based in Vienna, in Austria, in the European Alps.
Jon Krohn:	02:45	I think, if I remember correctly, Vienna has for several years in a row now won The Economist Magazine's number one spot for most livable city in the world.
Martin B.:	02:56	Yeah, actually, it is a very nice place. I have to agree. I'm personally coming from the countryside and I had in the early 2000s a hard time moving into the big city, if you will. I think that's the same across the world, but after a while, I really got settled here and I love it quite a bit. Yeah, we are here. There's also an award for the most grumpiest population or citizen, which Vienna I think has also won several times.
Jon Krohn:	03:26	Oh really? I lived in Singapore for 12 months and I think it's the most miserable nation in the world, which yeah, it's funny. Expectations management, I guess.
Martin B.:	03:41	Exactly, It's probably good marketing too.
Jon Krohn:	03:45	The meanest cities?
Martin B.:	03:47	I think looking for the perspective from city marketing, it could be an asset. People are proud of it as well.

Jon Krohn:	03:55	Nice. So, beyond moving from the countryside to Vienna, to the big city, Martin, tell us a bit about your journey in technology. So, how did you get started in it? What led you to co-found your company, Adverity?
Martin B.:	04:09	Sure. So, I got into computing at a, I'd say, fairly young age. I think eight or nine years old, I got on Amiga computer, if you will. So, I'm not that old to have a C-64, but still a very basic computer where I got familiar with basic programming, which is I had fun creating those mini games and typing stuff into the computer from magazines and the activity. In school, I picked up a lot of networking, managing the school network, and we did have quite some fun with the computing equipment there, some hacking and understanding how this stuff works. It was pretty old equipment at that point in time and no internet access. There's been some popular computing magazines that we had and books that we had to read and get into working with this stuff.
	05:10	Then really I moved to during school and afterwards, I did some integrated systems programming in school, which was a telecommunications degree, so a lot of signaling and mathematics attached to that. But at the same time, I joined a company that was working very closely with mobile network operators. So, we were basically creating apps for mobile network operators. There used to be a technology called WAP, which were microsites that could be delivered through SMS. From there, went into creating pretty large scale messaging systems. So, we, for example, created the software that powered the American Idol SMS voting.
	06:00	Also here in Austria, we powered a couple of radio and TV stations with our technology and we then exited this company in 2006, spent some time in a bigger corporation, an American corporation. At that point in time, they had a lot of activity and I think that the slogan

at that point in time was being part of every transaction. That was fun for two years, but still I wanted to do something like creating stuff, which is why we created an incubation type company. Experimented with some stuff, after which I did create a meta price comparison engine, if you will. The whole browser add-on ecosystem was sprawling at that time. So, we created an extension you could compare prices in real time and suggest alternatives.

06:58 Yeah, so Alex, my co-founder, and Andreas, we knew each other for a while and we had an interesting problem to solve, which was reporting on TV commercials. So, we did have a very hard time to get for a different project, hands-on data in regards to the success metrics for our TV commercials. So, we figured there must be something better than taking a CSV file and creating an ad hoc report and putting it into a PowerPoint presentation that we received like six weeks after the commercial's aired. So, this is really how we got started with something automating reporting. In 2015, we decided to actually incorporate this as a standalone company that's Adverity today.

Jon Krohn: 07:48 Adverity. So, it's got advertising in it, I guess, reflecting your advertising roots and then the verity... It's interesting. I don't know if that's just something that you chose because it sounds nice, but that's related to the Latin root for truth, right?

Martin Brunthal...: 08:03 Yeah, exactly. So, that's connected to that. At least it was the thinking behind that. It evolved as a brand, but yeah, .com domains, even at that point in time, they are very rare. So, yeah, we had to.

Jon Krohn: 08:17 That's a nice one. Yeah, the truth behind data, the truth behind advertising data in particular, it makes a lot of

sense to me. You were about to say something and I just interrupted you.

Martin B.: 08:26 No, that's good. That's exactly like the story in very short summary.

Jon Krohn: 08:35 Martin, how is your English so good? I mean, your English is amazing. I wonder how that happens. Everything that you described about your story, you're in the German countryside tinkering with old networking components. At what point do you become completely fluent in this incredibly technical way in English? How does that just happen?

Martin B.: 09:01 To be honest, I don't know. Also, to be fair, my English training, if you will, is not up to speed. So, I spent half a year in England during my studies as in an abroad. I'm not sure if that's still possible, but in Europe, you had a program where you could take or join another university abroad and we had a cooperation with Staffordshire University in UK. That's where I've spent some time. I tend to read a lot of English material obviously. I try to watch movies in original language. I personally wouldn't consider I'm very fluent, but it works for my daily business.

Jon Krohn: 09:48 I mean, I would say you're extremely fluent. It's so embarrassing for so many people from North America. So, many of us are in this monoculture and even I studied German for 12 years on Saturdays. Een before we started recording, I was embarrassed to try to do a little bit. I can order a schnitzel and a beer off a menu and make my way to my destination on the train when I'm traveling in a German-speaking country. But to be able to describe anything, 1% of what you just did in terms of your career and what you're doing at your company, if I tried to do that in German, there's no way.

Martin B.:	10:27	Thank you. Yeah. One fun fact is though my German accent is very strong, so some coworkers, when I talk German, I have to talk a little slower.
Jon Krohn:	10:40	So you're easier to understand in English than in German.
Martin B.:	10:44	Probably, at least in our company.
Jon Krohn:	10:46	That's funny. So, that's something to do with also your regional accent from where you came from.
Martin B.:	10:51	Yeah, I'd say so. I never spoke to the book, if you will. There's the concept of high German, if you will, or proper German.
Jon Krohn:	11:03	Hochdeutsch.
Martin B.:	11:06	Exactly, yeah. So, I don't tend to do that and I don't tend to do that with our kids as well, which helps them understand me in German as well. So, yeah.
Jon Krohn:	11:17	So it's like it would be in England, they would call it the Queen's English.
Martin B.:	11:23	Yeah, exactly. It's some parallel to that. Let's not get into that too much. There is a conflict in between the type of German you speak in Austria and Germany.
Jon Krohn:	11:34	Habsburg Empire, Ottoman Empire. I don't know. I'm making stuff up. I have no idea.
Martin B.:	11:40	That's fine.
Jon Krohn:	11:40	I'm just throwing out random nouns related to Austrian history that I know. Okay, nice. Let's get back to business, Adverity. So, when you founded Adverity back in 2015, the intention was to have more finesse around analyzing TV ad data, having something be I guess closer

to real time, something more advanced than a CSV file, but now Adverity has grown into over the past decade. Congrats, it's your 10-year anniversary.

Martin B.: 12:11 Thank you.

Jon Krohn: 12:12 It's grown into a leading data analytics platform. So, how did the original mission evolve over the years into the broader offering that you have today? In next conversation, I guess you can get into this a little bit now, but we are going to be focusing a lot on the conversational and agentic AI elements. So, you can get into those now or we can save those for the next question if you want.

Martin B.: 12:38 Yeah, sure. I mean the way we started was really about the workflow and reviewing how agencies initially at that point in time created reports for their customers. A lot of the work behind that used to be going into all the ad systems that they took care of, copying the data into a type of spreadsheet, using this to create a visualization, embedding this in a PowerPoint and creating a PDF to send around to clients. So, that was very roughly spoken, the workflow that we attacked. Going back to the root, so we created a solution where it became very simple from our perspective, basically a verticalized BI solution to deal with the automated reporting.

13:25 Our vision was that agencies would then use this to build dashboards that they can share with their clients and get near real-time reporting. That's also how we expanded it very quickly in terms of connectivity. So, initially, literally, you had to upload the CSV as well, but then we built out our data integration capabilities, which is today the core of our business really. So, we do support quite a wide range of integrations with various ad systems and also adjacent systems, measurement, finance systems, CRM systems, shop systems.

	13:59	All those data points that you need in order to do proper reporting is what we built out over time. Like I said, I think over the course of the last 10 years, I think reporting changed also how it's done within many of our customer or many of our customers. Dashboards are not that relevant any more than they used to be when we started. So, yeah, that probably could bring us to data conversations as well or how this might be changing.
Jon Krohn:	14:33	Yeah, for sure. Really quickly before that, so it sounds like Adverity still has some specialization in working with marketing data, but you mentioned other kinds of data there as well, customer relationship management data, financial data. Is that all with a view to enhancing a company's perspective on marketing or would you say that Adverity today could be used as a more general purpose enterprise analytics platform?
Martin B.:	14:59	Yeah, the platform itself could absolutely be used for other verticals as well. I think for us as a company, from a go-to-market perspective, and also when you look at the connectivity portfolio, there's obviously a bias towards marketing and we as a company in our DNA have a lot of marketing knowledge baked in. So, all our implementation consultants, professional services staff, they know marketing inside out. So, I think as a company, even though the technology could be applied to other fields as well, you need to do more than just extending your connectivity portfolio. You need to train people and stuff accordingly to be able to work in other fields. But that said, I think there is a marketing adjacent teams that are very likely to adopt this solution as well.
Jon Krohn:	15:45	This episode of SuperDataScience is brought to you by AWS Trainium2, the latest-generation AI chip from AWS. AWS Trainium2 instances deliver 20.8 petaflops of compute, while the new Trainium2 UltraServers combine 64 chips to achieve over 83 petaflops in a single node -

purpose-built for today's largest AI models. These instances offer 30-40% better price performance relative to GPU alternatives. That's why companies across the spectrum - from giants like Anthropic and Databricks to cutting-edge startups like Poolside - are choosing Trainium2 to power their next generation of AI workloads. Learn how AWS Trainium2 can transform your AI workloads through the links in our show notes. All right, now back to our show.

00:16:36 All right. So, now let's get into the more advanced topics that I just said that we'd put a pin in a few moments ago in terms of generative AI and agentic systems and how those fall into a platform like Adverity's. So, Adverity is currently launching its data conversations product, which lets users query data in natural language and even get proactive recommendations. So, it sounds to me like the talking in natural language, this is leveraging the generative algorithms like the GPT-powered experience that we expect in our tools increasingly today. Then the proactive recommendations, that sounds like an agentic system to me. So, do you want to talk about these particular features and how you're thinking about generative and agentic AI at Adverity?

Martin B.: 00:17:31 I think when you look at the proposition of data conversations, which we are about to launch, there is one overarching theme which is data democratization. This is not a new term or a new topic, but it used to be that data catalogs and BI tools and dashboards have been considered to democratize data access somehow. It turns out that those dashboards mostly are created once, not reviewed, are not up-to-date, and always are getting challenged by the user and then also don't give you a capability to drill down and ask more questions to understand what's going on.

00:18:15 So, usually you would see people looking at a dashboard for figuring out what this means, comparing it in their mind with yesterday's data, challenge that, go back into the source system and see, "Is this actually the data? Can I trust this data?" and all those issues. So, I think there's an evolutionary step happening here with technology like generative AI that allows you to actually use human language to ask questions about the data.

00:18:46 In order to get there, you need a lot of things done right before you can do that, but once you have a solid data foundation created, you can actually take this to quite a solid use case, which we have done so far. So, I think one of the strengths that we have is a deep understanding of the lineage source of data, meaning of data. So, we manage a data catalog. We call that data dictionary that has a clear connection to source and also a deep description of what's going on, what meaning is behind it given attribute. That helps us ground our system quite a bit to give propensity.

Jon Krohn: 00:19:28 Nice. Yeah. So, let's talk a bit about... We have listeners at home. A lot of our listeners are either hands-on data science practitioners like machine learning engineers, AI engineers, data scientists themselves, or people who are interested in building products or companies that leverage generative AI. What are the kinds of lessons that you've learned in implementing a product like data conversations at Adverity? What do you need to do? What are all the things you need to line up in advance of bringing in a large language model and having conversations work effectively with data? You talked a moment ago about the issues that you typically see without this conversation in place where people have a dashboard and it's not exactly the information you needed.

00:20:22 It's too fixed in its outputs, and so then people end up going and digging under the covers into the raw source data to try to really find answers, which adds strain onto the data analyst team. So, I get all of the advantages of being able to have a conversation with your data, but what are the things that you at Adverity, that are listeners if they want to be making this similar transition, what do they need to get right in order for that conversational aspect to work out?

Martin B.: 00:20:50 So I think one really critical piece is the quality of the data underneath. So, there's many aspects of data quality if you will. Also from an academic perspective, you can list those out, but from a more practical perspective, you need a complete data set that is also very well aligned with all the various sources that you have. So, harmonization plays a role in this as well. We built up actually a data quality component in our platform that helps you monitor all those issues that you can have in your data. There's specific monitors for data quality in marketing.

00:21:30 There's a concept called naming conventions, for example, for campaign names that we can monitor and act on in an intelligent manner. But there's also simple things like if you onboard a generic source from a database or from a REST API, all the data types need to be aligned, data formats need to be aligned. You want all your data to be harmonized in UTC, for example. You need to clean up some stuff. This is also why there's some transformations going on usually either by splitting up, combining various sources, and all those things, but I think it's very critical to get the quality right. You need to be alerted. If something's going wrong, you want to prevent, not saying dirty, but problematic data sets to hit your production environment. I think we can help in this discipline quite a bit.

Jon Krohn: 00:22:21 You could help in the discipline by having these data quality reportings built into the platform.

Martin B.: 00:22:27 Yeah, but also the multi-layer approach to this. So, we keep always a raw data set that can then be used as a starting point to reiterate on transformations for example. So, you can always go back to the previous state and improve your transformations. There's also obviously today an AI system helping you to compose those transformations. This is specifically always very useful for those type of generic sources, but it's a simplified data of wrangling exercise, if you will. Then once you're satisfied with that, there's a component that helps you monitor the quality as it flows through the system. There's an anomaly detection and all the things that you want to monitor.

Jon Krohn: 00:23:11 Right, right. Yeah. So, built-in anomaly detection would be key to this working out. There's a huge amount of breadth of capabilities that you could potentially get from a conversational interface. When you are designing a conversational product, how do you figure out, okay, this is the range of things that we're going to support or not support, and then how do you select the right large language model for that breadth of features that you decide to support? Yeah, let's start there. I have more follow on questions from that, but I feel like that's a good starting point.

Martin B.: 00:23:52 Yeah, I think it's useful and maybe one thing to add to the previous question in terms of quality, like I already said, the data dictionary descriptions, understanding of lineage is very critical as well. This goes also into the design of our conversations interface and how people can interact with that. We iterate very quickly. So, we are going through, I'd say, a pretty fast-paced development cycle with adding features every week. We have a dedicated team taking care of benchmarking and

analyzing the quality of responses. So, using frameworks to monitor that and the data science team is having a continuous test on...

00:24:44 We have a predefined set of responses that we expect from our questions and we can monitor on those and improve and test models as we go. To be fair, at the moment, we're committed to one model, but there is also the plan to use different models for different aspects of our capability. So, for example, we could use a different model to compile our SQL query, a different model to do the preflight qualification of a question, a different model to do the actual conversation. So, yeah, that's also possible.

Jon Krohn: 00:25:16 Nice. Obviously, the questions that I asked you were tricky because I'm trying to get at what are the things that people need to be doing in order to build these kinds of conversational interfaces like you did, but obviously, there's proprietary things involved.

Martin B.: 00:25:32 Yeah, I think there's no trade secreting in building, if you will. A lot of LLMs and the type of APIs they offer are similar in regards to their capabilities and you see all models reaching the same capability and basically the leaderboards change. Just every other month you'd have another leader, but everyone's catching up to the same state of quality, if you will. I think where it then boils down to is how you put the components together to create a compelling and exciting use case on top of that. I think in terms of how this works from a technical perspective, it's pretty straightforward.

00:26:22 You can qualify a user input into a type of question, select the model that you want to run with, basically feed it with a system prompt and additional information about the model, which is very critical to get the answer right, use this to create a SQL query, verify it's actually a valid query that can be executed, fire the query, use the data to

run some basic analysis and create a decent nice answer for the user. For us, the use case then circles a lot around the table that we generate from that response.

00:26:57 Because what our approach to this is, first of all, in terms of democratization, we are targeting two sides of the business. One of which is IT, and the other one is in the business user. Both have a requirement to access data. So, rather than going through a full chain of various teams, so it used to be that you had to create a ticket to get access to a set. The data set would then be prepared within two weeks and put onto a Snowflake table or whatever today, a Snowflake table. It used to be something entirely different. With this, you can actually run the query, create a table in near real time available for your further analysis and that's exciting for us.

Jon Krohn: 00:27:41 That is really cool. All right. So, you mentioned there's something interesting which is that a big part of these data conversations, whether your user is a relatively technical user, like somebody in IT or maybe data analytics or whether they're a business user who probably has no experience writing code. In either case, this data conversation is able to serve that user, but something that's interesting about what you were saying is this idea of it generating a table of results. So, that's something interesting.

00:28:11 So, something that I probably should have asked from the very beginning is what is a user's experience like when they have this data conversation? So maybe give us the user story of what your experience is like getting access to data in the Adverity platform before data conversations as well as after.

Martin B.: 00:28:36 To be honest, I can't speak of all use cases, but the predominant use case is to extract data from sources. Say for example, you want to run marketing reporting across

Google, Facebook, you name it. So, basically you extract data from those source systems, prep it in a way that is useful for reporting, and put it into a table. Some differences are to be made. Some put it into various tables or many tables. Some already adopt a one big table scheme. So, you end up with a database, BigQuery, Snowflake, you name it. So, we support quite a wide range of databases and have a full snapshot of the source data in your database. I think most will then connect us with a DBT model to do some further transformation, probably create subsets of data for different kinds of reporting.

00:29:34 I can speak of the models itself. So, that's up to the customer. Then in turn, this is usually connected with a BI system. So, mostly the things that the industry standard would be Power BI for many organizations. There's still Looker, Looker Studio, Tableau to name a few. That's basically where users would today access the data. The end consumer would look at a BI tool to run a query, to create a dashboard, or look at a pre-canned dashboard that someone else in the team has created.

Jon Krohn: 00:30:03 Can I ask a clarifying question? So all of these kinds of technologies that you just mentioned, like your data source being Google or Facebook for marketing data, taking those data from those various sources from a client, putting that into a Google BigQuery to allow you to have SQL-like queries over these vast amounts of data happen rapidly. You could also alternatively be using Snowflake for that. Then you mentioned DBT, downstream for data transformations, and then finally some tool like Power BI, Looker, or Tableau in order to be able to see, to visualize the data, to be able to get summary metrics that look nice. All that happens within Adverity. So, that is very interesting to me. Yeah, so confirm that that's the case. Yeah.

Martin B.: 00:30:58 No, I am just mentioning one of the cases and here it becomes interesting because you would look at this from a modern data stack perspective obviously as in a highly composable or a very modular stack of components that create a data solution for you. I think what you get with our solution, you don't have to use those tools.

Jon Krohn: 00:31:23 I see, I see, I see. So, that's what was confusing to me because I was like, "So all of these different technologies are supported within Adverity." So you're saying that basically that is a common workflow. Google, BigQuery, Snowflake, DBT into Power BI, Looker, Tableau, that's a common thing that a client of yours would typically be doing before they start working with Adverity. But then with Adverity, you get one platform where you do all of those things together and then you don't have to worry about dependencies changing between platforms. It relieves a huge number of operational headaches downstream for your clients.

Martin B.: 00:32:06 Absolutely. So, you don't have to use DBT with us. We can execute DBT on the customer's behalf as well. So, you don't have to manage an orchestrator or run some other complex machinery to run DBT. We can do this on your behalf as well. But I think one key point is we already provide a harmonized, aligned, and properly structured data set for any purpose in marketing. So, if you go with a default and create a database that receives data from us, it's ready for analysis in your BI tool. So, you don't have to do all the in-between work. Our transformation capabilities are pretty powerful as well, like I initially said. So, if there's some need to change the shape of data before it lands in your database, that's also a possibility that our customers have.

00:33:03 So, I guess it saves quite a bit of effort to create the data set that you require. Also, you can connect it natively with any BI solution that works through an old data

interface. We have a concept called data shares where customers can create a subset of data that then can be shared with an external application as well. So, a lot of what we do really is we have baked in governance and enterprise capabilities that make it very powerful to use as compared to something that you have to build on your own. Because what many people don't see is that there is a lot of work in creating and maintaining a full modular modern data stack.

00:33:54 Probably also going back into the overall topic of AI, I think people need to start concentrating on the business value and strategy on top of all this, if you will, I'm not saying magic, but this amazing machinery that you don't actually want to build a modern data stack with a lot of components and orchestrating a huge number of SaaS tools. I think that's certainly something that I'm very excited about.

Jon Krohn: 00:34:21 This episode of SuperDataScience is brought to you by the Dell AI Factory with NVIDIA, delivering a comprehensive portfolio of AI technologies, validated and turnkey solutions, with expert services to help you achieve AI outcomes faster. Extend your enterprise with AI and GenAI at scale, powered by the broad Dell portfolio of AI infrastructure and services with NVIDIA industry-leading accelerated computing, it's a full stack that includes GPUs and networking; as well as NVIDIA AI Enterprise software, NVIDIA Inference Microservices, models and agent blueprints. Visit www.Dell.com/superdatascience to learn more. That's Dell.com/superdatascience.

00:35:07 Right, exactly. So, what you're saying is that instead of investing your engineering time and money in your organization on rebuilding and maintaining these same kinds of complex data stack interactions that all of your other competitors are doing, that it would make a lot of

sense to focus on finding a solution. If you are working with marketing data, then Adverity could be a great solution for you and your data and then your engineering time and effort doesn't need to go into reinventing the wheel yet again, on a similar data stack to all of your competitors.

00:35:46 You can instead be investing on actually saying, okay, the data are going to flow properly. We can export them in various places if we want to and use them in other tools if we need to for some reason, but we can now concentrate on using the data that we have, getting some actionable insights from them and getting an ROI, getting a return on our investment in our engineering team.

Martin B.: 00:36:09 Yeah, exactly. Maybe also what is very interesting about the marketing ecosystem, it's very dynamic. There's lots of changes happening on the source side as well. So, you don't want to keep up with that. If you go with a typically horizontal ETL solution, keeping up with the changes on the source side, it might be a very tedious and problematic. So, if you have someone to rely on to get this data in without you taking care of that, that certainly already saves quite some money.

00:36:40 Maybe going back to the question around the data set itself and how you bridge the gap to other departments, say for example, finance or whatever data sets you have that sit in your organization that you want to connect in terms of reporting, I think you can see our solution, as in the marketing data set magically fills itself into your data warehouse and warehouse environment. We take care of everything there. It's a table that holds this data and you can then connect it with other sources that you have within new organization and still keep it isolated as much as you like.

Jon Krohn: 00:37:14 Perfect. So, now I have a better understanding of what the workflows are like for somebody who is trying to work with marketing data, trying to get actionable insights from marketing data, whether they are a hands-on data practitioner or whether they're a business user. With the data conversations product that Adverity is now released with this generative experience, how does that change everything for a user?

00:37:47 Previously, Adverity has already been making life easier by handling the interaction between all these different kinds of databases, data transformations, outputs, as well as like you mentioned recently a few moments ago on the podcast, keeping up with all the changes that data providers like Google and Facebook marketing data providers make. So, there was clearly already this value to Adverity before. What is the added value and what is the change in the experience like with data conversations?

Martin B.: 00:38:18 Yeah, I think so the most standout difference is that rather than looking for the correct dashboard or correct set of data that you want to look at, you can formulate a natural query and we figure this out for you, rather than looking for a specific data set on your own or really asking a different team to get you to the data that you need in order to answer your business type question. I think that's the first and most basic need that we address with data conversations is accessing data and putting it to good use. Because like I said, the tabular response can be immediately actioned upon by either continue to use it in, say, for example, a Google sheet or materializing this in a Snowflake table or BigQuery table for that matter.

00:39:09 So, you can immediately work with the data set once you're satisfied with the result. But on top of that, there's also business analytics questions that you can ask about the data set. We try to do the best in order to get to a solid

response so you can ask for the best performing campaigns that the system will figure out how to connect the dots and create a useful response for that matter. All the marketing type specialized or subject matter expertise that we have built is also built into the system.

00:39:42 So, that's helping you to get to the response a lot faster. I think the second huge benefit is you get a system that you can actually explore deeper. I mean there's been basic features for that or basically in many BI, I told you, that you can grow down into the data set and get more detailed data. But if you want to branch out into other data sets and understand for example why a trend is happening this way, you can do this with the conversation. You cannot do this with an explore function in say Looker or Adverity for that matter.

Jon Krohn: 00:40:19 Very nice. All right. So, basically you get more depth of response. You can explore data more deeply and it's easier. I guess for any user doesn't even, I was going to say particularly for a business user, but it's easier for anyone to be composing their questions in natural language relative to a SQL query or something like that.

Martin B.: 00:40:40 Yeah, absolutely. Again, going back to the enterprise capabilities, I think what's very critical as well is to give IT a tool that they can still manage on an audit rather than having a wild mix of SaaS tool connected together that don't necessarily give you the level of governance that you need in order to comply with a lot of regulations that we face today. So, yeah, that's another aspect of this.

Jon Krohn: 00:41:08 Nice. All right. So, for our core audience, people out there who are data analysts, data scientists, do you think that they should be worried about tools like this encroaching on their roles or do automated conversations like this around data actually free up data analysts, data scientists, machine learning engineers to be tackling

more interesting problems? I'm assuming it's the latter. I'd be surprised if you answered any other way.

Martin B.: 00:41:43 No, absolutely. I am totally in this. I have the same understanding. So, this is going to allow a lot of analysts, also data engineers to focus on more strategic stuff, topics that are important for the business and focus on what value you can deliver for your company rather than having to... But the technology allows you to do less busy work, if you will.

Jon Krohn: 00:42:11 Yeah. You're not just getting messages like Slack messages, emails from finance, HR, marketing asking you for, "Oh, can you just do this data poll for me real quick?"

Martin B.: 00:42:22 Exactly.

Jon Krohn: 00:42:22 You're like, "Well, there goes my day."

Martin B.: 00:42:25 Yeah, sure, let's do this. Rather than that, you can point your colleagues also to this solution that helps you focus on the real stuff. I think something that certainly will matter in this role moving forward is there's a lot about how you govern those tools and make sure that the responses are correct. So, having tools to monitor and supervise what's going on in your business, that's critical, but yeah, I don't think that any jobs are at risk. On the contrary, I think the profiles will change and I don't see a major issue. You have to obviously go with the trend in a way, but that has always been the case.

Jon Krohn: 00:43:11 Yeah, yeah. Okay. So, we've talked now a fair bit about generative elements and how being able to have a natural language conversation with your data can allow you to have an easier time of getting insights from your data. You can get deeper into your data. I'd like to talk about the other big buzzword in data science and AI today, which is agentic AI. so I don't know. Would you apply that

term agentic AI to some of the features that you have in the Adverity product that are autonomous?

00:43:45 It seems to me from my understanding of the product that there are processes that are automatically, you talked earlier about data quality, for example, doing anomaly detection. Would you consider that autonomous process that is looking out for issues, maybe flagging opportunities as agentic? If so, would you say that you have agentic features built into the Adverity platform?

Martin B.: 00:44:13 I have a very hard time with the term today because it's been used across the board. When AI entered the scene, it basically meant technology like ChatGPT, but very quickly, many companies adopted the term for non-AI type problems, basically statistical problems or machine learning type capabilities but that still haven't been or didn't make use of any-

Jon Krohn: 00:44:40 Any data model.

Martin B.: 00:44:43 Exactly. The same happened with agentic AI. I think there's certainly a common denominator, but many people would've used this as agency left and right. My understanding of this really is to give ChatGPT or a generative AI type capability access to functions in a way that they can take action on behalf of the user sometime in the future. There's this concept of establishing trust with AI solutions. I think we are not there where people would trust an AI solution to make decisions on their behalf fully.

00:45:24 So, there needs to be some level of control and monitoring baked in, but where they can actually call an API, use the result to even trigger some action without some means of user interaction. That's my take on this at the moment. Some autonomy can be baked into this as well, but for me

at the moment, it's mostly around actions and also interestingly retrieval of information.

Jon Krohn: 00:45:51 Nice. All right. Yeah, so this makes sense to me as a better definition of agentic AI. I guess you could ask any individual to define what agentic AI means to them. I was using it far too broadly just as people use AI to describe any data modeling process where I'm thinking about, okay, if there's any autonomous system. But I think if this is a completely programmatic autonomous system, we probably shouldn't be calling that agentic AI. So, yeah, so I liked your definition there of where it involves probably you have a large language model in the mix in an agentic AI system. Ideally, like you described, it has access to tools that allow it to be taking various kinds of actions. I think that's the key part there. It's basically maybe we could define it, I don't know if you'd agree with this, but an LLM that can take actions.

Martin B.: 00:46:43 Yeah, exactly. So, I think that's a much more realistic definition of this because on the other hand, if you define it too broadly, people have expectations for a system that is not able to behave in a way that they would expect as well. I think it's mostly also now in a phase where we establish standards. So, a couple of weeks ago you had all this hype going on about MCP, slightly another hype going on about A2A in the near future.

00:47:10 So, I think we are at the phase where actually companies are going to set standards in terms of how this might look like in the future. So, having a capability like MCP in place and having a lot of SaaS products, we certainly are going to do this as well, offer MCP capabilities to hook the company's tool of choice, which could be Gemini or OpenAI or whatever agent that the company will go with into your environment. I think that's very valuable and will probably truly help to get some agentic solutions out there.

Jon Krohn: 00:47:47 Nice. Yeah, so you mentioned MCP there casually. So, I did an episode on this recently, episode 884, if people want about seven minutes of detail on model context protocol, MCP, but this is an open source framework from Anthropic that provides a standardization for agents taking actions. So, it sounds like maybe fill us in a bit more on how it sounds like you are actually integrating that into Adverity.

Martin B.: 00:48:18 Yeah, so we will offer capability to connect this. It's in an early stage at the moment. You need cloud desktop in order to connect this in a proper fashion, but this will surely develop into a very good solution where you connect your chat agent with services and let them take action. Like we discussed earlier, it might not necessarily be a gen AI type action. It could be anything from statistics to triggering some action in a remote system or even consuming information from that system. So, in our case for example, getting a listing of data sources connected. If they're all properly connected, create links to reauthorize, really interacting with the system in a natural way. That's what we are building out here.

Jon Krohn: 00:49:09 Very nice. I love that. I mean MCP is one of the hottest topics right now that I come across at any conferences that people are talking about. So, it's cool that you're integrating that at Adverity as well. It can be tricky to see into the future and make predictions when things like MCP come out of nowhere and all of a sudden become a standard and we can anticipate that that thing, as well as AI advancements, LLMs continuing to, as you say, ratchet up on the leaderboard time over time and all the big players at the frontier of AI capabilities having more or less fungible APIs for a lot of different tasks.

00:49:52 So, the point of me saying all that is that it is tricky to be able to see into the future, but given your role as the CTO of a company that's taking technology and turning it into

more streamlined experiences for users, what are your predictions for how things will continue to evolve five years from now? How will somebody, whether they're a hands-on data analyst, data scientist, or a business user extracting data from their systems, their providers, what will that experience maybe be like in five years relative to today?

Martin B.: 00:50:36 Like you said, it's very hard to predict five years out. Being in this business for a while, I know that things can change very quickly, but I think overall a lot of focus will shift to strategy rather than having to do... We quickly talked about or touched on busy work, but there's lots of workflows that currently need some manual interaction that are fully automatable in a way. People will then have a chance to focus on the business and outcome of data and how it's going to be used rather than the mechanics behind that. So, that's certainly something that we will change quite a bit.

00:51:16 I think also we didn't talk about that, we focused quite a bit on AI, but there's also an interesting move in our industry as we forgot to storage and disconnecting compute from storage, which is a shift that is happening underneath what we do today. So, looking out into the future, the mobility and picking compute connected to data, if that makes sense, will certainly be something that will change the system quite a bit as well or the way we work today quite a bit. So, having a raw repository that sits in an accessible object storage environment in a standardized manner, because historically, that's not been standardized.

00:51:58 Today, the industry settled on Iceberg as a format for that and being able to connect this with various query engines depending on purpose. Likely in the future also with a generative AI capability, I think that's very valuable to businesses to pick the right solution for the right case.

So, that's certainly going to change as well. But five years, it's a long, long timeframe. We are talking maybe half year, year timeframe here, but yeah, I think mostly strategy and focus on the why rather than the how is probably something that will change quite a bit in the next five years.

Jon Krohn: 00:52:43 Right, right, right. So, back to your point earlier in the conversation about tools like Adverity allowing you to have your engineering team, your IT team focused less on integrating systems and keeping everything up to date and more on actually extracting business value.

Martin B.: 00:53:00 Yeah, I'd say that and less busy work, less repetitive task if you will.

Jon Krohn: 00:53:06 Less of the find me these data real quick messages.

Martin B.: 00:53:12 Real quick, I need it now.

Jon Krohn: 00:53:14 Exactly. Nice. So, we've now looked into the future a little bit, but another question that just occurred to me is given your experience now as a serial entrepreneur, you've spent a long times at tech companies in senior leadership roles like CTO, what are the kinds of lessons do you have from your past experiences that we can learn from? What are the maybe mistakes that have happened or things that you've learned that you now avoid based on that experience and that our listeners could avoid as well?

Martin B.: 00:53:53 Also a very good question because it depends. So, I think one critical thing is don't overthink stuff. If you want to build something and have something in mind, go build it because thinking about how nice it could be and overthinking architecture, I think we are victim to this or have been victims to this as well. I think you can get the ball rolling very quickly.

	00:54:24	Also today with the capabilities of AI-based coding tools, I'm not saying vibe coding. I'm saying try out various solutions, use what you have. I'm in the camp of use boring tech as well, proven tech that works and compose an exciting use case and keep talking to the customer because that's the most critical impact or most critical feedback that you need. So, from the start, we've been very customer-centric, taking feedback seriously and integrating this with our solution. So, that's probably something that I would always encourage in any company.
Jon Krohn:	00:55:04	Sensible approaches there. Yeah, listen to your customer and don't vibe code, but do take advantage of LLMs for code generation to be able to get a POC stood up more quickly and not just spend all of your time in the planning stages on a potential product or feature.
Martin B.:	00:55:25	Yeah, absolutely.
Jon Krohn:	00:55:26	Cool, nice. Great advice today, Martin. I've really appreciated it and it's great to hear the exciting things that you were doing at Adverity. Before I let you go, I ask all of my guests for a book recommendation.
Martin B.:	00:55:39	Okay, yeah, sure. I recently read a book about how Vienna influenced ideas in the modern world. It's called Vienna: How the City of Ideas Created the Modern World. It's an exciting book showing how a lot of concepts in architecture and many different fields have been created in this city. So, it was an interesting read and I also learned quite a bit from this book.
Jon Krohn:	00:56:08	Fantastic. That is exactly the reading that I wish I had all the time for because that thing about we really stand on the shoulders of giants in terms of what we're doing technologically, linguistically, scientifically. Vienna for sure plays a huge role in the renaissance of ideas that

has led us to a lot of people in the world not having to worry about shelter and nutrition, great child mortality rates and healthcare, all these kinds of things that we enjoyed today. Vienna played a big part in that and I would love to learn more about it.

Martin B.: 00:56:47 Yeah, absolutely. Good recommendation.

Jon Krohn: 00:56:50 Nice. All right, and then Martin, for people who want to have your thoughts after today's episode, how should they follow you or Adverity?

Martin B.: 00:56:58 I do have no social profile. I do have a LinkedIn profile, but I'm not hanging out on any of the popular spaces. An email is always appreciated and I respond to emails. Other than that, LinkedIn certainly works as well. Adverity, visit us on our homepage to get more insight on our product, but personally LinkedIn.

Jon Krohn: 00:57:24 LinkedIn for connecting with you. Then I guess, what's your email then if people want to reach out? That's a very generous offer to make.

Martin B.: 00:57:30 Yeah, it's martin@adverity.com.

Jon Krohn: 00:57:33 There you go. That's pretty easy. Nice. We'll have that in the show notes. All right. Martin, thank you so much for being on the show today. I hope you enjoy the rest of your day over there in Vienna. Yeah, maybe we'll check in on you and the Adverity journey at some time again in the future. Thanks for all your insights today.

Martin B.: 00:57:52 All right. Thanks for having me.

Jon Krohn: 00:58:00 Cool episode with the super successful but also super modest Martin Brunthaler. In today's episode, Martin covered his journey from programming on basic computers as a child to con-founding Adverity, a

marketing data analytics platform that simplifies integrating data from multiple sources and then gleaning actionable insights from those consolidated data. In particular, he talked about the concept of data democratization through Adverity's, new data conversations product, allowing users to query data using natural language rather than relying on fixed dashboards or SQL expertise.

00:58:36 He talked about the importance of data quality, anomaly detection, and proper data descriptions for effective AI powered data conversations, how generative and agentic AI tools are freeing data professionals from routine busy work to focus on strategic value creation and analysis, and his advice for entrepreneurs including don't overthink solutions, use proven technology, start building quickly and always prioritize customer feedback. As always, you can get all the show notes, including the transcript for this episode, the video recording, any materials mentioned on the show, the URLs for Martin's social media profiles, as well as my own at superdatascience.com/891.

00:59:13 Thanks, of course, to everyone on the SuperDataScience Podcast team, our podcast manager, Sonja Brajovic, media editor, Mario Pombo, our partnerships team consisting of Natalie Ziajski and Nathan Daly, our researcher, Serg Masís, our writer, Dr. Zara Karschay, and our founder Kirill Eremenko, who does a ton for this show behind the scenes. Thanks to all of them for producing another insightful episode for us today. For enabling that super team to create this free podcast for you, we are deeply grateful to our sponsors, you. Yes, you can support this show by checking out our sponsor's links, which are in the show notes.

00:59:49 If you yourself would ever like to sponsor an episode, you can find out how you can do that at

jonkrohn.com/podcast. Otherwise, support the show by sharing it on social media, reviewing it on your favorite podcasting platform, or YouTube subscribing if you're not a subscriber already. But most importantly, just keep on tuning in. I'm so grateful to have you listening and I hope I can continue to make episodes you love for years and years to come. Until next time, keep on rocking it out there and I'm looking forward to enjoying another round of the SuperDataScience Podcast with you very soon.