BBBT®
Boulder BI Brain Trust

Leveraging
Business
Intelligence

BBBT Podcast Transcript

About the BBBT

The Boulder Business Intelligence Brain Trust, or BBBT, was founded in 2006 by Claudia Imhoff. Its mission is to leverage business intelligence for industry vendors, for its members, who are independent analysts and experts, and for its subscribers, who are practitioners. To accomplish this mission, the BBBT provides a variety of services, centered around vendor presentations.

For more, see www.bbbt.us

Vendor:	**Diyotta**
Date recorded:	**04-01-2016**
Host:	**Claudia Imhoff**, Founder, BBBT
Guest(s):	**Sanjay Vyas**, Chief Operations Officer, **Ravindra Punuru**, Chief Strategy Officer, and **Jonathan Wu**, Chief Executive Officer
Run time:	**00:13:50**

Claudia: Hello, and welcome to this edition of the Boulder BI Brain Trust, or the BBBT. We're a gathering of international consultants, analysts, and experts in business intelligence, who meet with interesting and innovative BI and analytics companies here in beautiful Boulder, Colorado. We not only get briefed on the latest news and releases, but we share our ideas with the vendor on where the BI industry is going, and help them with their technological directions and marketing messages. I'm Claudia Imhoff and the BBBT podcasts are produced by my company, Intelligent Solutions.

I'm pleased to introduce my guests today. They are Sanjay Vyas and Ravindra Punuru. Sanjay is the co-founder and the chief operations officer and Ravi is the co-founder also and chief strategy officer for Diyotta. Welcome to you, both.

Sanjay: Thank you, Claudia. Glad to be here.

Ravindra: Thank you, Claudia. Nice to be here.

Claudia: All right, Sanjay. Let me start with you because there was a big announcement today. If you don't mind, tell me about it.

Sanjay: Yeah, today we are really excited that we have Jonathan Wu who is joining us as new CEO of the company. We have been working on the technology for the last three years plus, and we have proven the technology and market adoption.

Jonathan had been involved as a board member since last two years and it had been exciting for us to have him come in and lead the company to the next level.

Claudia: I know. I'm so excited so let me introduce Jonathan Wu. Congratulations, Jonathan. What a boon for Diyotta. What was the compelling reason for you to join the company?

Jonathan: Thanks, Claudia. I'm excited to be here. There are several factors that Sanjay pointed out. I've been on the board for about a year and a half, part of that as an advisor for two years for the company,

and I've gotten to know the founders quite well, as well as really taking a look at the technology and then speaking with their clients.

It's a very compelling story when you take a look at it. What I can contribute is my business experience, having grown Base Consulting Group and then merged it with Knightsbridge and then grow it. Then, eventually have liquidity through the acquisition by HP.

Given my knowledge of the space, given my background in terms of the business, I'm excited about jumping back into the operations of business again. I'm excited about this opportunity.

Claudia: It's going to be wonderful. It is so nice to have you back in our space. I have to let you know. All right, let's get into a little bit about what they're doing.

The data integration space is getting pretty crowded. You've mentioned in your presentation this morning that there are five principles of modern data integration and these are the foundation for Diyotta's technology.

What are these and how did they differentiate your technology from let's say, traditional ETL vendors, as well the newer data prep/data integration vendors offerings?

Jonathan: When you think about traditional ETL, you think about the technology that was created some 20 years ago. It was specifically designed for relational database management systems, moving data from one relational database management system to another.

The data is selected, ingested into a centralized server for processing and then moved to the target location for final resting. That architecture doesn't work in today's world. Out of the frustration of having to work with these traditional tools, the founders of Diyotta have created their product. This product is really based upon these five principles that take a look at the new word of data, and how to address the ever-changing and growing presence of information everywhere.

When you take a look at these five principles, number one, take the processing to where the data lives. I think it's critical because the data can reside in, say Hadoop, or it could reside in a relational database management system. It could be in the cloud.

Take a look at where you need to do some of the processing. You may need some minimal processing in one of these platforms before eventually having it moved to another platform, as an example.

These roles right into the second point, which is a fully leverage all platforms based upon what they're designed to do well. If you can do some processing at the source and then move the data to the target and do some processing there, you can do some balancing that takes place. You can't do that with traditional tools.

The other aspect is, and this is point number three, move the data point to point to avoid single-server bottlenecks. If you don't have a server in the middle that has to ingest, you alleviate the bottleneck. That was the design in the architecture of Diyotta.

It has agents at the sources as well as targets and then an orchestrator that sends messages to these agents to do the work and then to move the data from point-to-point. There's not a hub-and-spoke type approach. You don't have the bottlenecks.

That was number three. Number four is managing all the business rules and data, which I think is critical. By having this orchestration taking place, you have the visibility to all the business rules. Where the data resides, what the lineage is, and the metadata history. It's all centrally located where it can be shared and reused. The reusability I think is critical given the environments that we're dealing with now.

Then the fifth component is, make changes using existing rules and logic. What that means is the platforms are constantly changing. The business rules are constantly changing but the ability to evolve rather quickly and be able to address these changes is critical.

That's embedded in the architecture. It doesn't require significant rework that traditional ETL tools have to do or in certain cases, workarounds.

Those are essentially the five principles that help differentiate Diyotta in terms of modern data integration versus traditional data integration.

Claudia: It also differentiates Diyotta from the newer data prep tools. There are some newer data integration tools. I think your architecture is probably one of the things that does differentiate you from all of these other companies.

Ravi, let me bring you into this conversation at this point. What does your data integration platform consist of?

Ravindra: Thanks, Claudia. As Jonathan described about the five principles of modern data integration, we actually baked all these principle into the architecture to solve the problems, currently, we have in data integration space on working with the big data platforms like in Hadoop and MPP platforms like the Teradata, or Netezza, or any other MPP platforms.

Diyotta consists of three components. The first one is the orchestrated component and the metadata repository component and the agents. Orchestrated component is responsible for communicating to agents and sending rules based on the data movement, and transformation rules defined by the user using our browser-based modules.

The core difference or the advantage when you compare with other tools in the market today is: you can define rules for different types of platforms on the same workflow, meaning you can use the same tool, you can push down certain data integration rules on Netezza or Teradata, or you can push down some rules to modern data architectures like Hadoop.

It's gives you the flexibility to implement it to your whole solution, end-to-end solution, and also keeps the metadata end-to-end from the rules, which you implemented.

Claudia: What I like about it is the simplicity, and I'm going to get to you Sanjay in just a moment about the simplicity, of the architecture. You've got this, as you say, orchestration in the middle, you've got then the agents that are actually doing all of the real work on the platforms where the data resides perhaps or doesn't have to be but it can be. You've got this wonderfully distributed environment, with these lightweight agents, being controlled by the central hub. Is that right?

Sanjay: Yeah, exactly. That's correct, Claudia. The agents are very tiny software components which can be installed anywhere in Windows or Linux-based systems. The agents are actually doing the work to extract the data from your source point and load it to your target endpoints, which could be another agent collection or to your target endpoints. You're absolutely correct.

Claudia: Sanjay, let's go talk about the simplicity. You said that you wanted to simplify data integration. How does Diyotta do that?

Sanjay: What we have seen is the paradigm shift in data integration. Gone are the days when these projects use to take like 8 to 10 months. What we have seen with our customers is that they need agile development process. They need to get the data in the hands of their business in weeks or days.

That's what Diyotta brings to the table. It has simplified data designing and development patterns. You have wizards. You have a very intuitive interface. You don't have like a take on a desktop installation. You have everything browser-based so you're interacting with very easy tools to build all your designs without knowing the underlying technology.

That's the simplification patterns, which we are bringing to the data integration world.

Claudia: Staying with you, you have mentioned the customer case study. It's a big name, Sprint. What don't you tell me about that?

Sanjay: That's really interesting actually. Sprint, as you know, has over 50 million subscribers and has multiple product lines based on prepaid, postpaid and spectrum services they provide to their consumers. They embarked upon their big data journey to harness the data for several key businesses initiatives including fraud analytics, network forecasts and so forth.

In order to zero down the right solution, they wanted to make sure that they come across and solve these challenges like skills gap, total cost of ownership, keeping pace with the rapidly evolving technology, sustaining the production implementation, and balancing the time to market for their business consumers.

When they chose Diyotta, they were able to load high-volume, high-frequency data from the network into their enterprise data lake, and then extend that or to analyze those transactions in a near real-time fashion so that they can save millions of dollars in subscriber fraud which is a very key issue for all these telecoms.

On top of that, what we have seen is a rapid expansion in different use cases like a network forecast analysis where they are actually combining, unifying different technologies using Teradata and Hadoop and then IVR analytics where they are actually getting data from IVR systems into non-Hadoop systems.

They're expanding rapidly using Diyotta. That was a really good use case.

Claudia: It's a very nice use case I have to admit. Ravi, back to you. You did the demo of the product. I gather from what I saw that you allow developers to create what I would standard procedures, something like that, and be able to store them. How do see people using these?

Ravindra: Diyotta comes with the best practices like reusing the rules. Let's say if you have a rule, which needs to be used across your project, we

have a component called global object where you can define the reusable components all the way from your low-level functions to the design where you actually create your data flow. Then that part of the data flow can be reused, meaning, you can clear that particular data flow and keep in the global object, which can be shared to the developers where they can use it.

Similarly, you can standardize the approaches where they can leverage those to quickly bring the data flows to the QA system.

Claudia: Sanjay, you get the last question. Let's talk a little bit about the future. What's up for Diyotta?

Sanjay: We're seeing a lot of different trends in the market, in the data integration world, where people need data in the near real-time fashion or even streaming fashion: social media capabilities, cloud capabilities. That's where we're headed. We are going to have exciting announcements down the line.

Just in a few months, we are releasing our 4.0, which will have a lot of new exciting features.

Claudia: That's excellent. I look forward it. All right, unfortunately, we're out of time so that's it for this edition of the BBBT podcast. Again, I'm Claudia Imhoff and it's been such a great pleasure to speak first with my good friend Jonathan Wu, Sanjay Vyas and Ravindra Punuru of Diyotta today. Thanks so much, all of you for speaking with me.

Sanjay: Thank you, Claudia. It was a great time today.

Ravi: Thank you.

Jonathan: Thanks, Claudia.

Claudia: I hope you enjoyed today's podcast. You'll find more podcasts from other vendors at our web site www.bbbt.us. If you want to read more about today's session, please search for our hash tag on Twitter. That's #BBBT. And please join me again for another interview. Good-bye and good business!