



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:	ClearStory Data
Date recorded:	February 20, 2015
Host:	Claudia Imhoff , President, BBBT
Guest(s):	Kumar Srivastava , Senior Director, Product Management Andrew Yeung , Director Product Marketing Scott Anderson , Sales Engineer
Run time:	00:15:09
Audio link:	Podcast
Transcript:	[See next page]



Claudia Imhoff: 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 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.

CI: I'm pleased to introduce my guests today. They are Kumar Srivastava, Andrew Yeung, and Scott Anderson. Kumar is the Senior Director of Product Management, Andrew is the Director of Product Marketing, and Scott is the Sales Engineer for ClearStory Data. Welcome to all of you.

Scott Anderson: Thanks for having us.

Kumar Srivastava: Thank you.

Andrew Yeung: Thanks for hosting us here in Boulder.

CI: Andrew, let me start with you. You gave us a pretty good company overview. Why don't you tell our audience about the company, its history, and a little bit about the founders?

AY: Glad to do that. This is our first trip here to BBBT. To give you a little bit of the history behind ClearStory, this was a company founded back in 2011 by industry veterans coming out of Aster Data.

The problem that they saw was business intelligence as a whole really required a whole lot of IT involvement, and also from data scientists. They wanted to produce a solution that really made the process of combining diverse data together a lot more consumable.

CI: Most interesting. Now let's turn to the company mission and the three objectives that are driving your product. What are these and how do they bring value to your customers?

AY: Our mission statement is to transform data analysis by making it simpler to be able to, first of all, converge more data sources faster. Then enable frontline business users to do collaborative data exploration, to answer new



questions, and then to be able to share those insights across what we call interactive collaborative storyboards.

If I were to call out those three pillars to what really motivates users to come to our solution, it's those three things around what we call data harmonization, which is the smart blending of these disparate data sources, whether they're internal or external, the data stories that allow line-of-business users to continuously explore data to drive to new insight. Then, the interactive collaborative storyboarding, to really share those storyboards across a large audience.

CI: Well, let's talk about your customers a little bit. You gave us a few case studies. You also gave us the best use cases for your product. If you don't mind, I'd like for you to repeat those stories as well. They were quite interesting.

AY: If you look at our website, you'll see a web page that lists some of our customer base. Let me start with one particular story that comes to mind in terms of the use case in the CPG space.

This is a leader in dairy products and their top-level question was really fairly obvious, fairly simplistic. How are we performing daily by grocery store and what are the factors that are impacting them the most?

If you really think about, it, in order for them to answer these types of questions, no matter how simple these questions may sound, they really require understanding where all the different data sources need to come from.

In this case, it was the entire supply chain, whether they're internal as well as the external supply chain and logistics with their retailers and grocers. Ultimately, they also wanted to match those up with external data sources to be able to get to how their sales are performing against the broader market.

This is where we do the best with respect to the use cases with a lot of these disparate data sources, not just a number of data sources, but also the variety, the formats, the semantics and the data that users have had to



bring together to be able to get down to the answers that they want to get to.

SA: We love to call that the "cow to cup analysis."

CI: What is this?

SA: The cow to cup analysis... but across industries, whether it's media and gaming, whether it's food and beverage, CPG, retail, pharma, etc. Numerous applications with the large commonality around the need to harmonize, again, those multiple disparate data sets together.

CI: Is that why customers choose your product over someone else's?

SA: Certainly that associated with ease in which you can collaborate. You're not always the person with the answer in your head, or a subject matter expert necessarily. But if you can quickly and easily work with them across many data sets, you're going to be greatly empowered to get to those insights much, much faster.

CI: Certainly two heads always better than one, right? Kumar, let me bring you into the conversation to get a little more detail. You have an integrated analytics application and platform. Why don't we talk a little about how that works? The architecture that's underneath it, maybe some of the technology that you bring to bear on some of these problems.

KS: We are a cloud-hosted application, which means that our users and customers get logins. They log in to the system using a browser. We are using internally Apache Spark. It's an open source technology. It has a lot of benefits and value around faster analysis. That's the core of our infrastructure.

What happens when you think about these disparate data sets or data sources, what we are doing, we are trying to connect wherever the data might be. It might be on the premises of the customer. It might be in the cloud. It might be public sources.

What happens is we connect to these data sources. We'll bring in that data. We run a process called 'inference' over it, where we try and understand



what the data contains, what the values look like, what are the ranges, what are the types in that dataset.

Once we build this inferred metadata model about what the data looks like, we are in a really good position of enabling what we call "harmonization." Once these users enter the story expedia environment where they start looking at datasets, we are then able to use this informed metadata to actually enable users to harmonize or blend multiple disparate data sources together and, in that process, start exploring and analyzing these disparate data sources.

Because we run on Spark, we heavily use the concept of RDDs within Spark to enable far-cycle analysis to offer collaboration that's context-aware which enables you, as you collaborate with someone, you exactly see what they're talking about or what they're referencing to when they're discussing an analysis or an approach with you.

That enables high-quality, predictable insights that can be repeated to offer the same answer again. We do have a robust API layer, which means that you can utilize the information that ClearStory has, the intelligence, the metadata that we've created.

You can connect to it through an API and bring that information to sit right next to your data and make that data, where it's stored in its raw form, more usable and discoverable.

CI: The last piece, the collaboration was fabulous, by the way. I thought the interface, you did a very good job, Scott. The interface was quite nice. The last piece that is really interesting is that storyboarding piece. If you don't mind, spend a minute or two describing that.

That's kind of a differentiator for you folks. I think. The collaboration piece certainly is as well. But the storyboarding, it doesn't make sense to just present results to someone like a CEO without any of the context, without any of the story behind it of why is this happening? Is this a good thing or a bad thing?

It's just a number. I really liked the storyboarding side of that. Do spend a little bit of time talking about it.



KS: The core problem at heart that storyboards are trying to solve is that collaboration needs to happen not only during the process of coming up with the analysis or coming up with the insights, but also as you are spreading these insights to a broader set of users that might not have that capability to come up with the insights themselves.

Storyboards really are your way of delivering these insights to a broader user base, and then controlling that message in how they receive the information. Their concept came out of analyzing our current customers and the general problem in the market with things like dashboards which tend to be very static.

You cannot ask a question to a dashboard because you have to go back to IT. They have to go in and come up with the answer, actually change the dashboard and even sometimes create a new one to answer the question. You lose time. It's a lot of effort to go and talk to someone about just getting their answer defined and integrated into the dashboard.

Storyboards are designed to enable that interactivity.

It's not a static view of the information. It's not about a fixed set of APIs. It's about something, just like your business, it's changing what you want to track, what you want to look at to run your business or your product or your marketing changes.

You need these storyboards that are live views of your data. They can be added and enhanced depending on how your business is changing.

If a certain set of users have the ability and the capability to self-service and answer their own questions that will come up when they look at the data displayed in the storyboard frames, they can be proficient to go and self-service their analysis.

They get the benefit of the initial analysis and the story environment which they landed into, because the entire context about how the initial set-up, the version of the insight came up with is there, which means they are not only able to self-service and get the right answer, but the answer is going to be right... because they're able to use all the value, the context and the breadcrumbs that were captured to get to the initial version of that analysis.



CI: I thought that was the most interesting part is someone builds the foundation, probably a data scientist or a business analyst of some sort. Then the business users, by and large, can use that foundation. They don't lose it. They stay within the confines of that foundation to continue on.

KS: Exactly. We have collaboration on the storyboards themselves, which means that if you don't have that ability, you can still contact and connect with that steward or the author that came up with that insight and analysis and ask them questions.

If you cannot self-serve because you might not have the skill set, you can bring in the right people who were responsible for creating and delivering that information, ask them that question, get your own questions answered and move on to your next stage, where you then focus on implementing the insight into your workflow.

CI: Yeah, very interesting. Scott, let me bring you back into the conversation, because you were the demo guy. It was well-done.

You walked us through all three of the phases, the harmonization, then the collaboration and then finally the storyboard. What do you think differentiates your product from...?

There are many competitors, let's face it, these days that claim to have storyboards, that claim some kind of collaboration capability and so forth. What do you think differentiates ClearStory Data from all these other companies?

SA: If you take away and you strip away the marketing language, if you actually live and breathe ClearStory, this question is answered immediately, so love to be able to demo it for anyone else. Come and ask us, please.

CI: There is one on your website. We might mention that.

SA: It's the end end experience. The way in which we deliver, there's a lot of intelligence around bringing data in that we then use to bring the data together as we harmonize it.

The collaboration piece really ropes together all those different subject-matter experts and individuals across the organization. Then the



ability to land this in front of a much larger audience, but still keep them engaged.

Even if you think outside of the enterprise and look at an external-facing use case for presenting a storyboard as a portal to any of your distributor or your partner networks or some other business with which you're doing business, there's not much stickier relationship.

That account manager can sit and interact with you. Someone outside of the organization can put their question in alongside of the data. Now we can very easily and quickly address that and keep up with the pace of the business that we're trying to operate.

CI: That's nice. That is a significant differentiator for you folks. Plus you've got the scale and the speed and the performance of the cloud and the technology underneath your architecture which certainly doesn't hurt.

SA: It was a great bet we made three years ago with Spark. We were one of the first groups to capitalize on that. Wonderful lead there.

CI: Good choice. All right, Kumar, let's end with you. The future of ClearStory Data. What can you tell me?

KS: That's a great question. It comes down to, we exist. We enable enterprises to get to insights faster, and enable the organizations to utilize those insights faster and get to a higher business value. It's all about reducing the time it takes to get to your final destination, whatever it might be, from just starting to look at data.

We wanted to enable users to look at more data sources, different types of data, handle all the data wrangling and data prep. The time and money they have to spend to do that, we want to minimize that, reduce that as much as possible.

We think that the future is all about intelligence. We want to look at automating these processes, trying to figure out what a human would do except do that faster, because we will be automated.



We want to look at that, come up with the algorithms and the intelligence that can define that, and then offer this information to whichever part of your stack might need it.

We want to have a good API story that lets you connect to our back end and use the information we have. Whatever the way you store it might be, we don't want to replace it. We want to definitely make everyone in your organization have the ability to discover, analyze and come up with insights as fast as possible and then ensure that the right people will see the right insights and they're with the right context.

CI: You've got a bright future ahead of you, it sounds like to me. It's certainly very clear.

That's it for this edition of the BBBT Podcast. Again, I'm Claudia Imhoff.

It's been a great pleasure to speak with Kumar Srivastava, Andrew Yeung, and Scott Anderson of ClearStory Data today. Thanks to all of you for speaking with me.

KS: It was a pleasure coming here today. Thank you for hosting us. It was amazing.

SA: Many thanks.

AY: Thank you very much, Claudia.

CI: 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!