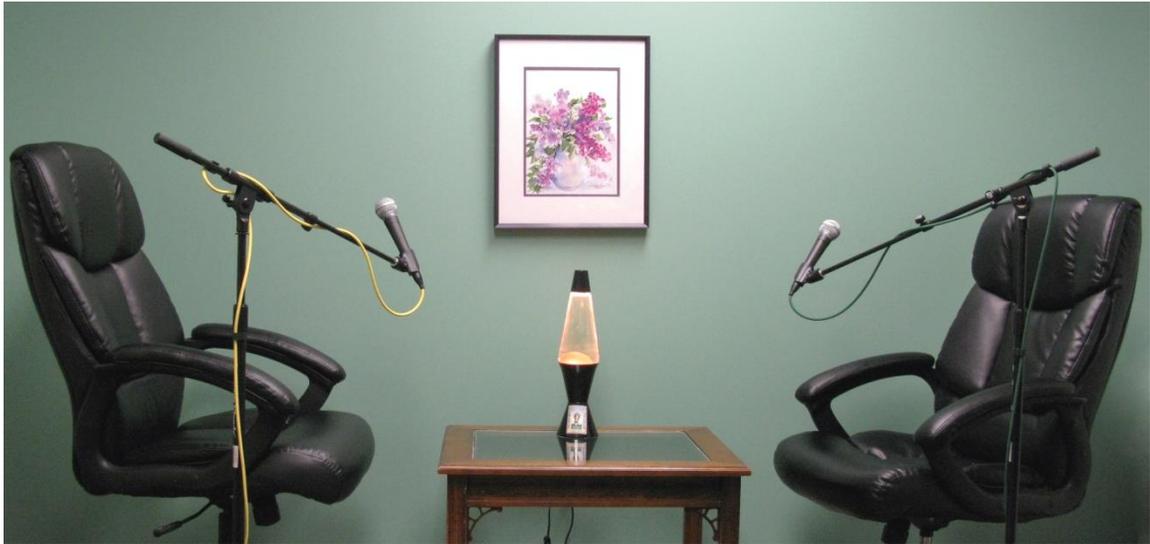




BBBT Podcast Transcript



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: Paxata
Date recorded: January 17, 2014
Host: Claudia Imhoff, President, BBBT
Guest(s): Prakash Nanduri, Co-Founder and CEO
Nenshad Bardoliwalla, Co-Founder and VP of Products



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.

I am pleased to introduce my guests today. They are Prakash Nanduri and Nenshad Bardoliwalla. Prakash is the Co-Founder and CEO and Nenshad is the Co-Founder and Vice President of Products for Paxata. So, welcome to you both.

Prakash Nanduri: Thanks, Claudia. It's wonderful to be here.

Nenshad Bardoliwalla: Thank you very much.

Claudia: All right. Well, Prakash, let's start with you. You started this morning's session with a very strong, very passionate vision that you have for your company. Please go over that a little bit for us, and tell me why it's important to you and important to your company.

Prakash: Absolutely. Claudia, it starts, at first, understanding how the business analytics space has evolved over the last 30 years. First, [it] starts with the fact that, while there have been some tremendous solutions in the enterprise around analytics, and most recently with self-service, business user based visualization analytical tools. The key problem towards data driven decisions in the enterprise still remain.

Why does that remain? It remains because a huge part of any analytics exercise is in the way we prepare data -- how we prepare the data so that it can be analyzed for a decision. That was the basic, underlying reason why I embarked on this journey.

The other very important factor was that I have a passion for creating new organizations with like-minded individuals, and I strongly believe that great companies are founded on three basic points.



One is they create a product or a solution that fundamentally changes people's lives. Second, they earn the respect and the passionate following of a strong set of customers and evangelists.

Third, last but not least, these companies create a team and a culture that makes the first two things happen. Those are the three reasons why I've embarked on this journey with Nenshad, with Dave Brewster, Chris Maddox, and others in this company, to make this a fundamentally different company.

We have looked at data preparation for 30 years. There have been so many different solutions around this, but the fact of the matter is, while most of the solutions, and there are some good solutions, these solutions have been targeted towards the IT users, to the deep technical experts.

They have not addressed the needs of the business analyst in sales operations, marketing operations, supply chain operations, financial planning and analysis. These are the people who day to day struggle, and are always asked to put data together so that there can be an operational meeting or a dashboard of some sort. They lose their weekends preparing data. We want to change their lives.

Claudia: That's a wonderful vision, and, certainly, I know there are a lot of people that appreciate that. So, you've emphasized that you are an adaptive data preparation company only, and you threw out a challenge that I'm going to just read. It's right off of your slide.

This is the challenge, "We believe the most critical challenge you face in decision making is no longer in analyzing the data. It's getting the data ready to begin with." That sure does throw down the gauntlet. Why do you think that's so?

Prakash: Because, if you look at three major types of analytical needs, there is ad hoc analytical need, there is packaged analytical applications, and there is the traditional enterprise data warehouses. There are three types of analytical needs in the enterprise. If you look at each of these buckets, there have been fantastic solutions.



You can't speak about some of the great applications, such as Anaplan, analytical application that's out there that's doing phenomenal jobs. We all know how strong data warehouse solutions have been in the market for a long time, and we also know that our friends at Qlikview, Tableau, Spotfire, and others have created a very easy and simple way for a businessperson to visualize and analyze data.

Those things are there today, but still, when you go into the enterprise, people have difficulty in making data driven decisions. Why? Because 80 percent of many analytical exercises is centered around data preparation.

What is data preparation? It is, again, pulling together my corporate data dump from a transactional system, my personal data, which could be in an Excel spreadsheet, an external data set that may be coming from a source like Nielson or Dun and Bradstreet or whatever. Having to merge that, that takes up a lot of the time. That takes up my nights and my weekends.

That's why we believe that the last piece of the puzzle, to make sure that we actually get to a data driven decision fast, is in data preparation.

Claudia: Excellent. I couldn't agree with you more. I think it's a very interesting area that you've gotten into. You also have a rather unique sales model, I guess. You call it a "frictionless" sales model. Explain what you mean by that.

Prakash: When we embarked on this journey, we did not want to innovate only on the product and technology side of things. We wanted to innovate on all fronts, on our customer experience and also in the way our customers can purchase our products.

When we say "frictionless", a lot of times the data preparation solutions in the industry have been around long enterprise sales cycles with deep IT and technology evaluations. Frankly, by the time the businessperson gets the answer, it's too late, because of this whole sales cycle and this whole lengthy approach to acquiring the solution.

What we have done is, just like other cloud based solutions are allowing people to very rapidly acquire the products with a low cost of entry, we have offered it in a multi-tenant, cloud based solution, where any individual



who wants to do data preparation can come to our website, very quickly do a trial, and sign up for our solution on the fly.

They can purchase our product with a credit card. It's very straightforward. It's made data preparation accessible to everyone and also easily adoptable.

Claudia: Nenshad, let me bring you into the conversation a little bit. One of the slides that you had was the, I guess, more traditional approach to data preparation, which is more of an IT centric type of approach, according to you, versus the more business centric solution that your company uses.

I want to get at the difference here, the differentiation. What is the difference, and more importantly, what's the benefit of your approach versus the more IT centric one?

Nenshad: I'll be happy to answer that for you, Claudia.

If you look at the more traditional, IT centric approaches to data preparation, they all evolved in the era of the relational model. They all assume that the underlying data must be converted into a common relational substrate in order to be able to manage through the pipeline process.

As you know very well and have been tracking in the industry, the variety of data sources that people are using today is actually much, much more diverse than we've seen before. People are working with data in key value stores, in graph databases, in document stores, columnar structures.

We think that in 2014 the average business analyst is pulling data in from Twitter, which is a JSON feed, pulling in Nielsen data, which might be in an Excel dump, and their relational data inside of their corporate systems. We think that it's very important that it's not just a purely relational approach or a relational translation approach. You really need to be able to deal with all the various data types in a native format.

The second really has to do with the fundamental difference between predefinition and being able to define things on the fly. In the traditional IT approach, that you would know well, what we do when we do data



preparation is we often design a data flow diagram. We actually use a procedural language to instruct the computer as to what steps we want to take.

We want to pull data from a certain table, we want to apply various expressions and transformations, and then we want to be able to merge it or join it with other data sources. All of this has to be decided in advance, and then it is scheduled in a batch based mode, where then that process can be run repeatedly.

For business people, thinking in advance that way is just not the way they're wired. If you look at the way that people work in Excel, they're actually looking at the data, and then making the transformation in the context of the data that they're working with.

In our business-centric approach, we use a data-centric model, not a procedural model. All the changes that people make are declarative. You just tell the system what you want to do, and it figures out how to do it.

That allows you to not have to pre-define an event, what different changes people want to make to the data, and that level of flexibility and the level of interactivity that a business-centric solution provides, is what makes it a lot more suitable for the ad hoc, very ephemeral workloads that people work on in an environment where business is changing very rapidly.

Claudia: Let's keep on with the IT versus business kind of scenario, here. We got into quite a discussion of your solution being able to support both, being able to support the very technical IT person and also the less technologically savvy business person, both of whom have data preparation roles.

It was an interesting discussion, I think your tools certainly does support these two environments to a certain extent, but we also got into the discussion of what happens when the business user uses this. How do we control what they're doing, do we want to control what they're doing. Are they technologically savvy enough to actually use your technology. Give me a little bit of your insight into these questions.



Nenshad: Sure. Happy to. The idea behind the solution, I think you very rightly point out, that we are not looking to bypass IT. Many of the software and service solutions that have evolved over the last decade have either only catered to the business side and gone around IT, or they've catered to the IT side and not really paid attention to the business. We believe that there is a happy marriage that we can provide.

The truth is, if you look at the average company, Fortune 500, Fortune 1000 company today, what happens in the business side is that they do have data that's prepared by the IT team, in the enterprise data warehouse, multiple federated data marks, etc.

If it is not in a shape or form that is for their current needs, they inevitably dump it to Excel, and then they have it, massaging that data so they can turn it into whatever they need. That is the current status quo.

On the other hand, having been in this industry for 15 years myself, what I've also noticed is that if you don't give people that flexibility in the tool that they're working with, they will immediately go to Excel. We've always been caught between a rock and a hard place in terms of providing tools that give business people empowerment, and allow IT to actually see what the heck people are doing with their data.

Our approach is to say, "Look, we know if we don't provide the business analyst in the context of what we call a project, which you can think of in some ways as a sandbox that is isolated for their use in the context of a project, they should have a massive amount of flexibility."

We know if we don't provide that, they're going to go to Excel anyway. We know they have that alternative. We provide a massive amount of flexibility, but at the same time, we also provide accountability through our data time machine and lineage capability.

What's interesting about the Paxata system is, as a business analyst, you can do pretty much anything you need to do to get your job done, but you can't hide under the transformations that you've made.

If your Q1 forecast suddenly looks a bit more optimistic, that's recorded in the Paxata system. If the data that came in from an external system needs



to be massaged in a certain way, that data is recorded in the Paxata system.

What we're trying to do is find a blend that allows you to emergently come up with a governance model that is constantly balancing between the tensions of standardization, that come from the IT organization, and the flexibility and the agility that people need on the business side. That's how we really think of breaking down those barriers.

Prakash: I also want to add that, at the same time, the IT folks have full control on what kind of data sets are available to what kind of user, so that they're in not a full free for all, it is managed and it is up to the department or the division or the entire enterprise to come up with those in line with their data governance rules today. That's the beauty of the system.

Claudia: I love it. I think, yes, you've certainly crossed, or [are] acting as a bridge between IT and business. You've also mentioned a number of my next question's features, and that is you gave us a very good in depth demo of the product. Thank you very much for that.

You have, you've already mentioned a number of the key features. Number one in my book, was that ability to govern what people are doing to the data. I thought that was very well illustrated. I wanted you to touch on, just briefly, we're running out of time, but just briefly if you could touch on some of the other key features like the idea of your "FilterGram" and your semantic integration and so forth. Just a few of them.

Nenshad: Sure, I'll be happy to do that. One of the key features you mentioned with the notion of what we call visual preparation. Again, we're trying to build a capability so that the average analyst in the enterprise has the ability to do data preparation.

Our system is designed to be interactive, data-centric, and declarative, with providing visual capabilities like our FilterGrams, which work with both textual data as well as numeric data, to make it very easy to identify where the outliers are.



What the distribution of the values?, etc., are commonly what would be referred to as data profiling like capability, but in a completely interactive consumer like interface.

Where things get really interesting for us, and one of the key features that we think is very compelling, is the notion of our use of semantic technologies to type the data, to enrich the data, and also to combine the data. In Paxata's solution, the average analyst does not need to understand the underlying structure of the data.

We can automatically detect, using the patent pending algorithms, that Dave Brewster, our Co-Founder at CTO, designed that allow us to detect the meaning of the data and automatically in a probabilistic way, detect the relationship across the data sets that allow the end user to very flexibly combine data to see the view that they want to see across multiple data sets.

We take that another level further, and are actually able to ascribe that meaning using technologies from the semantic web world, terms like ontologies, which of course are not used by business analysts, but certainly they have the concept of how cities relate to states, and states relate to countries, and that hierarchy of knowledge that people have.

Our system is able to detect that information and reconcile against other trusted sources of information like third party web sources, like internal master data management systems, or even other Paxata projects, and then based on that reconciliation, to recommend additional attributes that one might want to enrich on.

If I know that your cities and states are the correct combination, I can offer you population information or geocode information, that you can fluidly, on the fly, without any predefinition in advance, bring that data into the context of your data preparation task.

These are some of the things, the visual preparation capabilities, the automatic relationship detection, the automatic semantic typing, these are innovations that we're bringing to the market that we haven't seen other folks provide to date.



Claudia: I would agree with you. I think you've got a rather unique set of features and a unique product. Let me go back to you, Prakash, as a wrap up on this. You ended our session with your going forward plan. Again, keeping with your vision, there were four areas that you wanted to focus on for the future. First of all, what were they, and why do you feel that these important to Paxata?

Prakashi: It goes with our beginnings of what our vision is, what our strategy is. The four areas we're going to continue to expand on and continue to deliver on in 2014 and beyond are, number 1, our product innovation, keeping in mind direct customer feedback. This is not building it in a vacuum; it's about product innovation. Our solution will involve a number of dimensions, including the governance, including the enrichment, including the scalability, all of those aspects, including the user experience. So number one is product evolution.

Number two is our continued focus on being a strong ecosystem player, where we focus on the needs of the ad hoc analytics based in the enterprise, and continue to partner with the likes of QlikView, Tableau, Spotfire, and others to be able to deliver a data prep solution for these guys. That is what we will continue to do. We will continue to partner with the likes of Cloudera and others in the infrastructure space.

Number three is to focus on our go to market efforts. It's a frictionless land, succeed, and expand model where we can enable any analyst who wants to use our product to very quickly adopt it. That's why we have the multi-tenant approach. We have the Personal Share and Enterprise packages so that anybody with their right requirements can quickly adopt Paxata.

Last but not least, we will always be centered around the people who use our product on a day to day basis and derive value from it.

These are our passionate advocates, these are the people we work for. On the four areas of product evolution and innovation, continuing to be a healthy partner in an ecosystem, continuing to focus on delivering value by our land, succeed, and expand model, and last but not least, addressing the needs of our passionate advocates. That's what we're going to go forward with.



Claudia: I think it's a marvelous forward looking view and certainly in keeping with everything that you've mentioned before. Unfortunately, though, we are out of time, so that's it for this edition of BBBT podcast. Again, I'm Claudia Imhoff, and it's been a great pleasure to speak with Prakash Nanduri and Nenshad Bardoliwalla of Paxata today. Thanks so much, both of you, for speaking with me.

Nenshad: Thank you.

Prakash: Thank you very much.

Claudia: I hope you enjoyed today's podcast. You'll find more podcasts from other vendors at our website, www.bbbt.us. If you want to read more about today's session, please search for our hashtag on Twitter. That's #bbbt. And please join me again for another interview. Goodbye and good business.