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:	**Revolution Analytics**
Date recorded:	**March 8, 2013**
Host:	**Claudia Imhoff**, President, BBBT
Guest(s):	**David Smith**, VP, Marketing and Community
Run time:	**00:16:15**
Audio link:	**Revolution Analytics**
Transcript:	[See next page]
Transcribed By:	CastingWords

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

I'm Claudia Imhoff and I'm pleased to introduce my guest today. He's David Smith. It's the second time you've been here. David is the Vice President of Marketing and Community for Revolution Analytics. Welcome, David.

David Smith: Thanks very much, Claudia, and thanks for inviting us back again. It's great to be back here in Boulder.

CI: It's wonderful to have you here. It's been a little more than a year and I have to tell you, Revolution Analytics has had a pretty darn good year. Why don't you talk to me a little bit about the changes. What do you attribute the growth to?

DS: I think it's a couple of things. One of them is that big data has finally come of age. We spent a lot of time working with organizations who were just getting to grips with this idea of Hadoop and ingesting different data sources and munching data together. Having spent that investment, the next logical question is, "What are we going to do with this"? "What value can we actually get of all this data that we've spent this time and money collecting"? That value comes from predictive analytics being able to make forecasts about what's going to happen in the future, being able to make inferences to help people make decisions.

Using Revolution R and the R language to build bridge models around this big data is what a lot of companies are really, really interested in right now.

Another thing is, in terms of changes, we have a great leadership team that's joined Revolution Analytics over the last year. Our new CEO, David Rich, who came to us from Accenture Analytics, brings a lot of experience

to the company about real world deployments of analytical applications within businesses that has really helped us hone our message and our product line on that basis.

CI: You've also expanded the employee base. You're going to grow the offices -- even one in Dallas.

DS: That's right. We aim to double our employees by the end of this year and expand our international presence from Singapore to the UK, and also offices also here in the States.

CI: Wonderful. One of the discussions that we've had off and on with vendors that have come in, especially those that deal with big data analytics, predictive analytics, a lot of discussions about whether these types of analytics belong in the hands of the "everyday Joe" working at some company, or can they even handle these kinds of sophisticated analytics. Do they belong in the hands of a proper data scientist, not someone that is just mucking around with them? We go back and forth on it. My stance has always been that I think they belong in the hands of a more sophisticated user, somebody that's trained in statistics, or data mining, or whatever, and that they can create analytics that perhaps will show an "everyday Joe" something on their screen.

But that they really do need to understand the analytics behind it. I want your opinion on this as well.

DS: I'm glad you said what you just did, because I feel like a little bit of a heretic coming to the Boulder BI Brain Trust to say that I think that's one of the areas where BI didn't fulfill its promise. The idea behind BI was, "if we just give a business user, a salesperson, or a product manager, or an operations manager enough tools and algorithms so they can slice and dice data and apply sophisticated techniques to them, that all their problems will be solved."

I don't think that promise was fulfilled. The reason why is because number one, these people don't have the time to really delve down and work with data at that kind of a granular level, but also, they don't have the expertise. They have expertise in their particular domains. They know about their products, they know about their customers, they know about

their operations, but a data scientist is the person that can actually apply these advanced analytical techniques, like predictive models, and regression, and decision trees to data.

That's a really important concept, I think, for me. Because you need to have somebody who understands where the data comes from, what it is that we're trying to forecast or understand, and how to use all those predictive models in a way that's both effective and powerful, but is also robust and safe.

CI: Yes. What bothers me about turning data mining tools or data predictive analytic tools loose on the population at large is that many companies tend to dumb them down, make them so simplistic, or bury the algorithms so that people don't see them. I think that's wrong, because you can pick a bad algorithm, you can pick the wrong set of data, or you can do something that will give you misleading information.

DS: That's right. I don't think we should be putting people in the position of making them choose between a K-means clustering algorithm versus a neural net because that's not what they ought to be doing. They ought to be managing their businesses. The philosophy that we take is that absolutely, business users should be using predictive analytics, but they should be using the results of those predictive analytics through point applications.

CI: Absolutely. They should be consuming them, not creating them.

DS: Exactly. Our philosophy is, we work with a lot of our OEM partners and our application customers who have expertise in their particular fields. Who might have expertise in, say, marketing analytics and can create an application for marketers based on their own expertise of the industry and the expertise of their data scientists to deliver advance predictions and forecasts an analyses to those end users through a dedicated application. The end users may not realize, at the end of the day, the nuts and bolts of the advanced analytics that drive it, but they get good results.

CI: But they sure do understand the results to it.

DS: Yeah.

CI: Excellent.

        Well, let's turn to one of your customers. You spoke this morning about Upstream. It was a really interesting case study. A little creepy sometimes. Perhaps they know more about you than you think they do. Why don't you describe the case study for me?

DS: Yeah, absolutely. Upstream Software. They're great people, they're one of our customers based out in the Bay Area. They provide our marketing analytics application as software is a service.   It's a web-based application that their retail customers, like Williams Sonoma and Neiman Marcus, can log into to understand how effective their marketing spend is. I work in the marketing department, and sort of a common adage in marketing is that half of the money you spend in marketing is wasted -- and you don't know which half. That's no longer true, because with predictive analytics, you can actually figure out all of the different messages and communications that are being sent to your customers, correlate that with the actual sales that they make, and then figure out, of all of your marketing spend, how much of your sales can be attributed to direct mail versus online advertising, versus website, versus mobile, and use that to actually tune your marketing spend to achieve your corporate goals.

CI: That has been the bane of marketers forever.

DS: It has been.

CI: It really has been.

DS: It's being sold. It's really, really cool.

CI: Your example of a website ad that happened to you in particular was quite interesting. They knew where you came in from, they knew what you were looking at. They knew your Twitter feed, and so forth and so on. They could figure out kind of what your interests were, just from the correlation of all of those very disparate sources.

DS: That's right. This is kind of falling into the uncanny valley of predictive analytics, where we're actually seeing the results of predictive analytics algorithms that have this insight into you as an individual. Now, whether you find that creepy or whether you find that valuable depends on your particular perspective on this. Personally, I find it valuable. I really like the idea that now I can visit a website, and rather than getting ads about sports, which I have no interest in, versus ads for products that, hey, this is actually something that I would really be interested in, I personally find that valuable.

They're able to do that today because there are so many more rich data sources available to marketers and data scientists to build the bridge of models that sit behind this. We have data providers that know information about us as individuals, our interests, our products, the websites we visit, where we live. The retailers themselves have their own data sources around which loyalty programs we're members of, which marketing campaigns we've been exposed to, [and] what products we've purchased. By mashing all those data together, going through a data distillation process to build a predictive model, you can really get some results that, in my opinion, are quite magical.

CI: One of the things, that we also have to be careful of with these models, is that people do change. They get older, they have kids, they retire, whatever it is. They do change. The model that you've created for them needs to keep up with that. Right?

DS: Yeah.

CI: Fraud models come to mind immediately. As soon as somebody figures out that you've got the model down, they'll change the way they perform fraud. It's almost immediate. They know immediately that the credit card has been detected and they then try something different. How do you handle that kind of really rapid change in these models?

DS: That's actually a really good example of why you need humans in this process. Because it's humans that are reacting to the predictive models themselves, changing their own behaviors which, in turn, implies you need to make an update to the model. Black box systems can certainly give

you good prediction as things are today, but as things change, they really can't keep up with structural differences in the data. I outlined a process in the presentation this morning of booting the predictive models, which is necessarily on a static snapshot of data as it is today, but building into the process, number one, a system for monitoring those models to make sure that the predictions from those models commensurate with the actual outcomes for the models and having some kind of signal to figure out when those models need to be updated.

When that signal is triggered, to go back to the data scientists once again, to go back and look at the new data that's come in. Also new data sources that have become available, so they can build new predictive models that are maintaining the performance characteristics that they're looking for.

CI: That's brilliant. That's what's needed, isn't it, that closed loop if you will, of our feedback area?

Let's turn to Revolution Analytics itself, a little bit, and tell me about the Revolution R Enterprise. What is that composed of?

DS: Revolution R Enterprise is our software platform that makes the R Language ready for the enterprise. It starts off with that open source R engine. It's throw everybody out there that may have been using R. It's exactly that same engine, 100 percent compatible with R, and all the community packages out there where it's such great resources for innovation. But we make that run faster by making it run multi-threaded, and we add our own proprietary components to it to make it possible to deploy our inter-production applications. Just a few of those components include big data predictive analytic models to make R scale to big data, and run those models very, very quickly. Connectors to other data sources, like Hadoop and databases and other analytic platforms like SaaS. A deployment API, a web services API, so that an application developer working in, say, Microsoft Excel, or an iPad application or a BI tool, can use the languages used to extend those applications and call out to our Web services API and get the benefit of analytics developed by data scientists in R, without those application developers themselves needing to know R, which is pretty rare.

CI: Pretty nice stuff, indeed.

The other thing you've got that is remarkable is a pretty darn good partner ecosystem. You put up a slide, and I don't even remember how many logos are on there, but it was pretty impressive. You use partners a lot. How does that help your company, and what do the partners get out of this?

DS: Actually, I of think that who it helps the most is our customers. The alternative is, if you wanted to build an application built on bridges of analytics, is to go with a company like SAS or Oracle that provides the entire stack, from data to analytics, to the solutions and applications themselves. We take a <u>very</u> different philosophy. We focus purely on the predictive analytics, the R language, and the tools that data scientists to build those predictive models. We let our customers connect with data sources wherever they might be. There are a lot of new data technologies out there now, like Hadoop, like NoSQL systems, plus traditional relational databases, but also new streaming sources of data, which R can connect into, rather than having to lock it into some sort of proprietary platform.

We also take a different philosophy on the applications, the end user applications. We don't deliver solutions, we don't deliver applications ourselves. We enable our customers and partners to use their own domain expertise in the specific market areas to build applications on our platform, which they deliver to their employees, customers and partners.

CI: Let's talk about your customers a little bit, because I'm curious why they would choose you over R, perhaps, or over some of your competitors. What do you bring to the table for these folks?

DS: That's a great question. What I hear from our customers, why they've adopted Revolution R, innovation is something I hear a lot. People want to enable the data sciences resources they've invested in. They know R, and they want to be able to get access to this huge community of data scientists and statisticians and application developers that created tools that work with R.

What we do, though, is make our work on IT platforms, IT environments. We bring it to the data, to connect with the databases, the data appliances, the new data streams. Plus the server hardware and cluster hardware that IT departments want to be able to bring to the table in making these things run fast with big data.

I think another thing is that it's such a dynamic changing world, there's onset of big data, companies are really looking to capitalize on their data, and using that as kind of the key competitive differentiation. Those companies that can make forecasts or understand patents and changes in the market better than their competitors do, on the basis of the data, are the ones that are getting ahead.

You need to act quickly. Something that's really unique about the R language, it allows people to go from this concept of an application build on data to a working application in a very short space of time. Sometimes just days, sometimes a few weeks but in a much shorter space and time than competitive, off the shelf applications.

The bottom line though, is value. Especially for companies that are invested in legacy predictive analytic software like SAS. We bring to the table a much more competitive solution with innovation, with multiplatform, and serving the needs of data scientists at less than half the cost.

CI: That does seem to be the big ticket right there, isn't it?

Most interesting to talk with you, but unfortunately, we're about out of time. That's it for this edition of the BBBT Podcast. Again, I'm Claudia Imhoff, and it's been a great pleasure to speak with David Smith of Revolution Analytics. Thanks so much, David.

DS: It's been a pleasure. I hope to come back again in a year's time and show you where we've grown since then.

CI: I do hope so. I also hope you enjoyed today's session. If you want to hear more about today's session, please search for our hash tag on Twitter, that's #BBBT. Please join me again for another interview. Goodbye and good business.