



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:	Composite Software
Date recorded:	January 4, 2013
Host:	Claudia Imhoff , President, BBBT
Guest(s):	Jim Green , CEO
Run time:	00:17:36
Audio link:	Composite Software
Transcript:	[See next page]
Transcribed By:	CastingWords



Claudia Imhoff: Hello and welcome to this edition of the Boulder BI Brain Trust, or the BBT. We're a gathering of leading consultants, analysts, and experts in business intelligence who meet with interesting and innovative companies here in a beautiful Boulder, Colorado. We not only get briefed on the latest news and releases, but we also share our ideas with the vendor on where the industry is going, and help them with their marketing direction and messaging. The BBT podcasts are produced by my company, Intelligent Solutions, and are sponsored by the BI Network. I'm Claudia Imhoff, and I'm pleased to introduce my friend and my guest today, Jim Green. Jim is the CEO for Composite Software. Welcome, Jim.

Jim Green: Thank you.

CI: It's nice to have you here. It's been a long journey for Composite from the early days of enterprise information integration. You've certainly gone way beyond that. Tell me a little bit about how Composite's doing these days.

JG: We're doing fine, thank you. We are finishing up a record year in sales, and the product continues to be stronger. The most important thing is that our customers are actually using it in bigger and more expansive ways than they ever have before. It's really fascinating to see people take our work and do things that we didn't imagine they would do. It's a fascinating business.

CI: We're going to talk a little bit about a customer presentation that we heard. You're right, they're using it in ways that I would not have imagined. It's not just integrating a couple of pieces of data together, they're actually using it in very innovate ways across the entire enterprise now, right?

JG: That's right. Data integration and data management seemed to be very well understood in the middle of the last decade. Over the last several years, though, things have changed quite a bit. We're now seeing more and more people understand that they cannot get all their data in one place. There are many cases where an enterprise data warehouse is the right answer, and the best answer, but there are many others that just don't work, particularly with the expansion of XML data types, which don't



fit into a relational schema that well, and with the more recent emergence of Big Data, which don't fit into any schema all that well.

We have broken the boundaries of the relational system, which was a simplification mechanism that allowed things to be homogeneous. In today's world, you just don't have that cleanliness, that elegance that you had before. And so virtualization has stepped forward in order to actually help put all that together.

CI: Composite, as you mentioned, has been successful. You've had a banner in 2012. It's now 2013. What do you attribute to Composite's success? What's causing the traction? Is it the things you just mentioned, the fact that there are all these different sources of data, these unusual sources that don't fit nicely into boxes?

JG: As always, there's a combination of things that create a confluence. First off, business has emerged from the recession, so we're now seeing spending rise as a whole, and we're definitely a recipient of that. The second thing is that all new ideas have to go through an incubation period. During several years of the previous decade, we have people in experimental mode, much the same as Big Data is in experimental mode today. We've now emerged from that experimental mode, and people are now very comfortable taking us to the enterprise level.

We had a number of customers that would buy our product, and now, two or three years later, They're coming back and they're buying in great volume, because they've been successful with their initial installations, they understand it, they developed their methodologies, and now they're ready to deploy on a global basis, and they understand the business benefits of what's happening.

CI: Yeah. They've gone way beyond the prototype stage.

JG: They've gone way beyond the prototype stage. I think another thing that contributed is Big Data. What do you do with Hadoop? Do you create a separate IT infrastructure? What we're seeing is mostly no. People are actually to the point now where they're starting to understand that Hadoop is not a replacement for relational. It's a supplemental technology that can be used complementary to relational. It has to be



incorporated in, and assimilated in, to the existing IT systems and infrastructure, so that people can pursue an evolutionary approach to "How do you incorporate Big Data along with your transactional systems, along with your services systems, again, XML and cloud based systems?"

There needs to be some technology to step forward to reconcile. The idea of standardization is not possible anymore. You can't be all Oracle, you can't be all IBM. What you have to do is you have to reconcile all these different steps. Therefore, data virtualization is a really good reconciliation mechanism to bring a number of these incompatible systems together.

CI: Speaking of that, the presentation from Michael McNab of Franklin Templeton was quite impressive. He really has taken it way beyond the prototype stage. Is that typical? First of all, why don't you describe a little bit about what he's doing with Composite, how it fits into his overall "data as a service" architecture? Then, tell me if that's becoming more the norm in your customers.

JG: Let me answer that in a couple of ways. Franklin Templeton is one of a set of financial services companies that use us. I'm not sure, based on our contracts, which banks we're allowed to mention out loud, but many financial services institutions use us to handle their data. What we're seeing at Franklin Templeton is typical of what we're seeing with large scale financial institutions. They have a number of different asset classes, they have a lot of real time trading systems, they have thousands of reports that they have to produce. They have a lot of compliance issues. They're very interested in analyzing data in real time in order to do risk management.

The problem that a lot of these corporations have is stovepipes. The bond trading or fixed income trading systems is a totally separate system than equity trading, is totally different than their investment portfolio, et cetera.

Many people are now trying to figure out, "How do I leverage my data better? How do I put the data in a form so that I can do more extensive analytics? How can I avoid the replications and the cost of replications of data?"



Frankly, we're hearing a lot of people talk about the licensing fees that they have to spend on the Oracles, and all of the larger corporations, because they replicate data so much. They're seeing data virtualization being attractive, not only because of the agility and flexibility, but also because it can help them save hard dollars in terms of reductions of hardware, electricity, air conditioning, and software license fees because of the rampant data replication that has occurred.

What Franklin Templeton stands out as special is that they've figured it out. Mike and his crew have done an excellent job of putting together a more holistic architecture so that data management is dealt with as a science. It's dealt with as a topic by IT, not left up to the individual application developers who create more and more specialized capabilities.

They're not alone. New York Stock Exchange and a number of other financial institutions have also, in their own way, figured out how to deploy data virtualization in a systemic way across their corporation, which resulted in millions of dollars of savings as well as increased agility and dramatically shorter development times to meet new business requirements.

CI: What I found particularly interesting in terms of the benefits that Mike talked about, two of them caught my attention. The first one was that by building this "data as a service" type of environment, he was able to build, as he put it, an appetite for rigorous governance. Governance has been one of those banes, we all know we have to have it, we just don't want it. We don't know how to do it. The fact that exposing the data in the way that he has, making it available, has actually captured the attention of the business community to do governance, or IT to support the governance movement, I thought was fascinating.

JG: Yes. I promised you two answers. Let me give you the second answer, because it does deal with the data governance issue. There's a single term called "data governance", and there's a number of different aspects to it and a number of different views on it. There are some people who have created a sense of data governance within a BI tool. A BusinessObjects Universe is an example, where the data is normalized in such a way that the report developers can access it. However, many



companies find it nearly impossible to standardize on a single business intelligence tool.

CI: Right.

JG: Taking the views of the data, and taking it outside an individual BI tool, so that you can have multiple tools benefit from it, is a form of data governance.

There are folks who are very excited about Big Data, because the proliferation of data is gigantic, especially machine generated data, and so they're dropping off all of this data into big data depositories like Hadoop, that is now separate from their standardized, very carefully thought out data warehouse, which then breaks down data governance. Now there's two ways of handling this. One is to actually use the MapReduce capabilities in Hadoop, in order to do data reduction prior to moving it into the warehouse, so that Hadoop becomes a staging area.

Another approach that is being discussed at some length, is the idea of a logical data warehouse, where some of the data exists in Big Data and some of the data in EDW, or some in data services, as you mentioned earlier, and then combining that into a single system, so that you can present an organized view of the data to the consumers. Now you have a way of actually incorporating Big Data into your IT systems, and retaining some control of the data and its representation to the data consumer, which is another form of data governance.

So, there's different points of view on this, but, yes, everybody is trying to reduce the costs, make sure they have encompassed all of the data that's possible, create a simplified view of that data, and an organized view of that data, and, in some cases, an approved, or what you might call, a "curated" view of the data, so that they can empower as many different data consumptive applications as possible.

CI: And you just listed all the benefits that he was talking about, which was pretty interesting. We have a couple questions and a little bit of time left. You mention Big Data, several times, as having a pretty significant impact on the overall data integration environment, data management environment. How does that seem to change these requirements? Is it



just because it is so much data and it's in such a -- not in a standard -- format or not a traditional relational format? Are those the things that are really stressing this whole data integration environment?

JG: Well let me turn your words on you a little bit here. I don't know that Big Data has yet had much of an impact. I do know that Big Data has captured a huge interest level, and a mindshare, and people are very interested in it for a variety of reason. There's cost reduction of using open source capabilities as opposed to vendor supplied capabilities, which have become very pricey. There's the notion that you can store data in multiple schemas in a single repository. You can grab data that doesn't fit a row/column format or a tabular format, which is kind of rigorously imposed by relational data bases. There's a number of different interesting aspects of this.

But in general, most of the industry is still trying to figure out where and how it fits in. I wouldn't encourage any of the listeners to the podcast to come the conclusion that they are behind on Big Data. Because most people are behind on Big Data.

CI: I think the hype curve is way ahead of reality.

JG: Right, but in talking to a number of people it seems to me that the use cases is falling into three camps. This is very non-scientific, but I am starting to see clearer and clearer pictures of Big Data. First is, you've got data that just doesn't fit in relational a database -- an example might be long files, it might be click through on a website or stuff that's just not relational oriented. You drop it into Big Data, you do analytics against it, and you try to discern some insight from that. That's a completely standalone separate basis for Big Data.

The second, which is probably much bigger, market, is people who are trying to grab data such as that and push it through Hadoop, in order to figure out some reduction of the data, some insights from that data that you can then move into an enterprise data warehouse and make it available to people for their operational needs.

In this form, Big Data becomes a new form of ETL, you're grabbing it form a variety of sources. The inputs might be real time. You're working it and



cleaning it up before you move it into your standard data storage mechanisms. At that point, all the downstream system are unchanged, because they're used to going to the data warehouse and they continue to do that.

The third form, which has to do more with cost reduction, is your data warehouses are getting way to big. It's not only requiring you to buy more hardware, to pay more license fees, but it slows down queries because the massive volume of data that the queries have to handle.

So people are exploring the idea of Big Data being used as an archival system. So that you move data out of your data warehouse and you archive it. Now, in the event that you want to do a query that goes against Big Data and data warehouse, data virtualization is good. In the event that you wish to do a query that goes against the data warehouse and data warehouse archival, again data virtualization is good.

So people will need to sort out how Big Data fits into their overall scheme of things. But, should they wind up with data in more than one place, they will need a different access mechanism than the BI tools typically provide. We can help them by accessing the data across Big Data, across web surfaces, and across the data warehouse. We're on the downstream side of this, when people are trying to pull data out of their multiple systems. We're the data access guys.

CI: Yeah. Well, it's certainly a very interesting time for you. You're right, there are all these new ways of using data virtualization, you've mentioned several in our chat today, I think right now you're beautifully positioned to support Big Data, whatever that is. Certainly to support the type of analytics that it can offer and so forth, especially given the fact that the architectures are now extending beyond the traditional warehouse into these new and innovative technologies, you're perfectly positioned.

JG: Well thank you. We've been preparing for this a long time.

CI: Serendipity's a wonderful thing sometimes, too.

JG: That's right. Sometimes it takes years to create an overnight success.



CI: There you go. Alright. Unfortunately, that's it for this edition of the BBT podcast. Again, I'm Claudia Imhoff, and it's been a great pleasure to speak with Jim Green of Composite Software. Thanks so much, Jim.

JG: *You're welcome.*

CI: I hope you enjoyed today's session. If you want to read more about, it please read our blogs at our website. That's www.boulderbibraintrust.org, all one big word. And you can search for our hash tag on twitter that #BBBT. Join me again for another interview. Goodbye and good business!