



BBBT Podcast Transcript



About the BBT

The Boulder Business Intelligence Brain Trust, or BBT, 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 BBT provides a variety of services, centered around vendor presentations.

For more, see www.bbt.us

Vendor:	IBM
Date recorded:	February 19, 2016
Host:	Claudia Imhoff , President, BBT
Guest(s):	Bill O'Connell , Distinguished Engineer & CTO, Data Warehousing & Big Data, IBM Analytics
Run time:	00:16:13



Claudia: Hello, and welcome to this edition of the Boulder BI Brain Trust, or the BBT. 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 BBT podcasts are produced by my company, Intelligent Solutions.

I'm pleased to introduce my guest and my friend, Bill O'Connell. Bill is the CTO of Warehousing and Big Data for IBM, so welcome, Bill.

Bill: Thank you, Claudia, so happy to be here.

Claudia: I'm happy to have you here. Today's session was just incredible. It was a lot of really thought provoking ideas. I'm going to have to cogitate on it for a while. Why don't we start with IBM's strategy for analytics? It's three parts, as I understand it, so why don't you go over those?

Bill: We look at it as three different parts of how we think about how we solve problems for the business. The real question starts to be is how do we enable the business for more, quicker ways of being agile from an application analytical perspective. We'll look at the three pieces. We broke it in three.

One's cognitive, and is looking at how do we enable the business users, through their existing tools to start moving down a path of more recommending and more of a learning perspective, so it's guiding more of a personal assistant to the users through their visualization tooling.

The other part is cloud. As we move forward, from an enterprise view -- we live in a hybrid environment -- and as we move into an environment of how we move a company forward, we still have to live with an on prem environment. We have to start looking at how do we go into more of an elastic view of how we bring in cloud into this.



We could get into a whole discussion of what data goes where, but cloud starts to be an enabler here, and in many cases, a very big enabler.

A third piece is, how do we take that cognitive capability, from a platform perspective, and interject it up into our solutions, being either Watson Health, or Smarter Workforce, or Security Division?

Anything that deals with building solutions, how do we make those smarter for the personas of the people using it, not having to have a whole set of an army of people below it trying to build that stuff.

Claudia: It's a key theme throughout what we talked to today. Let me drill down a little bit. Many people don't know what cognitive computing is, and IBM is certainly touting it right and left. It's a very important part of this strategy, as you just mentioned.

How does IBM define cognitive computing, and how is it different from our traditional analytics or advanced analytics?

Bill: There's nothing new here from an analytical perspective. It's reusing piece parts we have. I hear a lot of definitions around cognitive, and some are better than others. I hear anywhere from machine learning and predictiveness or grade.

The real issue we need to provide here for our users is, how do we start letting our visualization on a glass recommend, guide, and start thinking of more of a personal assistant to help you as an SME as you're visualizing?

I use Watson Analytics here, for example. It's not a visualization tool for doing discovery. It'll guide you through it and make recommendations. That needs to go much further, but it's changing the shift of how we deliver tooling.

Claudia: In terms of how it differs from advanced analytics?

Bill: Cognitive is advanced analytics, but it's using piece parts. We've already had machine learning. We've always had text processing.



We've always done predictive analytics, and it crosses between different personas.

How do we start to put this together and build on top of that to reuse what we already have to start building up new ways of recommending and guiding to allow the business to be more agile and adaptive but not have all the different piece parts there to do this?

Claudia: The difference, I guess, is, in some respects, what we've been doing in the past is more programmable types of queries. I know what I want to ask. It's going to respond to that known question that I'm asking.

Cognitive is different in the sense that I don't know what I'm going to ask. You recommend something that I should ask, perhaps.

Bill: You're hitting a key question, and I'll just jump in to steer it a little bit, because this is a really hard problem. Hard meaning it's a challenge, not necessarily hard. We have to look at if I'm building an analytical project or an analytical application, I need to start from some data, but then I need to start augmenting it with other data.

It's not a wrangling issue of shaping it. We know how to do that. It's not an issue of bringing it in and start discovering it. We know how to do that. The biggest question people have is what data's out there? It's information finding information so I can start recommending you look at this, this, and this, and here's the value it's going to provide.

Things aren't even reconciled on business taxonomies. I may be in one part of the business dealing with commerce and in the same company dealing with healthcare. How do I provide lenses and guide you there so you can find it and who are the people using it so I get more of a collaborative nature?

Claudia: You presented to us a maturity curve, if you will. The ultimate goal point was this cognitive company. Most of us are stuck in the data warehouse. Not stuck. That's a little negative sounding. We have



done the data warehouse. We've got that programmable query fairly well knocked down.

What do you suggest as a starting point to move toward this more cognitive type of enterprise?

Bill: It's always a journey. You can't just go and boil the ocean and build this cognitive world. It's a journey. You always start somewhere, and it has to show immediate value. That immediate value has got to be seen from the business side. Then, you start adding in on top of that.

I mentioned that data finding data issue. We can't just say, "Here's all the data in your enterprise that includes third party data, and public data, and all across the different business. Here, we're going to guide you to it."

As we start to classify and categorize systems, plural, now we can start incrementally doing that. It becomes a much more easy and scope-able problem to start to provide that journey for the users.

Claudia: To go from their traditional data warehouse to this more cognitive environment?

Bill: Yes.

Claudia: Interesting. One thing that was brought out were four key announcements -- sort of leading us into this cognitive world. If you don't mind, just briefly, tell me about these four announcements, and how do they help a company move forward?

Bill: As we look forward in building out technology, in the end, we have to deliver technology that's consumable by the business user. If we start with building blocks, and those building blocks are used by even a higher level of tooling, those building blocks need to be presented and used by the consumption of the business.

The four we just announced, which aren't future discussion but they're available today, we announced an IBM composable enterprise, which is really around how do application developers look at this and



compose things in a much more easy and agile way from an application perspective.

Another discussion we did around is IBM Graph. Graph may not be used by many of the business users, but underneath the covers, when we build things, Knowledge Graphs and other things are very key components of how we deliver information and analytics, especially around inference engines.

We built a core infrastructure around that as an open source. We start to build the infrastructure and inference engine around that. That capability's available and announced.

A third one was what are we doing around different personas? One of our personas we have to deal with is the data scientist. Though most of the users are business analysts, we still have key ones that are data scientists, and that's our predictive analytics.

We've always been in the analytical space, but how do we start to shift and evolve this to be both a hybrid environment, run in the cloud and on prem, and be more of a workbench that's also very tied into open source and R, not tied to SPSS modelers. How do we evolve that to tie into a platform that's embracing open source from a workbench?

When you look at the predictive environment, the key issue there is, how do we enable that and put that up into the channels? That's where the value comes in around the interactions on real time and channels.

To do that, you'll also need metadata. So we looked at what we did around collaborative and deployment services, which is really how we managed metadata versioning models. It's extending that one step further, so it's based off of not even IBM. It's open source. We can manage that from a workbench perspective.

The last one was an analytical exchange. Analytical exchange is around public data, so one of the things we have to categorize and classify is on prem and third party data.



We always want to start somewhere quick and easy to provide a way that a business could look at public modeled data that both IBM and third parties make available through that analytical exchange that allows you to search, survey it, and shape, and pull into your tools.

Claudia: Let me go back to a question I asked earlier, and that was, how do we get started here? Like we said, most organizations today have a fairly traditional enterprise data warehouse. To move to this next level, do they have to throw it out and start all over?

Bill: Absolutely not. We live in a heterogeneous environment, so the issue is as we as a company look forward helping our clients, we have to look at how do we move them forward? There's a stepping stone that we need to do from a timing perspective that's up and down the stack.

As an analogy here, I'll use an analogy, is we wrote a Redbook last April, and we called it data reservoir. It's not a data lake as the industry talks about, but it's how do we abstract IT to business as more of a cataloging, classifying system, so I can start building out a house with a framework. I build a house of cards and it doesn't come back down again.

That really starts to be a governance metadata discussion, which is an anchor point here. That doesn't mean that I have to go and bring all my data into one system. That doesn't work. We can't bring in one EDW. We've tried it. It's not agile.

We've tried dumping it all into Hadoop. That doesn't work. It's not agile. It can't find anything. Even if you classify it, things aren't reconciled.

We have to take a step back. How do we start moving forward in the current ecosystem to start to lay out that framework, then look at what you have in place as systems to start to streamline them as necessary or put the right systems?

It allows it to be more flexible, because the business doesn't care where data is. We can abstract it so it doesn't force IT to over



consolidate, which usually ends up happening, because IT can't solve the business's problem. They don't know how to solve it. They say, "I'm going to bring it all together so you can find it," which doesn't work.

Claudia: It sounds like there are two key components, if you will, to this movement. The first one is the idea of the top down, the business is going to be driving this. You mentioned self-service in the presentation.

The second one, though, is the ability of somebody to monitor how they're using these information assets and then having the intelligence to put it together and say, "Here are the patterns of utilization, the collaboration and everything else. Let's start to use the intelligence that we garnered from that to create this cognitive environment." Is that true?

Bill: That's right. This is a large topic, in general, to have, because we're shaping a business. If I look at stepping stones here, those used, and I know she mentioned before, shopping for data, this is a business persona type interaction. There's a lot of bread crumbs out there that we could abstract today.

It's not just about what assets are out there to find things. Think of a Knowledge Graph. If I want to find information very quickly and agilely, I also need to factor in the collaboration piece here of how people interact.

I actually can start crowdsourcing to areas, know what they're saying as they're adding their own knowledge tags, as they're typing things in and putting quality information in their own tooling, we extract that as log data. We start to extract that, and it starts to be the early part of the life cycle that feeds the governance piece.

What we're really talking about here is, how do we guide and allow the users to be able to interact very easily, and cleanly, and seamlessly with the data and us, IT guys, and the tooling always do that? We have to simplify it all, as we can't build a bunch of tools that



the users put together. We have to build more of a solution persona base, not product base.

Claudia: That gets to my last comment. It was a quote that I attribute to you, and I thought it was a brilliant quote. "The most important part of collaboration is the data it leaves behind." I thought that was brilliant, because the data it leaves behind is the stuff that we generally never capture, right?

Bill: That's right. Normally when I think of analytics, I think of building the metadata and the data I need to provide analytics to the business. When you actually look at how the business uses the data, people are collaborating and interacting through there. The problem is that's hard.

So, we have to rethink this problem out of where we need to go, we need to bring not social data, social interactions of how business users work and collaborate within their teams and between teams and the data itself, the analytics, mix those roles together.

Now, I start to have these bread crumbs that are left around that makes the analytics much, more simple, and allows me to build more inference engines that make it very simple for the users. It allows that recommendations coming in, which is really quite hard today.

Claudia: It also guides that interface that the business user has. You mentioned they don't know what's under the covers, nor should they know what's under the covers. Basically, it's the interface, the glass.

Bill: That's right, and I used an example. Darren and I used a global retailer that's also a pharmaceutical. If you look at those, if you're an individual retail person on the commerce side of the person as a business analyst, looking at data I need to enhance my business around marketing campaigns, I have access to information I have in that line of business.

What about the growth there is of that same company in healthcare, pharmaceutical, and clinical? Not only you can't find that stuff, if you could find it, you can't find the people who know it. If you could find



those people, it's under different taxonomies. It's under healthcare, it's under patients.

We have to break that down so we can look at how do we allow a simple user interface? I actually like the Google Now analogy. How do you have a simple interface, from an analytical perspective, that inferences and guides me to where data is that I need that now, I can start shaping, and assembling, and storyboarding, then put in production?

Then, at the same time, also point me to people who are there so I understand it in a very simple, elegant interface. At the same point, also being able to not just do a pull, if I ask -- I have to already know the question to ask -- push.

If we could pull those together, which is very doable today, the user experience, all they see, is a simple, elegant interface.

Claudia: I love that. Like I said, it was an excellent session today. Unfortunately, we're out of time. We could probably go on for another 30 minutes. That's it for this addition of the "BBBT podcast." Again, I'm Claudia Imhoff, and it's been such a great pleasure to speak with Bill O'Connell of IBM today. Thank you, Bill.

Bill: Thank you, 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!