



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:	Datawatch
Date recorded:	November 17, 2015
Host:	Claudia Imhoff , President, BBBT
Guest(s):	Jon Pilkington , Chief Product Officer Dan Potter , Chief Marketing Officer Rami Chahine , Vice President of Product Management
Run time:	00:16:26
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.

I'm pleased to introduce my guests today. They are Jon Pilkington, Dan Potter, and Rami Chahine. Jon is the Chief Product Officer, Dan is the Chief Marketing Officer, and Rami is the Vice President of Product Management, for Datawatch.

Welcome to you all.

Dan Potter: Thank you, Claudia.

Jon Pilkington: Great to be here.

Rami Chahine: Happy to be here.

CI: All right, let me start with you, Dan. First of all if you don't mind, give me an update on Datawatch since your last visit. It's been a little over a year.

DP: Yes.

CI: You've done a lot in the year.

DP: We've been very busy on two fronts really. The first front is around self-service data preparation. Jon Pilkington and his developers have been very busy taking this 20-year Datawatch Monarch data-preparation engine and really extending it out to a much broader use case.

We've always been specialists at hard-to-get-at data—semi-structured, multi-structured content—extracting and preparing that. What we've done over the last year is we've opened up the data sources in which we support, relational sources, Hadoop big-data sources, other sources. There's many more sources that you can bring in and prepare.



We've also been focused on making it really easy for a business analyst to point and click their way to go from raw data to get it ready to be brought into tools like Tableau, or Qlik, or Datawatch, or Excel. On the back end we've worked with vendors like Tableau, and Qlik, and Microsoft, and others to be able to support the native export.

You're quickly acquiring the data. You're cleaning it up, preparing it, joining it with other information, and making it readily available in a format that can easily be consumed by those tools.

The second area that we've been focused on is our real-time visualization capabilities. We've made some great strides in the area of the Internet of Things, in particular industrial applications, so the applications that really require and value real-time insights, continuous monitoring.

An example would be in oil production where we're working with a major manufacturing firm doing fracking, and it's all about margin. For them it's all about cost containment and real-time insight to be able to make smarter faster decisions about things like steam injection. Those are some of the things that we've been focused on in the last year.

CI: It really shows, I have to admit. What I appreciated most about your BBT today is the fact that you started off, bam, right at the beginning with a demo. It was a demo of Monarch, your data prep tool. I found it fascinating. If you don't mind, I'm going to turn to you, Jon, since you did the demo there. Tell me a little bit about the features of Monarch.

JP: We're actually having quite a bit of fun with it. We decided to build an application that would be incredibly easy to use. That was the first tenet that we defined. Secondly, we wanted to get about the broadest reach of data that we could possibly get.

We took our 20 years of heritage of dealing with documents and being able to take those and automatically transport the data from a document into rows and columns. We added web scraping in so you can enrich your data with third-party web sources.

CI: I found that fascinating, by the way. I thought that was terrific.



JP: It adds a lot of agility for the end user. It sure does. Then we decided to take all the classic relational sources, big-data sources, and allow you to bring all this in and blend it together. One of the things that really does differentiate us is every step and every change that we make we record. Not only does it act as an undo/redo list. It acts as a load plan for our automation.

It acts as a log file, an audit file, for people that need to look at things from a governance compliance perspective. But, it also allows us to re-use all of those items to make sure that I get efficiency within the table that I'm using, within the work space I'm using, or with the entire automation script.

CI: I like that. The full lineage I thought was also quite good. A very mature, very sophisticated product if I could say so.

JP: Feel free.

CI: OK, I'll say so then.

JP: Thank you.

CI: All right. Well, Dan, let me go back to you because you have more than just Monarch. If you don't mind, tell me about the full offering from Datawatch.

DP: Our approach from a packaging perspective is to make it really easy to get started with Datawatch so you can do self-service data prep on your laptop or on your desktop in a stand-alone fashion, get started quickly. As you want to grow in scale you can share with other users who may be running on their desktop, or you can publish to a server.

It's the server capabilities that unlock a lot of the more enterprise qualities—high-scalability supporting thousands of users, the ability to fully automate those data-preparation routines and deliver prepared data to the right users at the right time. We also provide the ability to govern and manage those assets that are typically not managed.



In the self-service analytics world, you see a lot of people using Excel extracts as a source of information and obviously with our heritage with documents and other multi-structured content... we felt it was important to be able to store and manage those assets as part of the process.

Because we do that, we can also provide the governance to that: You can see exactly where that information came from. You can drill right down to the source and see if it's coming from a legacy ERP report. You can drill down to that line in the report and see where it came from. So, that's on the server for preparation.

On the visualization side it's very similar. You can start simply on the desktop. You can run the Datawatch Designer and start to create your own visualizations. You can directly connect to data in motion sources.

If you're doing an IOT-type application, you can connect to a Microsoft Azure, or a PubNub, or a PTC ThingWorx.

These are all emerging cloud-based infrastructure for gathering, and filtering, and providing sensor data. There's a lot of ease of use in starting on the desktop, and as you move to a server-based deployment, then you've got rich HTML5 applications that you can do this continuous monitoring, viewing, and easily disseminate this to thousands of users all using a browser, or a smartphone, or tablet.

CI: The other thing that I greatly appreciate is the transparency you have with your pricing model. You showed us exactly what it would cost. If you don't mind, go over that.

DP: It shouldn't be a secret. There shouldn't be a mystery to pricing. If it is, then...

CI: Something's wrong.

DP: Something's wrong. Again, it's easy to get started with Datawatch.

On the desktop versions of our products the data-preparation licensing is \$1,600 per user per year. It's a subscription model that gives you full update rights going forward.



On the visualization side, likewise.

It's a subscription licensing model. It starts at \$1,000 per user per year. Again, it's very simple to get started. As you start to move to a server-based licensing it's based on the modules which you license and the capacity of the machine.

The lowest price in the starting point is \$25,000 to start to automate those data prep and deliver the prepared data to different users through a browser or different ways.

CI: As I understand it there's an all-in-one package, and you can pick and choose what that is.

DP: Our approach is let's make it modular. If people want to start to add specific server functionality—the automation, the content management governance, the visualization—they can do it in an à la carte fashion.

We also have what's called the complete server. The complete server is the full range of capabilities that you can just turn on to all your users.

CI: The other thing for those who are not sure they want to spend money yet is what?

DP: The Freemium offering.

CI: There you go.

DP: We have a free version of Datawatch Monarch. We call it the personal edition, and it provides a very robust set of capabilities for free.

We want business analysts to be able to see the value of Datawatch.

We're very unique I believe in our ability to quickly go from any type of data to preparing that data and delivering it in rows and columns, really clean, to a Tableau, or Excel, or other visualization or advanced analytics tools.

And, the proof is in the pudding. We want people to download the free version. We want them to see that to get to value very quickly.



As a result people will start to say, "Hey, I want some of the more advanced capabilities..." and they are. It's been a very successful strategy for us.

CI: It sounds like it. Rami, let me bring you into the conversation. You gave us some customer examples. If you don't mind, talk about a few of those.

RC: Sure. A couple examples come to mind, and I'm going to talk about the customers that do actually leverage the most out of our platform. One of them is our customer, Time Warner Cable. We are being used within the finance organization. What they're dealing with are reports of their financial and billing statement that comes through a third party.

Those reports could be as large as 170,000 pages per report. They use our software to extract data out of those reports, build models with our desktops, upload it to the servers, have the servers then crunch the numbers with their Excel data, their structured data within their data marts... and then distribute the consolidated view of their whole finance billing to hundreds of accountants within the organization. It's also loaded within their data marts. It's around 600 million records monthly. They're using actually Tableau to visualize their billing data.

Then from there all their artifacts in these files that are generated with the Datawatch server are stored and governed within the Datawatch content server as well.

Another example that comes to mind is our OEM partner, Broadridge, that actually took our software and created a SaaS offering of it. They provide the service to a lot of brokers within 60 of their own customers where the brokers can submit their broker statements to their service. Those broker statements can then be turned into Excel data and sent back to those brokers.

Within that they're doing 300 million sources of processing data a month, so a lot of data is being processed through the servers. It's probably one of our biggest deployments. It's in tens of servers being clustered together to deliver that.



Again, the artifacts are being stored. They're one of our biggest lineage customers. They use the lineage feature to show if there's any abnormality in the data.

CI: Very good. Very interesting I have to admit. All right, Jon, back to you. You talked a little bit about challenges I guess that companies face and how your product supports these challenges. Let's start off with the challenges.

There were five areas, and you had some very specific information about how these are challenges for corporations. Why don't you tell me about them?

JP: Great, thanks. Data's ugly. It's awfully hard to deal with. I've been dealing with it my entire career. When you start getting into data from disparate sources, from different systems, and even just the cleanliness of that information, how do you quickly take that information and blend it together? How do you do it efficiently?

That's really what we're trying to bring in. As part of doing that the number-one issue that we see is untapped potential of data, which sits in documents. When you talk about documents, the question always becomes, "Well, how do you get information from a PDF document? Why wouldn't you just go to the source?"

The reality is data doesn't always come from inside the four walls of the organization, and you don't always have query capability to those sources. A lot of times what people don't realize, every single document is actually a date/time-stamped data mart.

Now as you start bringing that information together you can actually recreate things that you couldn't through a regular system.

For example, Salesforce. If you go to Salesforce and say, "What is my pipeline?" it shows you your pipeline today. If you say, "I want to see my pipeline on June 1st," how can you tell?

So, internally we take the snapshot of that report every day. It's a PDF, and we put them all in as data. Now we use our visualization tool...I can go



back on June 1st and see what was my pipeline. Nice, easy ways of dealing with information like that.

Then once you have this data about 80 percent of the time that you start dealing with it, it's on data prep. It's not on the visualization. It's not on the analytics side.

It's just getting this dirty, rotten, filthy, smelly data into a data set that's reusable. That's really what we're about. We want to make that fast. We have customers where we've had estimates of a small company that's about \$100 million, 80 people. They actually save over \$300,000 a year in data prep lost time using our product.

We have some larger organizations that it's a multi-multi-million-dollar problem for them and turn around as an organization. Then once you have all this data and once you've done it, you're going to need to do it again, you need to turn that into a recurring process. We've been doing that for years as well.

How do I automate that process and quickly now either turn it into my own personal data set, share it with others, or put it in the container where everybody else can access it? Then trading off risk for speed and getting information faster but still have some control over it.

Self-service analytics as a whole is a bit of the Wild West as we all know, but being able to track who made the changes, what the changes were, have an understanding of that lineage, and always to be able to go back and make a call to it. Really those are the things that we're all about.

CI: All right, Rami. Let me go back to you then. What's in the future for Datawatch?

RC: I think part of our future at Datawatch is to leverage our history. We've been in the market doing data prep for around 20 years. The data prep market is an evolving market with no actual leader yet evolving out there. We're moving fast in our product lines.

We're adding all the features that we've had in the product line, but we're putting in a concept of easy-to-use more-intuitive products. The future of



Datawatch is laser focused on becoming a large player within the data-preparation market, delivering products that are easy to use for the business users, products that are used on a daily basis, and also products that not only land on the business-user desktops but also can be used on an enterprise level with all of our servers.

Just pushing that knowledge that Datawatch has, and has collectively gathered from lots of use cases and 40,000 customers, into the market to be a thought leader in this space.

JP: I think that's well put. The way I see it, we simply want to own the space. We believe that anybody who works with data can benefit from using our tool from a person doing Excel all the way up to the rim of a data scientist.

We're going to keep advancing our capabilities extending out into cloud services as well, and allow people to get at information in any way that makes it easy for them to do.

Cl: I'll tell you. I am very happy with you guys. I see a very bright future for you. I think you've done an excellent job of presenting what you have, and I encourage people to watch the video coming out soon about your BBBT session with us. Unfortunately though we're out of time, so that's it for this edition of the BBBT podcast.

Again I'm Claudia Imhoff, and it's been such a pleasure to talk with Jon Pilkington, Dan Potter, and Rami Chahine. Thank you all for talking with me today.

DP: It's been a pleasure. Thank you, Claudia.

JP: Thank you very much, Claudia. This is great.

RC: Great to be here.

Cl: 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!