# 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:** | **Lyzasoft** |
| **Date recorded:** | **July 26, 2013** |
| **Host:** | **Claudia Imhoff**, President, BBBT |
| **Guest(s):** | **Scott Davis**, Founder |
| **Run time:** | **00:08:52** |
| **Audio link:** | Lyzasoft Podcast |
| **Transcript:** | [See next page] |
| **Transcribed By:** | CastingWords |

Claudia Imhoff: Welcome to this special edition of the Boulder BI Brain Trust podcast. I'm Claudia Imhoff and I'm here in Grants Pass, Oregon, at the Humphrey Strategic Communications annual Pacific Northwest BI Summit. I'm pleased to have the opportunity to interview each of the vendors attending this event.

CI: With me is Scott Davis. He's the founder of Lyzasoft. Welcome, Scott.

Scott Davis: Thanks, Claudia.

CI: We've had an interesting three days here at the Pacific Northwest BI Summit. One of the big topics, of course, was the cloud. The future of the cloud looks bright. We're very interested in what the cloud can offer. But one of the areas that we do still worry about is security in the cloud, and especially when we start talking about putting the entire data warehouse stack in the cloud. There are so many points of security breaches that could happen. Let's talk about those.

The first thing, of course, is getting the data in. The data has to travel to the cloud. It has to be integrated. It has to be put into a database. Where do you see security in that instance?

SD: There are some great advances that have been made in the last year or two, some good products that are out there now, to allow companies, for instance, to create highly encrypted VPN tunnels to their virtual private clouds. Amazon's made great strides to provide infrastructure options. When you provision a virtual private cloud, you can actually make it a single point of ingress, single point of egress.

You can put, for instance, Viada Surface. It's a great VPN provider. It can be inside the VPC and it gives you all of the control that you ever had to do, an IP point rejection, so that you don't have to have an open access.

You can do all the things you do with VPN in a traditional router environment. You can do between the data source, which is typically your production systems and say, the data receiver in the cloud, which might be a data integration platform or storage platform. Encrypting, getting the data in was job one. We've made great strides there.

CI: Let's keep going down the pipeline here. Once you've got it in, the next step, of course, is to allow the business community, the data scientist or whoever it is, the business analyst access to this information. Again, another area of potential security problems.

SD: Yeah, and we're seeing, again, the infrastructure is evolving to meet these needs. For instance, paying identity now has the ability to deliver single sign on and corporate identity in the cloud amongst many different components that you might use in a session. You don't have to constantly log in to your DI platform and then log into your storage platform and log into your reporting platform. You just have seamless migration across these different clusters that might even be on different platforms using federated identity. It's another interesting software app.

CI: That's nice because I think that's the complaints that many people have had is, "Jeez, I got to sign in here. I got to sign in there." Each of those is a point of weakness or could be a potential point of weakness.

SD: I think it's an important point to bring up, that the cloud has some important security advantages in certain circumstances. A great example is, you cannot hack a box that does not exist. We talk a lot about transience, for instance. The idea that you could, for instance, provision compute resources for a batch process -- which, for instance, almost all ETL is a batch process. You could provision what we call a transient cluster, bring it up, use it for a batch process, get the data off of that to another platform. Those boxes are then sanitized and decommissioned. They don't exist anymore, so they don't represent a security threat.

You have back up and think about, "Where are the vulnerabilities, and how do we address those things?" The cloud offers some interesting advantages, actually.

CI: The only other place is just the fact that the data is now outside of your firewalls. It's being stored in some remote location. Let's talk about that in a moment. The other problem that companies have had, especially European companies, is that there are laws that say their data cannot leave the country.

SD: There are some laws like that. I would say the most onerous one right now is that EU companies are reluctant to have their data inside the United States. For instance, AWS offers an infrastructure in Ireland to get around that. We can do exactly the same things in Ireland that we can do in Virginia or Seattle. For our European companies, we simply do that.

CI: Just move it closer to them.

SD: Exactly. When you store that data, there are also some interesting capabilities that are available. For instance, if you store data in Amazon's S3 service, you can choose to have that data on-disk, encrypted. It's not just encrypted in flight, it's actually encrypted on the disk. We've seen some interesting things, where the service provider managing the environment makes sure that that data is encrypted but provides the mate key only to the service user, so even the service provider can't decrypt the data that's on-disk.

CI: Does that cause any kind of performance slowdown? There must be a little bit of it.

SD: Sure. Encryption/decryption is always going to be a performance issue, but one of the benefits of the cloud is just that resources are not a problem.

CI: Yeah, just expand or contract.

SD: It's so, so cheap. We did a DI project for a company, and we used 800 nodes in this DI project, and we could do a month's worth of ETL work in six or seven hours for $12.

CI: Wow. Try doing that in-house.

SD: It hearkens back to the days of jobs and logs, when they were looking at how do we build a GUI and everybody else in computer science is saying, "My gosh, compute is so expensive. How can we waste it on this?" These guys are just betting that it's going to get cheap. We're at this same point where some things that we always assumed were really expensive aren't. You just need to back up and think about how to do things differently.

CI: What I find interesting is that what used to be a problem is fear of security. My data's leaving my clutches, has almost evaporated because of these advances. Now, I think the cloud has a tremendous future. I think people cannot afford to store it on premises. They really do need to rethink their entire paradigm of storage and access of data.

SD: I think that there are some circumstances in which the cloud doesn't make sense.

CI: Oh, do tell.

SD: Some people do need to think about it. A good example is, highly interactive, for instance, visual data discovery atop large data sets. You essentially need to have the user, the compute, and the data, in very close proximity because the transmit time for that data becomes a latency issue.

CI: That's a good point.

SD: I just think you have to look at it at a case-by-case basis but for things like transient batch processing, the cloud has some huge advantages for things like, we know a lot of companies that aren't using tape anymore. Use S3 as your tape. It's cheap. It's there. It's accessible everywhere. It's encrypted. You're not going to have any media decay. Do you know how many tapes we've recovered and found that the data decayed? Yeah, we have the data on tape but it's not really usable, so why are we doing that?

CI: [laughs] Like I said, I think it's a very, very bright future. I'm actually quite optimistic about what's happening in that area. I wasn't so much a few years ago, but now I am. I think there's a bright future. Thanks so much for speaking with me today. Again, it was Scott Davis, the founder of Lyzasoft. Thanks, Scott.

SD: Always fun.

CI: Thank you for listening to this special edition of the BBBT podcast, and thanks to Scott Humphrey for giving me this opportunity and for hosting the Pacific Northwest BI Summit.