



DATA VERACITY

VIDEO TRANSCRIPT

THE IMPORTANCE OF TRUST

Dominic Delmolino: Hi Gus. Welcome to our continuing talk about the Tech Vision from Accenture for 2018 and today we're going to be talking about this topic of data veracity. Which for me means this idea of data integrity, data that we can trust; data that's objectively verifiable and we can rely on it to make decisions. So, I know you're an expert in looking at this. You've considered this problem for a while and the challenges associated with data integrity. What does it mean for you and where do you see our clients going with it?

Gus Hunt: So first off, everyone talks about the '3 V's of Data' – that's been around for a long time, right? Volume, variety and velocity.... and veracity is actually been the fourth 'v' that's been out there for a long, long period of time, right? The ability to understand your data; where it's come from; have integrity in the data. Data integrity becomes absolutely the essential in this thing. Has my data been modified in any way, shape or form? You know, and if you really want to get down to it, a concept -- I would use the term 'correctness'. Is the data itself actually correct? And so, for the ability for businesses to really use data effectively, this concept of 'veracity' is absolutely essential.

DD: Sure, sure. So, it sounds like more than just ensuring the data I have is correct but not only that data's correct but maybe other data I'm starting to take advantage of; how does it compare? How do I balance and contrast and say, "Do I get multiple-source verification?" So, it sounds like it's gotten more complex than it used to be.

GH: It's gotten extremely complex. And in fact, that's sort of the good news and the bad news thing, right? Is that with technology today, we actually have the ability to do some of this compare and contrast that we could never have done before at scale with the volumes that are out there, right? So, that's the good news. The problem we get into I think with data and data veracity is that along with – as we said the integrity of the data, which I think is a nice solvable problem. This is where blockchain comes in and a lot of things are in encryption and hashing, signing and all the things that actually get assembled into this concept behind blockchain that's there. So, the ability to guarantee the integrity of my data, meaning that it has not been modified from the point of collection and delivery to when I get to use it, I think is a fairly solvable solution. I think the harder problems are around data bias in systems the thing is – I mentioned before – data correctness. Is the data actually itself correct, right? Again, these get really tough because bias creeps in on every side, every aspect. From the data we collect, to the questions we ask of the data are problematic. So, even though when we get in there, it really lends itself to needing to train the people themselves on how to use data really effectively within their environments. You hear what I'm saying? That's how you'll deal with that problem.

DD: So, I want to go back, you mentioned bias data. First off, how does that harm what you're trying to do? And how do you recognize the bias there? How do you address correct it? You know, deal with it. How do you deal with lies?

GH: You know it's interesting. Bias is the single, hardest problem that we deal with in data and because bias creeps in at every single level across the board. And so, this is where, again



training of people on the systems... transparency. So, in other words when big data analytic systems or AI systems or machine-running algorithms compute to an answer, we can't accept those answers blindly. We have to have the ability to say, 'Show me how you got to that answer?' so I can dig through it and actually see what transpired. And so, that's that transparency that also then combined with this other thing I call 'assurance' right? You know, this is part of data veracity. So you, I want assurance when I'm online that you are you. You see what I'm saying? I want assurance that the data has integrity. And then from there, I can make decisions about how much I want to invest in trust, right? So, trust being this objective outcome. But, with trust to do trust requires this level of transparency as to how data has been used. You know, I think you, provenance and lineage of pedigree, and all those things. Where they came from; how can I follow those things? And all that together is what will allow me and let the human understand what has transpired and see into it and then they can make judgements about bias.

DD: So, over the years we've had master data management; data governance activities. It sounds like there's a new emphasis on this trend. If I was going to get started as an agency, what's the one unique thing I should probably pay attention to that's maybe different than what I've done before?

GH: I would say, that as an agency the single, most important thing you want to be able to do, is to at least have data integrity. Period. No exceptions. That is, because the ability to suddenly modify data is a very pernicious thing and it's the new form of cyber-attack. When your data won't compute to the answer you were expecting or gives you wrong data because someone has suddenly modified it, that's a really, really bad thing.

But if I have data integrity, meaning I really do apply blockchain and things we've talked about, so I'm sure my data is what my data is. Then I can take all of these next steps to begin to make other decisions around it. Without that, the rest is useless.

DD: Okay great. Sounds good. Well, thank you so much.

GH: Yes, I'm delighted.

Copyright © 2018 Accenture
All rights reserved.

Accenture, its logo, and High
Performance Delivered are
trademarks of Accenture.