



Data trends



Guy Holmes
guy.holmes@katalystdm.com

My version of the truth

I was on a flight this week and, as usual, after wearing out my finger on the call bell seeking more red wine, I turned to the inflight entertainment to absorb as many documentaries as I could. It makes me feel better to say I watched documentaries rather than movies.

On this flight there was a great documentary on the Hubble Space Telescope, one on the great comedian, actor and director Mike Nichols, and a range of others. But the one that really caught my attention was called '(Dis) Honesty – the Truth about Lies' by the behavioural scientist Dan Ariely. This documentary looks at how and why people lie using data from various scientific behavioural studies carried out on a global scale.

It really got me thinking about data, how we create it and the ways we and others form opinions based on data. Can the same data tell more than one truth? Is anything ever categorical, or can two very different conclusions be drawn from the same data? If you can draw two different

conclusions from the same data what factors play a role in directing people to one conclusion over another? Does the fact that we deviate from the hard data to tell our version of the story really constitute a lie?

Seismic data analysis is the perfect arena in which to look at how the same data can tell two (or in fact a lot more) different stories, and how each story can lead to very different conclusions. We all know geophysics is not a perfect science, and that an 'interpretation' is just that - someone's view of the data based on their opinion, experience and, of course, all other related data they have used in current and past projects. But the fact that two very experienced geoscientists can reach often widely different conclusions using the same data should be a concern for everyone. What can lead to this?

Here are some contributing factors:

1. Geoscientists are infamous for being protective about their data. Many geoscientists simply don't like to have their work criticised by anyone else, least of all by another geoscientist. This can lead to selective reasoning or the selective sharing of facts that one party knows will support their conclusion rather than detract from it.
2. Most of our software tools are designed to try and support the story we want to hear. Take this word processing software I am using to type this article right now. It certainly will highlight spelling mistakes or grammatical issues it detects in the sentences, but it offers no opinion on the quality or accuracy of my story. Essentially I can type whatever I want into this software and, as long as I don't make a spelling mistake or a grammatical error, it will happily allow my drivel to

be published and form popular opinion (well probably not that popular).

Geoscience software can be very similar. Feed it a list of values and the software might check that the values are reasonable for the type of data you are analysing and ensure that there are no formatting errors in the data (letters where numbers should be etc.). It will then produce a result. However, what if that list you uploaded was a list of your last 300 top scores in Candy Crush instead of the gravity measurements over a prospect?

3. In the world of big oil, teams of people often work on individual parts of a larger science project. All of the various results are then brought together to create a 'final' result. Once that result is agreed upon it can become very difficult to challenge and, ultimately, it may become the *de facto* truth. Imagine eight or ten scientists working together to create a result, and two or three of them using a little creative license with their data to ensure they meet the team deadline for the drilling project. Just think about how that could change the trajectory of a result! Some of you don't have to then imagine drilling a duster and trying to reason with management as to how you arrived at your result to drill in the first place.

None of the above really constitutes lies, but the after effect of them often leads to having to find a more convenient truth about how we came to our conclusions. In essence, as per a quote from Chris Jami – 'Just because something isn't a lie, doesn't mean it isn't deceptive'. The irony about the geosciences is that we often reach a conclusion with the support of data, but not necessarily a conclusion that supports the data. After all, I'd be lying if I said geoscience was about the facts!