We hear a lot of about Big Data, and a lot of people claim many different things about it. I am so sick of all the hype, that I wanted to inject some sanity into the mix.

**Big Data is New!**
Working with large volumes of data has been possible for a long time, and it might surprise you in what different areas that large volumes of data have been used. For instance, back in the 1990, before the time of big data. The Commodore Amiga 3000 was used to make the worlds first 3d virtual world gaming environment. I went to the local RSL club, and you paid your money, and you could go into battle against another player, while you where encased inside a gaming cage. The volume of data and calculations used in gaming machines, pre the birth of big data, is in some cases much greater than the big data used to solve problems today. Because of the power of these gaming systems, and the realization of their processing capabilities, pattern search software is being written into their code, and these machines are now mining big data datasets.

**Never before in human history have we had so much data.**
What a load of rubbish. Firstly, lets look at collecting and storing data. Mother Nature has been collecting and storing data since the beginning of time. For example, tree rings, sand stone, background radiation, ice cores, star light, etc. They all contain data about events that occurred in the past. When we look around ourselves, we see colours, sizes, shapes, materials, etc. Every moment we are alive, we are processing and producing data. There has always been great volumes of data, and there always will be. What is different these days, is the way we deal with Big volumes of Data, and how we record it, etc. While we process it differently, value it differently, and use it differently. It is not Big Data that has changed or suddenly appeared, it is our use and realization that has changed.
Are you a big data fool?

Big Data has to be greater than 1 GB, to be Big Data. I really am yawning with excitement at this one. Big Data is relative to the task at hand, and the resources to do it. If I were to sit you down to do keyboard entry, and ask you to type in these 20,000 Australian farm addresses. And you had to type in every one of them, then that is Big Data. If I ask you to go and buy a commercial produced CD list for $65, and uploaded it into the database, that is not big data. But it is the same data, but a different effort goes into it. Of course big data sets to a corner shop owner are different to a big data set for a multinational. That does not mean that relative to scale and resources, that the same amount of effort is not required.

My End Bit

It does not matter if its called Big Data, All Data, or Broad Data. Large data sets are going to be useful to people and companies. We are in an interesting time; we can deal, explore, take advantage of, and connect with other data sets. But please be aware of the sales pitch hype, Big Data is now new, nor a trend, nor a fad, etc. It has always been, and I recommend that you find talented people to employ as you can, seek out both commercial and open source software options, and use them together.

Thus my Rant ends.

Published by

Tony Nolan OAM
Data Scientist
Published • 3y

Reactions

3 Comments

Add a comment...

Don McDowell • 1st
Consulting Expert - Strategic and Advanced Intelligence for Europol

So refreshing to hear the voice of sanity and, moreover, REALITY, for a change ... thanks Tony. I am sick and tired of hearing people in the intelligence, risk and analysis game harping on about how technology has all the answers and is replacing the need for human input. Indeed, the plenary conference session in Warsaw in Oct 2014 was the subject “Technical Analysis versus Human Analysts” with its governing premise being that we humans were becoming obsolescent. Related to your …see more

Like Reply

Achievement Heights Academy and STUDIO
Dr. N’Quavah R. Velazquez | CEO, President and Founder

Great post that holds big data to the light. Similarly, enjoyed your comments Don McDowell on data mining, human thought/analysis.

Like Reply