The Temptation of Data

Everyone wants happiness, no one wants pain.
But you can't have a rainbow without a little rain.

Not sure where that's originally from but it sums up BigData.

We as individuals are constantly instantly detectable by a plethora of data points. Every time your smart phone shakes hands with a cell network or connects to a Wi-Fi anywhere, a "smart" anything making a connection (like your fridge and your thermostat "talking" to each other over the home network), every credit card transaction and cash withdrawal, Facebook interactions, wearable medical devices that broadcast your health and location, websites and their cookies, RFID tags on things you've bought, driving on an electronic toll road, walking past a camera with face recognition technology, the Internet of Things … no man is an island.

Likewise, every business has its own set of data points that also feed Big Data.

Big Data describes the volume and velocity of structured and unstructured data points. If it can be identified and reduced to a number, it's part of Big Data. But it's not the raw amount of data points that matters, or their speed or granularity. What matters is what organizations think they can do with the data.

My views on the collection of Big Data slant towards the libertarian: each of us should have the near-absolute right to control how our own data is collected and used, even if most of us don't exercise that right. But we'll save that thought for another day, and instead look at the rainbows and the rain of Big Data.

Data is a succubus in digital lingerie just filmy enough to obscure the good parts. The Dance of the Digital Seven Veils!

What every Big Data analyst wants is a better view, a glimpse of the goodies. The techies, the marketing wonks, policy advisors, mutual fund managers, they all get seduced thinking, "Oh the wonders I could create if only I had data that told me how to ….". It's that tantalizing seduction that is the rainbow of Big Data.

There is no shortage of companies analyzing your network usage to deliver you targetted advertising, You google for a local guitar store, suddenly Gibson and Addario ads pop up in your browser. Waze delivers updated customized advertising depending on where you're driving. Walk into a mall, retailers send you up-to-date offers. And on a larger scale, companies anonymize, aggregate and sell a broad array of data collected from billions of data points. Forensic data analytics work on the data to tell what happened and perhaps why; predictive analytics work from there and try to extrapolate to when that event can happen next.

One unusual application of Big Data comes from LeoNovus Inc. (TSXV: LTV). While LeoNovus' website uses all the expected hot-button buzzwords ("creating the infrastructure needed for next-generation geo-dispersed distributed data centers that will enable cost efficient, highly reliable, secure cloud computing"), what the business plan really entails is using other companies' existing infrastructures (servers, pipe, spare CPU cycles) to create distributed data centers. (As an aside, it would be very hard to find a deeper more reputable management team in any start-up. If you believe in betting on jockeys rather than horses, Leonovus could be for you.)

This phenomenon isn't limited to the online world. The latest edition of **Mining Trends and Development (Fall 2015)** has an article on *How Analytics is Transforming Mining*. If there's any industry that could use an efficiency boost, it's the extractive industry. IBM's Dirk Claessens (GM, Global Industrial Products) uses the example of data points being extracted from a failure of car dumpers at a port, to

determine the cause of the failure and how to predict / avoid another failure.

This example shows the problem with Big Data: it is soooo seductive. Answers swirl like a veiled dancer, just out of perception:

- our latest burger didn't sell well in St. Louis. Why not? I don't know, let's get more data
- let's get more data to see how we can reduce overhead while maintaining service levels in our used car lots
- can we hedge the Canadian dollar against copper? I don't know, let's get data and see
- Right now, this second, someone from Vancouver is using a credit card in Bogota to buy a three thousand dollar emerald: legitimate or fraud? Allow it or deny the transaction?
- Even *The Globe and Mail* is chasing data, with this weekend's edition telling investors how to mine data for historical dividend information

It is this seduction, the temptation of data, the fantastic potential returns, that leads to data's two largest weakness.

First, we are entranced into believing there is an answer to any question, we need to believe there *has* to be an answer, and so it is possible to waste weeks and months chasing that which can never be caught. This leads to the inefficiencies that Big Data was ironically supposed to eliminate.

Second, as users and targetted demographics, we are willing to concede on protocols, security, data access points, passwords … just as long as we get our processing hands on that data.

This isn't about the generational perception of privacy, the role of government, cybersecurity insurance, cloud computing or its progeny fog computing; it's about the rain that goes with the Big Data rainbow. And here are, in my opinion, two of the largest data risks facing companies through the next

couple of years.

The first is BYOD. That's short for Bring Your Own Device. I use an Android phone. The CFO uses a Blackberry that is two generations of technology old. The VP Exploration in the field uses a current Blackberry and a crappy old Linux laptop. The CEO is blissfully unaware of any of it, she just dials the phone, thumbs the email, and hopes for the best.

In this scenario, there is no central policy governing how users are supposed to protect the corporate data on the disparate devices. If the VP Ex leaves his unlocked Blackberry in a taxi in Manila, the potential repercussions to the corporation are massive. The corporate intellectual property has been lost, perhaps forever. Even if it was locked, that won't stop a skilled hacker for long.

And what happens to the corporate data if that VP Ex becomes a disgruntled ex-employee: does the data stay on that personal phone? When the VP Ex decides to upgrade and sells the old phone to an unknown third party? When the CFO doesn't install the latest security upgrade, making the data vulnerable to malware and hacking?

And imagine the corporate problems if that same lost Blackberry, sanctioned by the company, turns out to host child pornography pictures.

BYOD is a massive problem. Every company should have a corporate policy addressing it.

The largest problem, though, through 2016 and 2017, will be CASL, or Canadian AntiSpam Legislation.

Pernicious: causing great harm or damage often in a way that is not easily seen or noticed.

CASL is by far the most pernicious legislation I've ever seen, making it a prime example of the Law of Unintended

Consequences.

Every business email you send, every business text message, every Facebook message for business, every cookie on a web browser, every software upgrade for your computer, tablet or phone: they're all covered by CASL. The law as it now reads is that you need someone's consent BEFORE you send that person a business email. A failure to be able to prove you had that consent gives rise to regulatory liability.

This isn't merely theoretic. The Canadian Radio and Television Commission has levied significant fines or extracted significant settlements for breaches of CASL. The latest was last Friday morning when Rogers Communications, with its A common shares trading at around $52 per and with a market cap of almost $6B, was fined. How much was Rogers fined?

Two hundred thousand dollars. For a breach of CASL.

The reason why Rogers was sanctioned? Corporate emails did not always have a fully functioning "unsubscribe" button. I'm not making this up. Here is the link to the government of Canada's website and the Rogers decision.

On that same webpage is a reference to the June sanctions against Porter Airlines. Porter was the subject of a $150,000 attack by the CRTC. What heinous act did Porter perpetrate upon Canadians? Porter could not prove that all of its email blast recipients had consented to receiving it, and the unsubscribe feature was hard to find.

Again, in case you doubt it, here is the link to the CRTC's decision.

This is staggering in its implications. I'd be willing to bet that 99% of you reading this are NOT in CASL compliance. That makes you and your role in Big Data a possible next target for the ambitious and aggressive CRTC.

CASL came about in part due to our love of BigData, and our desire to chase the answers that it might yield. Now we're paying for it.

We will be looking at CASL in more detail in a future article, but for now realize that you need a corporate policy on this. CASL is not an IT matter: it's the duty of the CFO, General Counsel or Chief Compliance Officer.

And that person must balance the siren call of Big Data against proper corporate compliance.