Big Data in business – father learns of teenage daughter's pregnancy from retail chain

A checklist for increasing and changing legal issues for Big Data -a major issue for corporates

February 2013

What's "Big Data" and why is it huge?

A great summary of what's happening for Big Data and business (ranging from finance and retail to airlines and health) is in an October 2012 Harvard Business Review article, *Big Data: The Management Revolution* [1]. Big Data moves on from analysing smaller sets of data, toward gleaning business intelligence and advantage with three key differences, the three Big Data "V"s:

- **Volume**: vast quantities of data can be mixed and matched.
- **Velocity**: All this can happen real time. HBR give an example of, on the busiest pre-Christmas shopping day of the year, Big Data allowing stores like Macy's to know their likely sales even before customers enter the stores. How? By tracking, by GPS, people's mobiles when they arrive in Macy's car parks. That's just the start: Macy's could track customers around the stores, fulfilling their needs according to profiles from huge databases. And much more.
- **Variety**: Big Data can be sourced, mixed and matched, from multiple places. Google's combined use of data sourced from all its platforms – from YouTube to Gmail - is one example: an example that shows the potential legal risks around mixing and matching data, with *reports* [2] coming out this month that European privacy regulators are looking to take action against Google on its policies around such use of data.

So what's the problem?

With the large commercial benefits of big data come some potential downside to be managed by senior managers, boards, IT departments and lawyers. While contract, IP and international issues are important - and dealt with below - privacy, data protection and security are likely to be major concerns to handle.

School aged pregnant daughter

For a good example of how big data can be challenging, there's the New York Times article, *How companies learn your secret* [3]. The article focusses on the retail chain, Target, which sells products from lawn mowers to baby clothes. Target analyses much of its data, such as purchase histories derived from sales on customer loyalty cards.

Big Data in business - father learns of teenage daughter's pregnancy from retail chain

The article reports an angry man demanding to see a Target store manager about his school-aged daughter getting Target ads in the mail for baby clothes and cribs. The manager rang a few days later to apologise. Says the article:

> On the phone, though, the father was somewhat abashed. "I had a talk with my daughter, he said. "It turns out there's been some activities in my house I haven't been completely aware of. She's due in August. I owe you an apology."

Target's use of big data

Using predictive analytics, Target are able to work out whether a woman is pregnant early on in the pregnancy, and can even predict the likely birthdate. It's normally very hard for retailers to break buying habits (such as changing a customer from one retailer to another or from one brand to another). But that's more likely during a major life event. If the retailer can target the mother as early as the second trimester, it can get in before others.

A Target employee gave an example: if in March a woman, hypothetically Jenny Ward, buys cocoa-butter lotion, a purse big enough to double as a diaper bag, magnesium supplements and a blue rug, there's an 87% chance she is pregnant and due in August.

Then, how do the marketers use that data? Sending out a flier saying, "Congratulations, Jenny, on your first child due later this year", will meet a hostile audience. She'd think that Target is stalking her.

Even a mailed catalogue devoted only to baby products can spook the mother. As the article notes:

> ...for pregnant women, Target's goal was selling them baby items they didn't even know they needed yet.
>
> "With the pregnancy products, though, we learned that some women react badly," the [Target] executive said. "Then we started mixing in all these ads for things we knew pregnant women would never buy, so the baby ads looked random. We'd put an ad for a lawn mower next to diapers. We'd put a coupon for wineglasses next to infant clothes. That way, it looked like all the products were chosen by chance.
>
> "And we found out that as long as a pregnant woman thinks she hasn't been spied on, she'll use the coupons. She just assumes that everyone else on her block got the same mailer for diapers and cribs. As long as we don't spook her, it works."

So this gets across the line by disguising the real agenda. Same use of data: just disguised.

OK. But of course this happens all the time: it just looks like Big Data will expand it exponentially.

Wider implications

Expand this story out to a wider range of information, sharing of data between companies, and so on, and it can be seen that organisations face legal and reputational risk on a grand scale. For example, one legal commentator [4] highlights the real prospect that anonymised data across massive and combined data bases can be de-anonymised by sophisticated analytics. Anonymised data provided, seemingly in compliance with typical privacy and data protection legislation, can be at risk. That's a legal and reputational issue.

Many of our *articles* [5] have been devoted to privacy and security legal risk. Security experts are pointing to a new set of greater and different security issues raised by big data, not easily handled by traditional security measures. Add to that the shortage of expertise in the big data area and there are risks for organisations.

Big Data in business - father learns of teenage daughter's pregnancy from retail chain

Checklist

Big data analytics will lead to increased sharing of data between companies. That adds to the things to consider, not only as to privacy and security, but also as to protection of valuable IP in data bases. Here's a checklist of issues to consider.

Big Data legal checklist	
Privacy, security and confidentiality	• Privacy/data protection legislation • Confidentiality law • Best practice security • Where relevant, appropriate stakeholder opt-in to use of data • Contracts with external suppliers and customers
Copyright	• Protect IP rights in databases (e.g. limit use of database by third parties to specific purposes; try to retain IP in databases and information created and derived using the database). • Contracts with users of data such as other companies (covered in our upcoming article) • Dealing with IP rights the database provider has.
International	• For example, international issues from cloud computing • Off-shore use of data
Contract	• Issues above • Licence scope and limits on use as to supplied databases • Limit liability • Payment and price • Rights in commingled data (i.e. data that has been mixed with other data). • Rights on termination (as to database, commingled data and information prepared from data).

1. http://hbr.org/2012/10/big-data-the-management-revolution/ar/1

2. http://www.globallegalpost.com/global-view/eu-watchdogs-target-google-privacy-policy-42805163/?utm_campaign=D1_19_02_13&utm_medium=email&utm_source=newsletter_d1

3.http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html

4. Paul Ohm, Broken Promises of Privacy: responding to the surprising failure of anonymisation (2010) 57 UCLA Law Review 1701. See also his interesting Harvard Business Review blog, *Don't' Build a Database of Ruin*, 23 August 2012

5. http://www.wigleylaw.com/assets/Uploads/MSD-kiosk-debacle.pdf

Wigley+Company
PO Box 10842
Level7/107 Customhouse Quay, Wellington
T +64(4) 472 3023 E info@wigleylaw.com
and in Auckland
T +64(9) 307 5957
www.wigleylaw.com

We welcome your feedback on this article and any enquiries in relation to its contents. This article is intended to provide a summary of the material covered and does not constitute legal advice. We can provide specialist legal advice on the full range of matters contained in this article.