



How to gather snowflakes: big data, AI and predictive analysis of customers

15 NOV 2017

By: Gareth Stokes

Ask people to think of something unique, and it doesn't take long before they mention the snowflake: fragile, perfect, utterly individual and completely different from every other snowflake that has ever fallen.

And surely we're all unique too – isn't it true that no one else has exactly my combination of needs, interests and supposedly secret desires?

But snowflakes – for all their individuality – are still fairly predictable. Every one will share common features: they form from water vapor under certain temperature and pressure conditions; they are always six-sided; they will all melt and re-freeze under the same circumstances. Mathematical models can even predict something as seemingly chaotic as an avalanche.

So if my buying patterns make me a unique precious snowflake, does that mean I'm completely predictable?

Who owns customer data?

Any business wants to have closer contact with its customers. Being front of mind when the customer is making a buying decision is by far the best way to drive more sales. There are many ways of doing that - classic advertising, sponsorship of events, creating useful content that demonstrates your expertise - but the very best opportunity to prove your worth to your customers is when you're already in contact with them.

Integrating the data received via those various channels and effectively exploiting that information to derive insights and make strategic sales decisions is the equivalent of 'controlling' such a data blizzard.

Every time the customer asks for a quote, or browses your website, or puts an item into a virtual shopping cart but doesn't then complete the purchase provides an opportunity to engage with them and gather data about what they're doing. With the right tools, and the right communication strategies, this data can reveal new insights and create new opportunities with that customer.

However, you can only do this if you have that customer data in the first place. If your business is interacting via channels that it controls, such as its own shops or websites, then you will have access to lots of data sources: not only completed purchases, but enquiries and analytics about things like browsing habits, viewed but not purchased items and abandoned online shopping carts.

Increasingly, though, customers are interacting via third-party channels, not all of which are prepared to share all of this data. If your products or services are purchased via an intermediary - an online marketplace, broker, aggregator, app store or "zero UI" personal assistant - then the real owner of that customer, and the associated data, is that third-party channel provider that gets the data. They "own" the customer and are depriving your business of potentially valuable insights.

This customer "ownership" also creates a significant and more immediate commercial threat. If the intermediary strikes a better deal with one of your competitors, there is little you can then do to stop that intermediary funneling those customer opportunities to your competitor instead, which runs the danger of leaving you 'out in the cold'.

Big data and tackling the four Vs

Assuming that you can get access to this rich lode of customer data, the next challenge is being able to use it to develop meaningful insights. A single snowflake might be beautiful, but in its millions it becomes a snowstorm, reducing visibility to a complete whiteout.

You may have access to vast amounts of analytics, social media feeds, customer purchase history data and any number of other potential sources. Creating a system that can use these sources to deliver new insights and drive additional opportunities is far from simple.

The Big Data challenge is often framed as four Vs:

Volume: the simple fact that there is a vast amount of data to be collected and analyzed

Velocity: both the speed with which the data is fed into any system designed to analyse it, and the need for answers to be delivered quickly, often effectively in real time if they are to be usefully presented to a customer while the customer is actively engaging with you online

Variety: purchase history data, analytics data, social media feeds and responses, store visits, instore video analysis... the types of data available for analysis are extraordinarily disparate

Veracity: both the "truth" of the data being fed into the system, and the ability to derive truths from it. Is the data complete? Are there hidden biases or self-selecting samples that need to be taken into account? Can false user accounts, deliberate social media trolls and the like be filtered out?

All of this presents a data analysis challenge no human can sensibly undertake unaided. Instead, a system capable of ingesting the data and performing much of the analysis by itself is required. From an infrastructure perspective, access to cloud solutions avoids the need to build and manage the high-bandwidth networks and servers to process the data, but that still leaves the problem of the analysis itself. The rise of various forms of artificial intelligence - often marketed under such labels as machine learning or cognitive computing - provides an answer for this element.

Most businesses will not have all of the necessary expertise to implement such a project in house and will need consultants and/or systems integrators to design and implement the solutions. Even where the expertise is present, there will still be a need for cloud contracts and software licenses. All of these contracts need particular care. If they are intended to have a transformative impact upon your business, getting them wrong is not an option. When drafting AI-related contracts, the parties need to realize that typical precedent contract terms (such as liability limitations,

termination provisions or service levels / service credit regimes) have been developed to protect against human error. They will not anticipate or protect fully against the types of failure modes that wholly machine-provided services might suffer.

Getting it right - not just a technology problem

We are already used to seeing "customers who looked at this were also interested in..." recommendations on websites. These fairly obvious attempts to drive additional sales are just the tip of the iceberg.

With the correct AI-driven analysis systems, all those companies we buy from - retailers, service providers, banks, insurers - can know us better than we know ourselves. Based on the fact that 1,000 consumers have followed similar purchasing patterns before us, they can predict your follow-on purchases with precision. Notoriously, this includes spotting purchasing trends that may indicate a significant change in personal circumstances - such as noticing a pregnancy earlier than actual family members have, due to a small change in shopping habits.

But simply having data and technical capacity doesn't necessarily mean you should be performing certain analyses. There are both legal reasons (arising from duties under data protection and privacy laws) and customer engagement reasons (being perceived as invasive or sinister) for restraint.

In the European Union in particular, the new General Data Protection Regulation aims to curtail the worst excesses of intrusive data analysis. Given the global nature of trade, retail and Internet-delivered services, we cannot assume that everyone will adhere to the same high standards. However, those players that don't will soon be caught out. There is no shortage of examples of businesses whose bottom lines were seriously harmed in the wake of widespread negative publicity over data (mis)use.

Next steps

Anyone who has ever gone tobogganing knows that when you're dealing with a very large volume of snowflakes it is easy to take a tumble and end up bruised. If your business is not taking full advantage of the data at its disposal or is embarking upon a big data or analytics project, get in touch with your usual DLA Piper contact to find out more about how we can help.

Our technology, data, commercial and other lawyers often work closely together to support our clients in addressing the challenges, and seizing the opportunities, that big data, AI and predictive analysis present enterprise as it seeks to more effectively manage its data for the benefit of its customers and those wider audiences whom it looks to engage / attract.

For more information about issues raised in this article, please get in touch with the author below.

Keep an eye out for our upcoming TechLaw Series events which we will be regularly updating.

AUTHORS



Gareth Stokes

Partner

Birmingham | T: +44 (0)20 7349 0296 [UK Switchboard]

Gareth.Stokes@dlapiper.com