Research excellence

Mind and the machine

TNS

It may be immense, fast and mind-bendingly varied. But researchers must remember: Big Data can no more speak for itself than the smaller sort.

Mind and the machine

Big things can be intimidating. Research cannot allow Big Data to be one of them. We stand on the edge of the most exciting and transformative period in our industry's history: 90 percent of the data in the world today was created in the last two years and "data taps" such as mobile, social and POS will continue to pour out raw information for us to work with at an ever faster rate.

However, if our response to the new era of data is to retreat behind number-crunching technologies, then clients and indeed humanity as a whole, will be much the worse for it. It may be tempting to conclude that human intuition must surely give way to computers and algorithms when it comes to keeping up with Big Data. But now, more than ever, we need to recognise the immense, unique power of our own minds when it comes to dealing with information – and deciding how to act on the basis of it.

90 percent of the data in the world today was created in the last two years

Share this f 𝕏 in

Mind and the machine
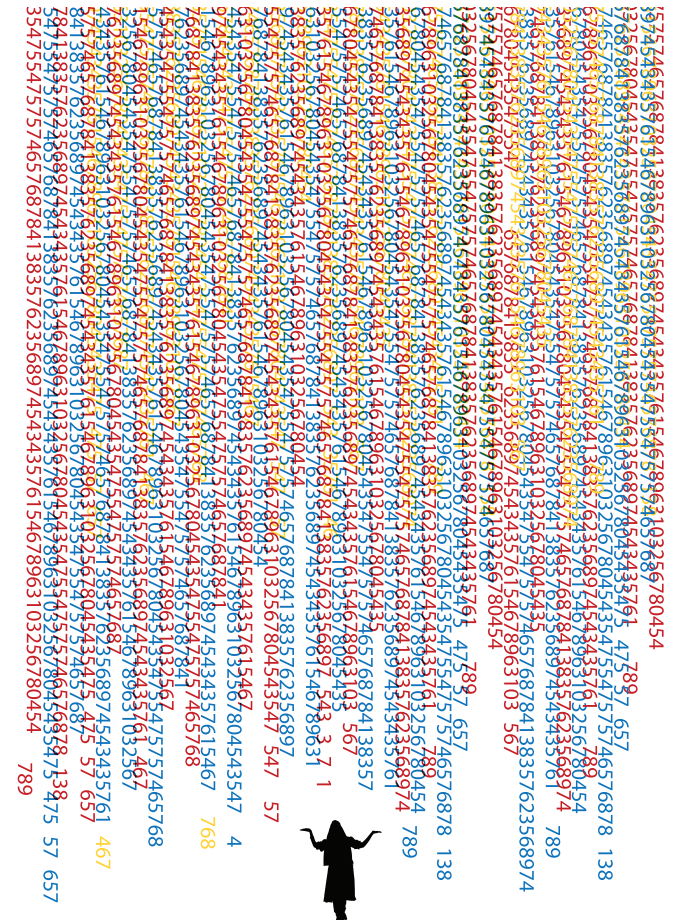
So what do we mean by "Big" exactly?
Big Data wouldn't be half as intimidating if it were just a question of having more numbers to deal with. But Big Data is bigger than that. It represents the coming together of several different themes, each of which would be fairly paradigm-shifting in its own right.

First of course, is the sheer scale of the data now being produced and stored. Walmart currently handles more than 1 million customer transactions every hour, in databases estimated to contain more than 2.5 petabytes. Such an organisation may soon have created more data every hour than research surveys have ever delivered. With data storage doubling every year, there appears no constraint on the amount of information that we are dealing with.

Connected to the size of data but equally significant is the fact that it now generates itself. Data no longer needs to be created through a questionnaire carefully crafted by a researcher, or painstakingly collected by a field agent; it is created and stored simply by virtue of things happening. It's broken free of human control – and therefore isn't limited as to how big it can get and how fast it comes at us. Data's Velocity, the speed at which huge volumes of it can be generated, is every bit as breathtaking as its sheer size. And the speed with which it is available raises the opportunity and the demand to work with it in real-time.

Yet perhaps the most challenging shift of all is that this size and speed is combined with an explosion in variety of data forms. Big Data comes in all shapes and sizes. Researchers are leaping on new types of data source – and new types of source are leaping on us: from mobile activity to Twitter feeds, geo-location information, facial expression capture and much more. We are quickly moving from dealing in numerical scores to dealing in shapes, movement patterns, expressions – and human language. And such data does not come readily packaged for analysis; using it must involve translating it as well.

Mind and the machine

You created it: you deal with it

Faced with such challenges, it's tempting to believe that computational power, which has taken the lead in creating this new world of information, must also take the lead in defining how we deal with it. In this view of the world, the researcher starts to look less like a person, more like supercomputer in a bunker: one where we simply have to feed in the right question or combination of questions, plug it into the river of Big Data – and wait for the answer to pop out. But there are significant dangers to this approach. If Big Data ends up becoming processed and commoditised data, then we are all in trouble.

Digesting really raw data

It's a mistake to believe that data can ever speak for itself. Data always speaks with a human voice; it can't say anything otherwise. Every statistic that we deal with is the result of subjective judgement about the problems that we should try to solve, what we think the answers should look like, and what data forms we can enlist to help provide those answers. And these judgements are human ones.

In the Big Data era, the human imagination continues to play an essential role in envisaging what our many different data sources can be made to do, and in aggregating, translating and coding them to enable them to do it. To take a very simple example, Google can predict a flu epidemic by spotting spikes in sea ches on cold and flu emedies. This is a tremendously cool thing, but it only works because somebody realised that this pattern is significant – and that it correlates to something meaningful and useful. Similarly, micro-location data gives TNS a powerful new tool for mapping movement around stores – but it is only powerful because we have established an understanding of what these movements mean.

In his book The Signal and The Noise, US election poll guru Nate Silver devotes a chapter to global warming and the fact that it would be impossible to find any evidence of this in the notoriously unstable climate record, were scientists not armed with a theory telling them exactly what to look for – and which data to prioritise. It's an important reminder that...

...the **bigger** data is, the more it needs to become articulate.

Mind and the machine

From data creators to data curators

In the old days (of six months ago), the raw numbers that we sat down to analyse weren't really raw at all; they were shaped by human hands even before they came into existence. The art of designing a questionnaire involves finely balanced judgements on which questions to ask and how to ask them. Whether to score preferences out of five, seven or ten can trigger some pretty serious debates with good reason – these things have a big influence on our ability to spot patterns, make connections and provide meaningful insight. And judging how to ask questions should, of course, be closely related to the challenge of what you are looking for.

In the Big Data era, we are no longer data creators, designing the structure of information from the outset; instead we are data curators, working with information that has been generated independently. As such, we will face many new challenges and require many new skillsets. However, as we evolve the role

cueators

of research, we must continue to apply the same standards to independently generated Big Data that we would if we had created it ourselves. And this will require leveraging much hard-won experience about how data works. The skills that once went into the design of research instruments such as questionnaires will remain crucially important in aggregating and selecting data sources, and deciding exactly how they relate to one another. For now, this might involve incremental improvements such

as linking spend and retention data to customer experience surveys, as we already do at TNS. In the future, we will find mo e and more scenarios where the data we aggregate does not include traditional surveys at all. In all of these contexts, it's not just a question of being excited about what data can do. It's equally important sometimes to step back, look at how complete and representative a given set of data is, and ask ourselves rigorous questions about what questions it is really qualified to answe .

Share this f ⏵ in

TNS

Mind and the machine

The continuing evolution of analytics
At TNS, we've already evolved from the era of ad-hoc analysis, when researchers collected data with little reference to how it would eventually be used (and then looked through it in the hope it would reveal something useful). Today the design of the instruments for a particular piece of research is informed from the start by the challenge of how best to answer business questions.

The conceptual framework that we use for any type of analysis reflects how the human brain naturally makes sense of information. This framework consists of four different ways of looking at any set of data, whether it was generated through research or arrived, Big Data-style from independent sources. "Dimensions" and "Landscape" address the structure of information; the first seeking out common themes across a data set (the key themes defining a p oduct category, for example), the second looking more closely at

competitive relationships, owned and disputed territory and areas of opportunity. We then build on this structural understanding with more action-oriented means of addressing the data: "Groupings" to segment the subject matter and "Drivers" to reveal the variables that influence relevant results, including causal connections that can be far from immediately apparent.

This approach may be structured, but it retains grounds for flexibilit . It provides a checklist for where and how to look for patterns and themes. In the Big Data era, we will learn to look for different types of patterns in vastly diverse forms of data, but human reason remains the key driving force in identifying them and drawing purposeful connections between them.

Computational muscle can give research the scale and speed that we will increasingly require in the Big Data era, but it is important to distinguish between automating processes

and expecting machines to design them in the first place.   e must not fool ourselves that Artificial Intelligence (AI) is eady to take on the task of formulating questions and crafting the algorithms to answer them. After all, even those that welcome the concept of a technological singularity in which human-designed AI surpasses that of humans themselves, don't envisage it happening until at least 2045. That's a long time to wait to take real advantage of Big Data.

Share this f 𝕏 in

TNS

Mind and the machine
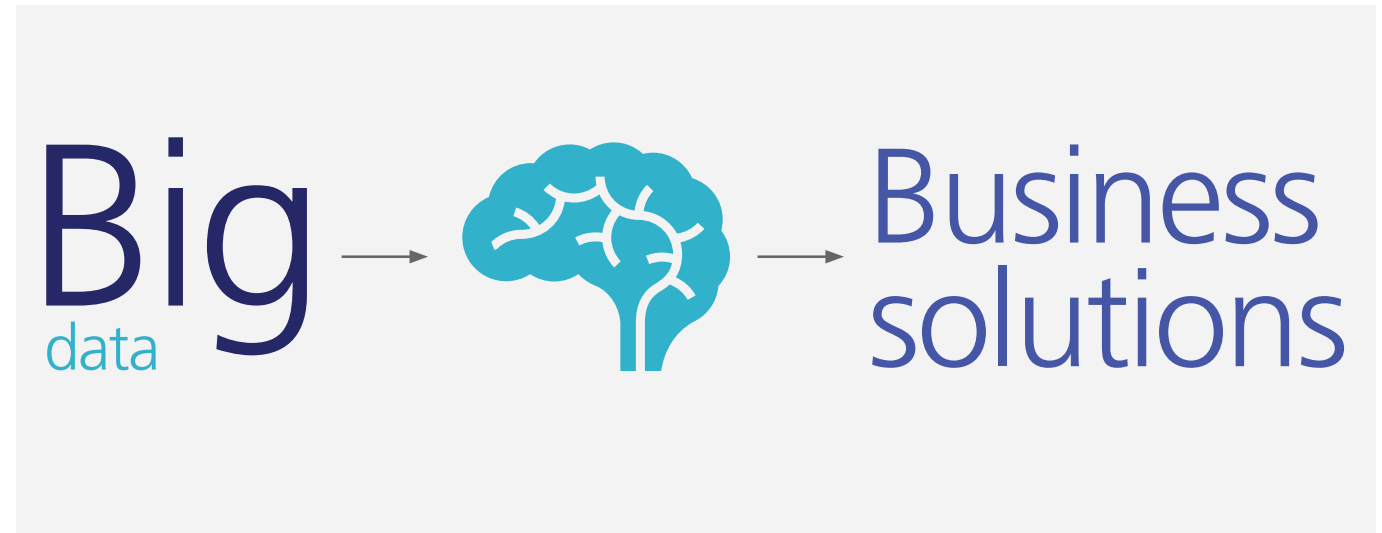
Data and the human imagination

Imposing structure on Big Data will throw up some intriguing challenges – and these challenges will involve logical leaps and lateral thinking for which the human brain remains our best available tool. What is a meaningful means of scoring a positive tweet or Facebook rant? What aspect of somebody's location is actually relevant to the client brief – and what other sources of information can be integrated or overlaid to give context to this information? The location of a car by itself is meaningless. If it's a car unable to fi into the WalMart parking lot on Black Friday, it becomes a whole lot more interesting.

Big
data

→

Business
solutions

When we talk about deploying computational power in the Big Data era, we must therefore be pretty clear about what we are asking computers to do. We must continue to exercise our judgement as to which information is valid and valuable, and how its many varied forms can be coded in meaningful ways. As data curators, that's our job. But by unleashing the power of today's machines we can dramatically increase the scope of data that we can use, the range of

questions that we can ask, and the speed with which we can answer them. Big Data can unleash the potential of human insight and human reason in ways never envisaged before.

The greater computational power that will enable us to make the most of Big Data must be harnessed to an expanded role for the human mind. Depending too much on non-human processing power creates two potential

dangers: that we define in advance what it mus look for and how it must look for it, leading to standardisation and blinkered, undifferentiated thinking, and that we confuse correlation with causation, failing to exercise human judgement about which results are meaningful and which are not. The challenges of the Big Data era will be challenges for the human imagination and human judgement as much as for IT infrastructure. We need to welcome them as such.

Share this f 🐦 in

TNS

About Opinion Leaders

Opinion Leaders are a regular series of articles from TNS consultants, based on their expertise gathered through working on client assignments in over 80 markets globally, with additional insights gained through TNS proprietary studies such as Digital Life, Mobile Life and the Commitment Economy.

About TNS NIPO

TNS NIPO advises clients on specific growth strategies around new market entry, innovation, brand switching and stakeholder management, based on long-established expertise and market-leading solutions. TNS NIPO is part of TNS. With a presence in over 80 countries, TNS has more conversations with the world's consumers than anyone else and understands individual human behaviours and attitudes across every cultural, economic and political region of the world.

Please visit www.tns-nipo.com for more information. Or follow us on Twitter @TNSNIPO

Get in touch

If you would like to talk to us about anything you have read in this report, please get in touch with Paul Oosterveld.

Paul Oosterveld

Consultant Models & Methods
t +31 (20) 5225 480
paul.oosterveld@tns-nipo.com

Paul Oosterveld is Consultant Models & Methods at TNS NIPO. He ensures that TNS NIPO defines and delivers best practice approaches to addressing our clients' business issues

References
1. Nate Silver, The Signal and the Noise
2. Daniel-Kahneman, Thinking Fast and Slow

Share this f 🐦 in