AS THE BIG DATA AND ANALYTICS WAVE BREAKS, THE PROBLEM IS STAYING ANCHORED

Abstract

In this white paper, we'll look at the transformational science of big data and analytics. Covering technological developments that are dragging big data from its subterranean roots to bask in the business-trend-setting sunlight, we'll show how sense can be made of once-static blocks of corporate information. Properly structured, analytical tools can be brought to bear to yield insights not only on customer behaviour, but also business process transformation and much more besides.

The paper also covers the Gartner-christened concept of dark data and how scientists are looking at concepts like orthogonal data to achieve new ways of looking at and understanding the world. As the dust settles, we complete the picture by describing how new ways of storing and retrieving data at the edge as well as from the centre can help break down barriers like organisational siloes.

Unprecedented big data tsunami

Around since 2005, the phrase 'big data' originally referred to any large data set stored in a dumb data warehouse. Almost impossible to manage and process using the business intelligence tools then available, big data was a slightly disparaging term. It carried the sense of 'whatever would do we do with it?' Fast-forward 15 years and the question's done a quick 180 to 'whatever would we do without it?'

"We stand in the path of an unprecedented big data tsunami stirred up by IoT, 5G, social media and sharing apps, autonomics, cloud interconnectivity, and digital cities along with their digital twins, to name but a few."

The change occurred, of course, through the complementary effect of two colliding technologies: the development of cloud-based storage architectures; and the advent of advanced analytical tools designed to make sense of hidden wealths of information. That road is littered with throwaway stats like "90% of today's big data has been created in the course of the past two years"and "the digital universe will be 40 times bigger in 2020" and "by 2025, IoT will make up more than 95% of all real-time data."

They're all coming true. We stand in the path of an unprecedented big data tsunami stirred up by IoT, 5G, social media and sharing apps, autonomics, cloud interconnectivity, and digital cities along with their digital twins, to name but a few. The exploration of orthogonal data sets, part-obscured at obtuse angles to conventional data, will lead to undreamed of analytical possibilities.

Changing the basis of competition

A December 2016 report by the McKinsey Global Insight team entitled **The Age of Analytics: Competing in a Data-Driven World** built on the consultancy's seminal work on the same topic in 2011. The latest report posits that most companies are still capturing only a fraction of the potential value of data and analytics. That was three years ago, but there's a whole bunch of evidence around in 2020 that that view still holds true in spades.

McKinsey found companies with digitally native competitors, like retail, tended to be further ahead than those without others biting at their heels, such as public sector and healthcare. Not surprising when you think about it. In fact, big data and analytics are changing the basis of competition, with leading organisations using such capabilities not only to improve core operations but also to create new business models. A surprise is that traditional retail banks are among the laggards, even though they've got a host of digitally native competitors. McKinsey estimates $260 billion as the potential value they could unlock globally.

> "Retail banks are among the laggards, even though they've got a host of digitally native competitors. McKinsey estimates $260 billion as the potential value they could unlock globally."

The barriers most face in extracting value from data and analytics are largely organisationally based. Many still struggle to incorporate data-driven insights into day-to-day business processes, obstructed in part by siloed data. Meanwhile, they're challenged to attract and retain talented people like data scientists and business translators who can combine data savviness with industry and functional expertise.

For those determined to stay at the front of the pack, machine learning (ML) and artificial intelligence (AI) are data and analytics-related sciences with value potential everywhere. Research has shown 45% of work activities could potentially be automated by current techniques, with ML and AI the enabling technologies 80% of the time. Natural language processing breakthroughs could expand that proportion even further.

Lighting up dark data

Until recently, big data thinking had been focused on three Vs. Coined in 2001 by Gartner analyst, Doug Laney, they were volume, velocity, and variety. But as companies digitise more processes, we've seen the definition expand. In fact, dependent on who one listens to, big data can now run up to seven dimensions: variability, veracity, visualisation, and value have since been added.

The problem with including visualisation and value in the Vs is that most data is sitting there unseen and unused, which Gartner refers to as dark data. In a **Forbes article from 2013**, Svetlana Sicular of Gartner, explained the concept: "Like dark matter in physics, dark data cannot be seen directly, yet it is the bulk of the organisational universe." She could have added much of it was only dark because no one was shining a torch on it. Or maybe a laser, as a more apt metaphor.

> "Like dark matter in physics, dark data cannot be seen directly, yet it is the bulk of the organisational universe."

Which is where orthogonal data comes in. For example, as Svetlana indicated, who would have thought of examining elevator logs to optimise building occupancy, or viewing seemingly innocent Tweets to pinpoint shoplifting activity, or analysing team members' email messages to unlock the secrets of successful project management? It is such intellectual leaps that contribute to V for value.

In addressing velocity, the sheer volume and variety of big data has fuelled faster databases like Exasol and MemSQL, while the need for standardisation has seen widespread adoption of JavaScript Object Notation (JSON). Meanwhile query accelerators such as Apache Impala and Drill speed analysis across multi-format, structured and unstructured data. This is all good news for those looking to optimise the up-to-seven Vs for best-possible big data ROI.

Inventing enabling architectures

Except for the most sensitive environments – usually central government, the military and core financial services systems – the days of on-prem big data are over. For the rest, achieving the scalability implied in the preceding paragraphs in traditional data centres would be unutterably expensive.

While public cloud domains might be OK for high concentrations of low criticality data, colocation with a company like Telehouse could arguably be the ideal big data solution. The transition process to colocation might also be used to introduce and drive storage standards worldwide, offering a relatively painless way out of issues that lead to unsearchable siloed data. It also offers seamless and secure connectivity choices to access third-party information repositories.

"The transition process to colocation might also be used to introduce and drive storage standards worldwide, offering a relatively painless way out of issues that lead to unsearchable siloed data."

It's as well to remember, also, that a proportion of big data (some would argue most) will increasingly be found at the edge of the infrastructure rather than the centre, especially true when IoT takes off. See our white paper on edge computing and connectivity to follow.

Tools under development will afford massive data integration capabilities, breaking through organisational and technological barriers, reaching out to the edge rather than only concentrating on the core. They will enable new business models and deeper-than-ever-before analytical insights across previously inaccessible data living in stovepipes.

The possibilities are limited only by data scientists' imaginations. Telehouse stands ready to help invent the enabling architectures.

Telehouse: guiding big data and analytics journeys

Providing access to a highly connected digital ecosystem of service providers, Telehouse guarantees a highly secure IT operating environment. It also offers the ability to access seamless and secure low-latency and direct connections to multiple cloud services such as Amazon Web Services Direct Connect and Microsoft Azure ExpressRoute when traffic must not be exposed to the internet. There are significant latency advantages too.

> "We have over 40 data centres worldwide in major locations, as well as a highly connected ecosystem of more than 800 service providers."

Well positioned to help organisations access leading cloud services, with over 30 years of experience in providing end-to-end global solutions for which London was and still is the crucible, Telehouse offers guaranteed performance globally. That's based on the following service USPs:

- We have over 40 data centres worldwide in major locations like London, New York, Paris, Frankfurt, Singapore, Turkey, Russia, and China as well as a highly connected ecosystem of more than 800 service providers.

- Our unparalleled reach means we're not only the perfect launch pad for low-latency-dependent applications like IoT but can also help enterprises serve adjacent territories with less-advanced infrastructures.

- The proximity of our Docklands data centre campus to the City of London and Canary Wharf makes us the optimum low-latency option for financial services, where AI and big data are rapidly making themselves felt.

- Over 530 carriers, ISPs and ASPs run services through the Telehouse London hub, and we host the London Internet Exchange (LINX) carrying a huge and growing proportion of UK internet traffic. That gives reach in abundance, including private network interconnects.

- World-leading cyber protection guards Telehouse customers against DDoS, network, application layer and zero-day attacks, and seamlessly integrates with existing on-premise devices and cloud-based platforms for faster attack mitigation.

The most significant USP, however, is that partnering with Telehouse puts at your disposal the imagination and know-how of Telehouse experts to assist you on your own unique and continuing big data and analytics journey.

Contact a Telehouse expert

At Telehouse we can help. We own the space where internet service providers congregate to interconnect and where enterprises come to introduce elegance into their storage and compute strategies. And it's where we can bring the finest minds to bear on helping you solve the developing big data and analytics issues covered in this paper.

Contact Us

Telehouse International Corporation of Europe
Coriander Avenue
London
E14 2AA

T: +44 (0) 20 7512 0550
E: sales@uk.telehouse.net
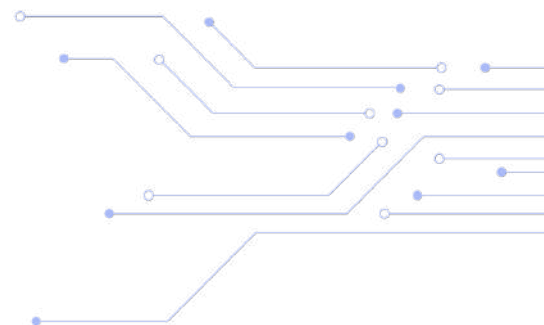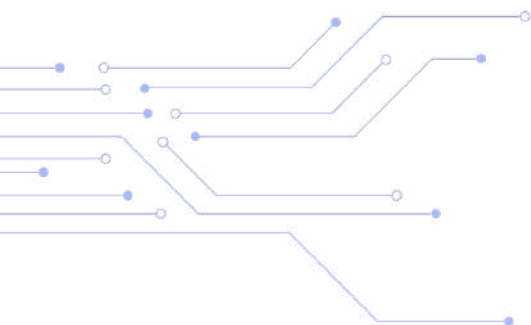
TELEHOUSE

3O
Empowering the connected
world for 30 years