

The big picture of analytics and big data

In this paper we look behind the buzz surrounding analytics and big data to shed light on what it really involves, why it's different and the key challenges to making the most of it.

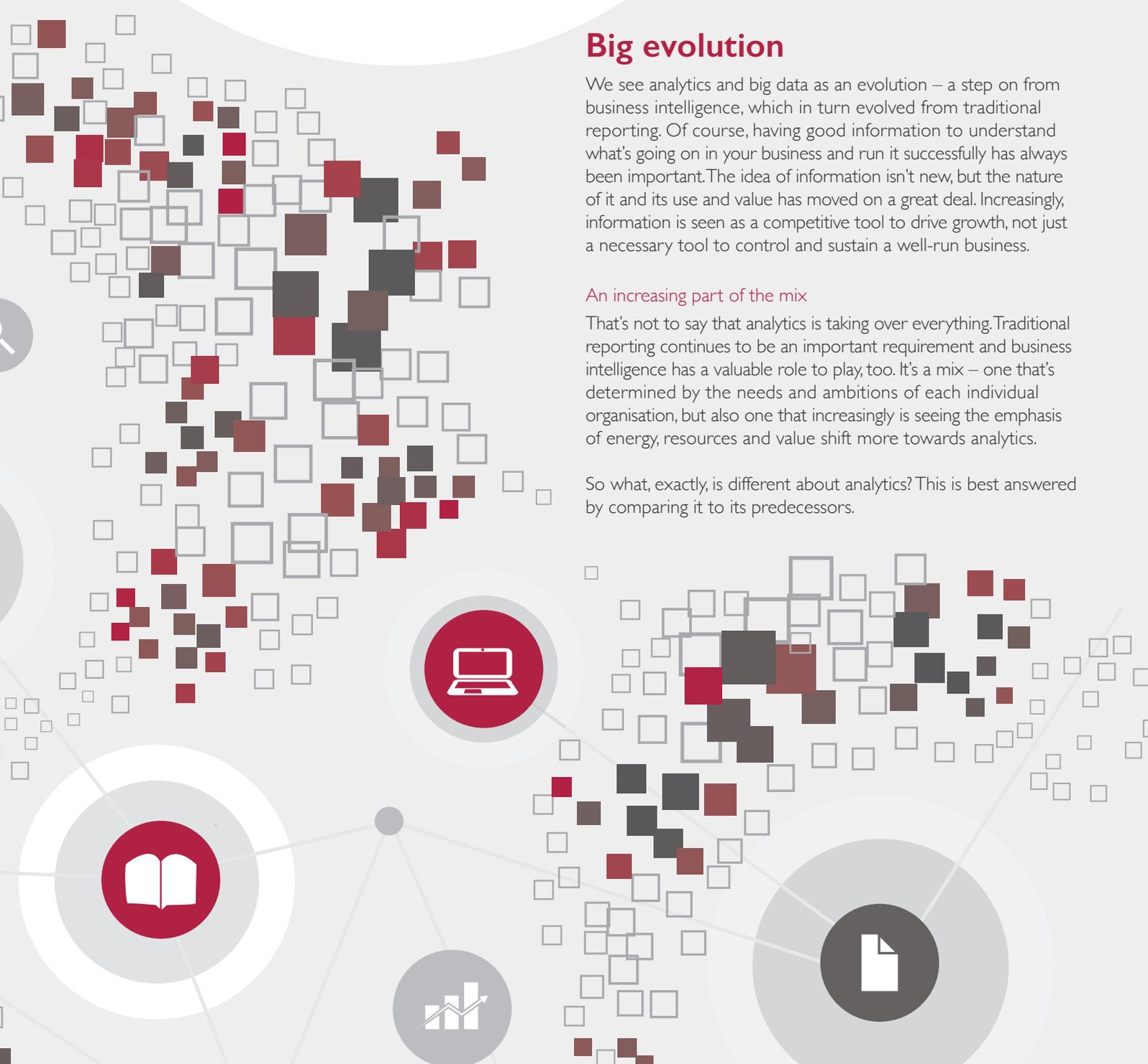
Big evolution

We see analytics and big data as an evolution – a step on from business intelligence, which in turn evolved from traditional reporting. Of course, having good information to understand what's going on in your business and run it successfully has always been important. The idea of information isn't new, but the nature of it and its use and value has moved on a great deal. Increasingly, information is seen as a competitive tool to drive growth, not just a necessary tool to control and sustain a well-run business.

An increasing part of the mix

That's not to say that analytics is taking over everything. Traditional reporting continues to be an important requirement and business intelligence has a valuable role to play, too. It's a mix – one that's determined by the needs and ambitions of each individual organisation, but also one that increasingly is seeing the emphasis of energy, resources and value shift more towards analytics.

So what, exactly, is different about analytics? This is best answered by comparing it to its predecessors.



Reporting

Traditional reporting focuses on pushing information in fixed formats to defined audiences. Monthly finance reports are a classic example. The emphasis is on capture and control. It's static, historic, a matter of record. It's essentially about the past – what's happened.

Business intelligence

Business intelligence made static data more interactive. Fixed-format reports became interactive and online, with more ability to slice-and-dice, drill down and explore. Information became easier to digest with dashboards and other data visualisation techniques. But the information was still essentially backward-looking, based on a defined area and amount of information – typically sourced from internal systems.

Analytics

Analytics looks not only backwards but also forwards – considering the future. It often involves predictive models to show future scenario options, forecast the outcomes, and so help you understand how best to drive your business forward. It is inherently dynamic and its power and purpose lies in enabling you to use information far more effectively than ever before as a living operational and strategic tool. If traditional reporting is a rear view mirror, analytics is a head-up display. As such it's front office and front of foot, with a strong bias to action – analysing what-ifs to identify therefores.

Moreover, the nature of the information is fundamentally different. It involves not only much larger volumes, hence the term big data, but also greater variety – looking not just at your business but at your customers' data, for example. The data sources are more disparate and, often, externally sourced. So the net is cast far wider and more dynamically, with a great variety of unstructured data, such as social media, constantly streaming in.

From a matter of record to a major strategic asset.



From traditional reporting...	to Business Intelligence...	to Analytics and Big Data
“Push”	“Pull”	“Predictive”
Mostly fixed format	More interactive / self service (drill down, slice-and-dice, etc)	Interactive and fully business led
Typically finance-centric offering little to other functions	Applies to all business functions, ‘front office’ (customer facing) and ‘back office’ (finance, HR, etc)	Applies mostly to front office – often customer relationships and product development
Internal & structured data only, with little or no external collaboration	Still mostly internal and structured data, but bringing together more data sources (crossing internal silos of function and geography)	Combines internal and significant external data (e.g. spatial, statistical, 3rd party data); often unstructured (e.g. social media), and often using very large data sets
Implemented as an after-thought to ERP programmes	Implemented independently as a peer of ERP programmes (not subservient)	Implemented as a business capability, with dedicated analytic team(s) built into the organisation
Technology generally not a differentiator	More technology differentiation and choice, generally still ‘on premise’ (not as-a-Service)	Many specialist, highly differentiating technologies and tools; increasing use of open-source and cloud-based approaches.
Backward looking / “rear view mirror” for control (what happened?)	Still backward looking, but looking at causality (what happened and why?)	Forward looking / “head up display”, looking at correlations to predict future outcomes (what will happen?)
Focussed on controlling and sustaining the business (bottom line)	Some elements of bottom and top line focus	Focussed on competitive advantage and growing the business (top line)

Three key differences at the heart of big data

1. The whole data – every last bit of it

With big data you can look at all the data – every last bit of it, rather than just samples. This can lead you to see some very different and valuable insights, as well as avoiding drawing some very wrong conclusions! Whenever you look at a sample of data, you run the risk of some sort of unintended bias in the sample, which colours your analysis.

In the 2008 US presidential election for example, many of the opinion polls called the results wrong. It turned out that the polling was done by ringing up people on landlines, which excluded people who only had a mobile – younger, typically more liberal people. The polling was skewed and therefore failed to predict the result of the election. The fact that you can look at and make use of all the data, rather than just a sample size, is an exciting step change in the world of applied information.

2. What, not why

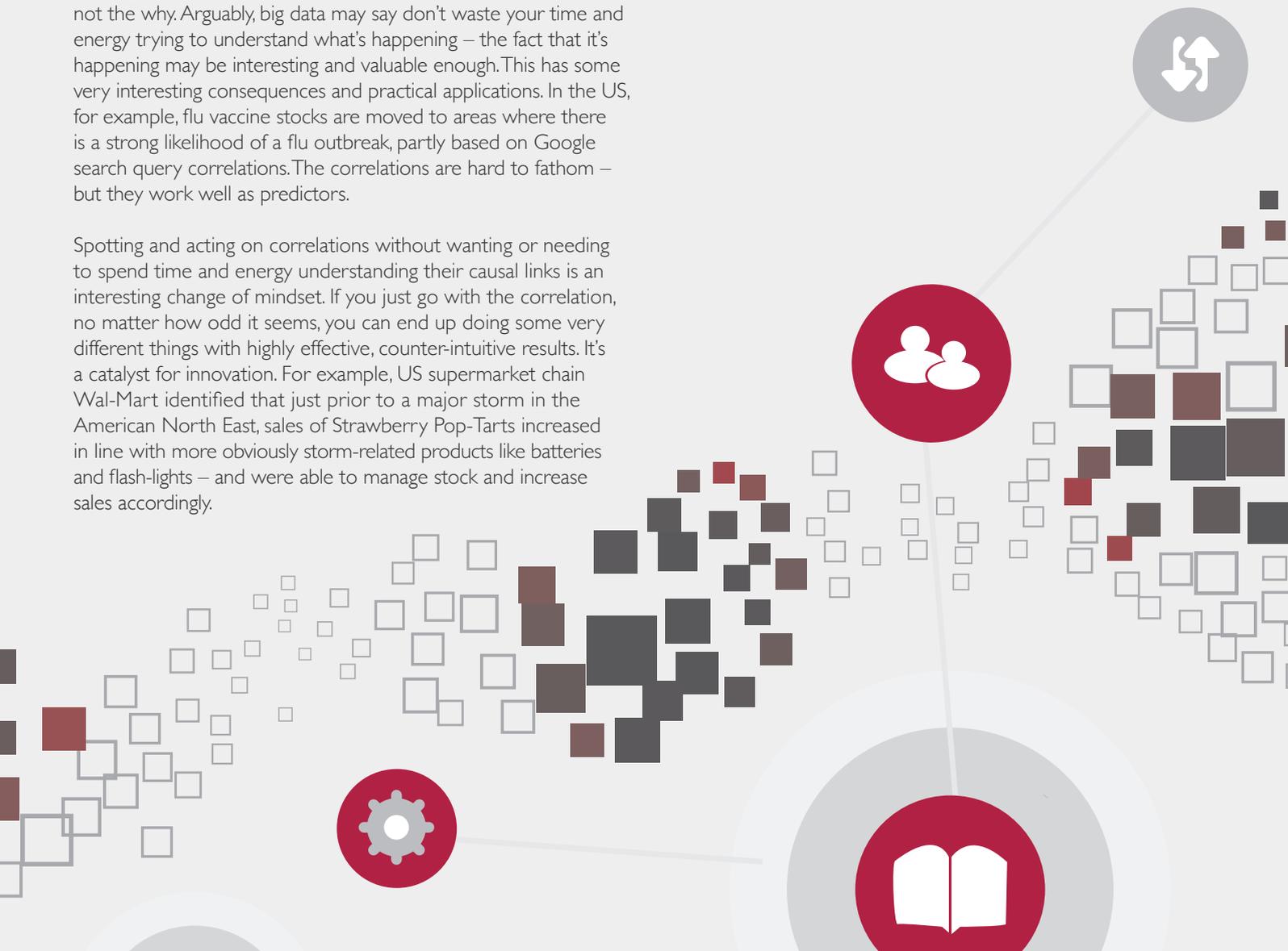
Historically, business intelligence has focused on looking for causal links – finding out why. It's a very left-to-right, linear, process-based way of thinking about information. Big data swerves this by tending to focus on correlations, rather than causes – on the what, not the why. Arguably, big data may say don't waste your time and energy trying to understand what's happening – the fact that it's happening may be interesting and valuable enough. This has some very interesting consequences and practical applications. In the US, for example, flu vaccine stocks are moved to areas where there is a strong likelihood of a flu outbreak, partly based on Google search query correlations. The correlations are hard to fathom – but they work well as predictors.

Spotting and acting on correlations without wanting or needing to spend time and energy understanding their causal links is an interesting change of mindset. If you just go with the correlation, no matter how odd it seems, you can end up doing some very different things with highly effective, counter-intuitive results. It's a catalyst for innovation. For example, US supermarket chain Wal-Mart identified that just prior to a major storm in the American North East, sales of Strawberry Pop-Tarts increased in line with more obviously storm-related products like batteries and flash-lights – and were able to manage stock and increase sales accordingly.

3. Buy versus build in the world of perpetual beta

Big data is increasingly open to everyone – from start-ups on a budget to multinationals with deep pockets. This is helped by another differentiator for big data, which is the growth of on-demand, cloud-based services such as Amazon's RedShift and Google's Big Query.

Indeed these pay-as-you-go services reflect a key change, provoking one of the big questions surrounding big data. With such large data volumes and potentially lots of sophisticated computing power required – should you do it yourself? Or are you better off buying it as a service? Using a service can avoid major capital investments in buying big databases and the necessary number crunching power. What's more, the services are inherently dynamic and elastic, fitting in with and supporting the "perpetual beta" (i.e. constantly in development) culture of big data. But using an external service may present a range of other questions and challenges, such as data provenance, ownership, security and data transfer times. So big data is presenting some big questions on build versus buy – and the best route for you will inevitably depend on your particular situation.



Tackling the challenges

Big data clearly offers many great opportunities across both strategy and operations. But as we know from talking with senior information leaders across a variety of leading organisations, it also comes with a number of big challenges. These are the top issues on the minds of CIOs, CTOs, Heads of Analytics and other information leaders:

Getting the balance right

The balance of build versus buy is discussed above, but another major challenge is how best to strike the right balance across traditional reporting, business intelligence and big data analytics. Many want to redirect effort (in both headcount and funding) away from reporting towards more analytics, but there may not be a simple way to do this. The skills sets and culture are very different, so you are unlikely to be able simply to re-use or re-assign the same people – it may require a deeper reorganisation, which leads us to the second major challenge.

Creating an effective organisation structure for analytics

Big data can be a major headache when it comes to structuring your organisation. What kind of structure do you need? What sort of people should you employ? What skills and capabilities should they have? Where do you put them in the organisation (should it be owned by Marketing, IT or neither)? One key insight here is that you shouldn't expect to be able to do analytics well by simply re-using or extending your reporting team or your business intelligence centre of excellence. The differences inherent in new data sources (as discussed below), typically require different skills, culture and management approaches.

So you may well need to employ fresh people – such as data scientists, analytics operations experts, data visualisation experts, and those with sufficient understanding of the statistical analysis fundamentals. And critically, you may be better off placing them with people at the sharp end of your business – your marketers, your supply chain experts, your product development whizzes, for example. Don't put them in an information team silo; sit them side by side with business folk. That way, together, they'll be able to make the most of big data's great potential to drive insights and actions that transform your strategy and operations.



Old data versus new data

The information used for analytics and big data may be quite different to that used for traditional reporting, as illustrated below. That has a number of big implications beyond organisational and structure skills. For example, effort that is typically invested in cleansing, modelling, structuring and aligning data used for reporting may be better invested in innovative querying and insight generation techniques for analytics. Put simply, rather than spending a lot of effort tidying your loft so it is easy to find things, you might be better off simply paying someone who is very good at finding the things you want in a messy loft. The true value in analytics, however, may often come from combining old and new data sources.



Old data	New data
Data is typically highly structured	Data is often unstructured, and may be of new types (e.g. social media)
Requires good data modelling	Less focus on data modelling
Modest data volumes (compared to big data)	Data is typically large volume (big data)
Strong focus on consistent data quality	Data may be of more variable quality
Data source is usually internal (e.g. ERP)	Data source is typically external
High dependency/link to master data (e.g. product hierarchies)	Less dependency/link to master data (although typically still some)
Structural change to the data is rare and typically slow to deliver	Constantly changing structure and sources (requires a much more agile approach)

Mobilising and managing big data projects

Making a business case for investment in analytics is hard. By helping you find the questions you didn't even know to ask as well as discover new, often counter-intuitive insights, it is by its very nature somewhat speculative. There may be gold in those hills, but it's hard to say just how much before you start to dig. This makes it important to start small, demonstrate the value, and then be prepared to develop in many small, evolutionary steps. Analytics projects are different – inherently open-ended, iterative, fluid, hard to pin down and constantly evolving. The very scale and dynamism that makes big data so powerful can also create big headaches for traditional project management. And no one is ever forced to use an analytics solution purely to execute their business process – they have to want to use it because they see the value. All this means different management approaches are required together with the right change management. And focusing on end user take-up and value realisation is essential.



Plotting your route

There is no easy answer on plotting the best way forward with reporting, business intelligence, analytics and big data. We recommend developing an overall enterprise information roadmap that highlights which information capabilities will really drive value, and a rough timing and sequencing for developing them. We'd also recommend thinking as carefully about the organisation structure and management approach for analytics as you would the technology and data approach. The order of development needs to take into account the underlying data dependencies as well as the business needs. Once the roadmap is set at a high level, the analytics capabilities need to have flexibility to iterate and evolve within that framework – but with just enough discipline and rigour to maintain control.

Whatever course you plot should ultimately serve the aim of harnessing the power of information to reveal valuable insights and actions for your organisation.

