

From Big Data to Big Insights

Blend innovation and agile governance to create an excellent customer experience



Organizations need to focus on best practices, data quality and agile governance to turn the promise of Big Data into reality for CMOs, CIOs, CDOs and their customers.

Introduction: *An inflexion point in the way we use data to serve the business*

Everybody's talking about Digital Transformation, the means by which organizations can use the latest binary tools to create powerful ways for prospects, customers and partners to interact with their brands via apps, websites, networks, portals and other services. It's not overdramatizing digital transformation to say that its successful application will often be the difference between success and failure in the 21st century.

Everybody's also talking about Big Data, the name given to a set of ways in which the new sources of information, from clickstreams, beacons, sensors, metadata, machine data, unstructured data and beyond, can be used to drive insights for businesses at unprecedented speed.

Type 'Big Data' into Google Search and you get over 66 million results; type it into Google News and you get nearly three million. Big Data fascinates because it promises to

deliver what business executives have expected from IT for some time now: game-changing insights derived from raw information that has been gathered from thousands of sources.

With Big Data, CIOs can go beyond the traditional processing, storing and sharing of data and enter a new dimension of leveraging that data to predict and inform business strategy as to which markets and geographies to enter and exit, which to prioritize, how to see where opportunities lie, and how to sniff out risk factors. In addition, Big Data has tactical applications, with examples from the consumer products and retail sector, such as inventory balancing, how and when to promote products, and adjusting products and services offerings.

In this paper, we want to look at Big Data and show its potential for two specific audiences: the CIO - and, today, often the Chief Digital Officer or Chief Data Officer (CDO) who is the architect of tomorrow's data management requirements - and the CMO, responsible for driving the customer experience when interacting with the brand. We will look particularly at two areas here: master data management, and how to instill information governance when integrating multiple data sources.

What's Big Data?

Big Data is a term that has soared in popularity in recent years. Some skeptics might suggest that it is just a reinvention of other terms for analyzing data streams, but Big Data really is something new. First, we have access to many more sources of data from our mobile devices, websites, metadata and sensors equipped to previously dumb objects. Second, data volumes are growing exponentially. And third, we can do something useful with that data because of new powerful technology platforms, such as the Hadoop ecosystem, and because compute capacity is now enormous and cheaper than ever.



Where are we today with Big Data?

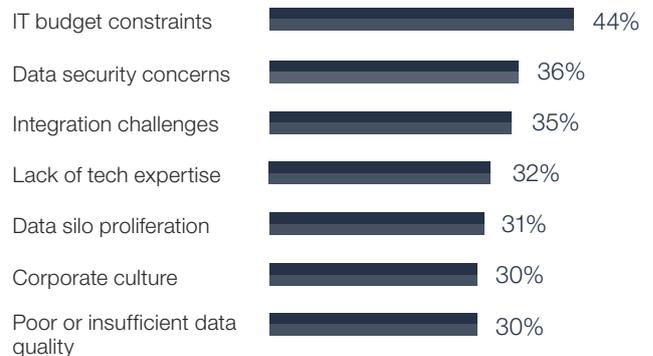
If Big Data were a person, some might argue that it is going through its awkward teenage years. The term itself, in its current meaning, might well date [back to 1998](#) and that term has been in fairly popular [use since 2011](#). But if there is general awareness of what Big Data is, our ability to get value from Big Data projects isn't always up to scratch and organizations haven't yet exploited self-service tools.

For example, [2016 research](#)¹, conducted by IDG Research Services on behalf of Informatica and Capgemini, found that just one-third of CxOs say that one of their concerns is governance and data quality, and that's likely to be one of the reasons they're not always getting enough benefits out of the data lake, the flat architecture repository for storing the wealth of structured and unstructured data. The implicit suggestion with data lakes and the concept of data democratization - with greater access and use provided to the business user - is that this combination is the solution to so many data issues. However, this should not be at the expense of security and privacy.

What's behind this failure to get the full benefits?

The research revealed that IT budget constraints were the number one reason, followed by data security, integration challenges, lacking technical expertise, data silos, corporate culture and data quality; see Figure 1.

Figure 1: What are your organization's top challenges in operationalizing Big Data?



¹ The Big Data Payoff: Turning Big Data into Business Value. A joint report by Informatica and Capgemini, conducted by IDG Research Services, June 2016

But if the full value from Big Data has not yet been unleashed, then there are clear signs of promise with improved decision-making capability, improved collaboration, productivity, agility and reliability and security all seen as gaining, as illustrated in Figure 2.

Figure 2: Which benefits has your organization gained from Big Data initiatives?



We need to talk about Agile Governance

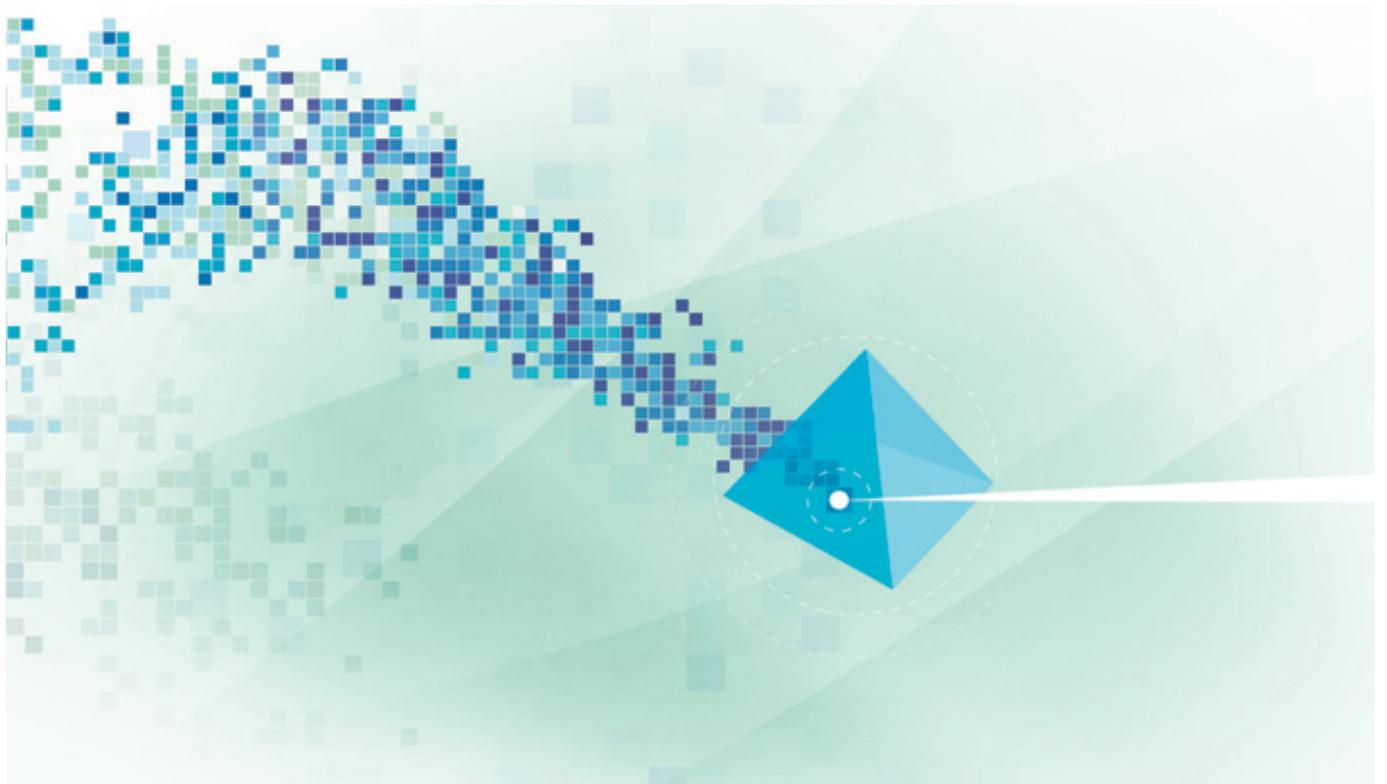
The excitement over Big Data can be so great that some might be in danger of getting ahead of themselves. In addition, some experts hail 'data democratization', in which end-users in the business gain access to powerful

decision-making, with the data and the means to exploit it. However, data scientists today can still spend up to four-fifths of their time in data "wrangling" – in other words, making data manageable. Also, data is powerful and can be dangerous if not carefully managed, so it's critical that information/data governance goes hand in hand with data analytics.

"The conundrum is that data lake adoption is growing, but satisfaction might decline because too often we tend to drop the 'guardrail', in favor of democratization of the data, self-service and a false sense of the degree to which businesses can act in an agile way," says Malay Baral of Capgemini. "That guardrail is governance, data quality and standards. The question we're often missing when we talk about Big Data today is: 'What about data governance?'"

Data governance goes beyond mere documentation and technical integration, and incorporates people and processes, all working towards protecting the integrity of strategic information across its lifecycle.

Baral says a relatively lightweight, schema-less approach where data is identified and simply tagged, almost like a hashtag identifies trending topics on Twitter or a library classifies books, will make information searchable from the outset and also prevent a long 'wait state' between business asking questions and data scientists providing answers.





Data lakes aren't magic, of course, and just putting all your data into, for example, Hadoop, isn't going to solve your information-access problems. It's simply going to give you a 'digital landfill' problem, where you have one place where you can't find anything useful. If you don't know where the data is located, you cannot affect the business outcome.

Essentially, a data lake should be able to provide the right data, for the right purpose, to the right people, and at the right time. Cataloguing (through search, machine learning or crowd-sourced annotation) is the first step to achieving that, but it cannot be the end of the governance perspective. Governance subsequently should provide self-service data identification, management, automation and more.

"You can't run data governance in a 'command-center' approach anymore. So we propose that governance for next-generation data lakes is catalog- or library-based," Baral adds. "Everyone is creating oceans of data and at the same time, too often it's garbage-in, garbage-out."

Otherwise, he says, data scientists will spend too much time gathering up data and not enough working on generating insights.

"You don't need to enforce tight data governance from the beginning. You can have light, agile data governance when you're acquiring data from many sources and putting different layers into the data lake, so that the right people can get access to the information they need, and only selected people have access to other more sensitive information, for example.

"Democratization doesn't mean everybody has access to everything, but the wealth of data can be overwhelming, particularly if it lacks sufficient context. What is important is to achieve a balance between the context provided by Master Data Management while keeping the compelling promise of it all."

Another 'burden' in data lakes may be having too many technologies and it is all too easy to download the latest cloud or open source tool. But what should be a greater focus area is getting the job done, not getting the technology working. Baral recommends standardization, based on the ODPI common reference specification wherever possible, to marshal environments. By reducing the risk of the moving parts it becomes possible to shift the effort towards the outcome and away from the technology.

The role of Master Data Management

Master Data Management (MDM) isn't a new field – it's been around since the 1990s. But it is certainly coming into its own in the Big Data age, as tools mature and businesses fully understand the need for an overarching way to manage information that is business led and that takes into account the resulting organizational implications (such as the creation of data steward roles).

MDM was once relatively simple: these were tools for 'matching and merging' information to verify it and clean it before it moved downstream into various usage scenarios, creating a single version of business-critical information around parties, places or things. Today, however, MDM is the calm at the center of a perfect storm of stimuli that includes:

- The drive to improve the customer experience, often via Digital Transformation processes and programs
- New and changing rules and regulations, such as the European Network and Information Security (NIS) Directive and General Data Protection Regulation (GDPR). (See the box: What's GDPR, and should I be worried?)
- Increased use of cloud computing
- The rise of marketing technology software
- More mergers, acquisitions and other business combinations
- Use of social networks and other tools to expand audience reach.

What's GDPR and should I be worried?

The General Data Protection Regulation will bring in sweeping changes to the ways organizations manage the data of their employees and customers in European Union countries, when it takes effect in May 2018. GDPR has attracted attention for several reasons. It enshrines the 'right to be forgotten', providing users with far more access to the way data about them is processed and shared. Explicit consent must be granted for data to be collected and there must be clarity regarding the purposes that the data will be used for. GDPR has very strong penalties, with offending organizations open to being fined by up to four per cent of annual revenue. In effect, GDPR will force a vast number of organizations to change processes and workflows, and to build in 'privacy by design'.

But MDM has to change in order to fit into this new environment. We still need consistency of structured data such as name, address, telephone number and email address. Now, MDM needs to pull together the structured and the Big Data in a scalable way, to capture much of what organizations know about groups such as customers: their habits, interests, what they spend money on, how they pay, and ways they are willing to engage.

Today's MDM must expand to be equally at home within structured and unstructured data environments. It must effectively manage core information but it also must cater for contextual information (the type of music you like, whether you use discount coupons and codes a lot, for example) with augmented variables based on broader data that analyses your loyalty, lifetime value as a customer and so on.

"Reimagined MDM needs to absorb all this ancillary information, be available on a choice of cloud and on-premise platforms, and data must be presented in a user-friendly way so that marketers, sales executives, customer service reps and others don't need to be data scientists," says Suresh Menon, SVP & GM MDM, at Informatica.

The future of MDM will be a comprehensive picture of the individual, tied in to others such as friends, family, colleagues, influencers with context (your location, activity and who you are with), to build up a 360-degree view, he adds, pointing to the examples of TransAmerica Insurance for improving marketing campaign effectiveness and, looking at operational data, the example of GE Aviation's industrial solution for predictive maintenance.

Integrating data will be necessary to glue together all these sources and categories of data in various file formats, sitting on different applications, in various versions, and on multiple hardware platforms and operating systems.

"Reimagined MDM needs to absorb all this ancillary information, and data must be presented in a user-friendly way so that, marketers, sales executives, customer service reps and others don't need to be data scientists to predict the next logical action to take."

– Suresh Menon, Informatica

MDM in Action

Hotels group [Hyatt](#) grew its revenue by almost one-fifth by strategically managing its guest information preferences in MDM records.

Software company [Citrix](#) developed a ‘fortune teller’ ability in its marketing, thereby dramatically increasing its lead-to-conversion ratio.

IT infrastructure giant [EMC](#) – now part of Dell – saved tens of millions of dollars by turning its data ‘swamp’ into a more manageable and clean data lake.

[Network Rail](#) was able to map all its UK train line assets and gain valuable insights that helped it modernize its services and anticipate risks.

[PostNL](#) effectively reinvented itself from a postal carrier to become a package distributor.

Menon says that with MDM and integration deployed appropriately, ‘the age of annoyance’ where we as consumers are bombarded with irrelevant, misconstrued and poorly-targeted information and offers, will diminish. (See MDM in Action box.) Information must be relevant to the business to help them plan and predict the next best offer or action for the customer.

The role of Master Data Management

With Big Data matched to Agile Governance, the CIO, Chief Data and Chief Digital Officers can finally deliver the strategic insights that have always been close to the top of lists of what business leaders want from their IT counterparts.

By collecting data, analyzing it carefully (and, clearly, quickly) IT leaders can help business leaders to:

- See market opportunities
- Identify key demographic differences
- Plot geographic location differences
- Understand buyer behavior changes
- Improve customer service
- Track logistics and supply/value chain management processes.



For this data to be optimally useful, CIOs will need to address MDM and integration challenges in order to make all data reliable and a meaningful source of decision-making support. They will also need to recruit, or collaborate with data scientists and others who can interpret information and detect false positives.

CIOs can also help their colleagues in sales and marketing by creating and managing the platforms that enable precise targeting. The new sales and marketing opportunities are enabled by an in-depth knowledge of customers, their preferences, the ways they like to be approached, content formats and information that they find compelling, and the stage of the buyer journey they are at. With these marketing data platforms in place, CIOs can build strong relationships with Chief Marketing Officers, heads of sales and their respective departments².

However, CIOs need to pay equal attention to opportunity and risk. Creating the ‘guardrail’ of governance, data quality and standards outlined above will help their organizations to avoid falling foul of regulators or upsetting customers and partners with intrusive behavior.

² <https://www.capgemini.com/insights-data/real-time-dialogue>

What does the CMO get out of all this?

Marketing, at least in the eyes of some sales executives, was once seen as a discipline that was overly-focused on brand positioning and top-of-funnel engagement rather than converting prospects, driving sales and winning more share of wallet.

In the modern digital enterprise, however, CMOs are able to show greater transparency and act as a leading partner with sales on the customer journey. CIOs are increasingly working hand in glove with CMOs to build systems that track every stage of that journey, and deliver measurable results.

In today's ever-more digital economy, there is a need to enhance the customer experience so that the prospect or customer can interact with the brand across multiple channels: websites, mobile apps, content such as blogs or videos, but also, and even today often at least as important as these, the in-person face-to-face channel. Business-to-consumer organizations need to cover all the bases, delighting customers with a dynamic and consistent appeal that stretches across channels such as wearable devices, phone and website to contact center and retail branch.

With Big Data projects and initiatives that are based on a solid rock of reliable, high-quality data, CMOs can connect thousands of data points that, taken together, provide an atomized, micro-segmented understanding of each individual customer. This in turn can be mapped against patterns of behavior sourced from individuals with similar demographic characteristics.



Vendors can identify specific groups, for example young, savvy, 'techie' customers, and market to them accordingly (and even in real-time) with, say, loyalty schemes and special offers of products via channels they know will be attractive because that same approach has been validated many times before.

By exceeding customers' expectations with this insight-driven approach, businesses don't annoy or waste the time of their customers - whatever their segment or grouping - but provide them with propositions that are relevant to their lifestyle, when and where they want them.

The Golden Moment: GDPR as a catalyst for Digital Transformation

We said earlier in this paper that Big Data has not yet realized its full potential. If it is to do so then that might require a stimulus for change. Arguably, for companies that hold data about citizens from Europe, that stimulus could be GDPR.

Andrew Joss, Head of Industry Consulting for Informatica EMEA, certainly thinks so. Joss believes that GDPR initiatives (and Digital Transformation projects) could propel organizations to get their houses in order, auditing their information stores, so they finally have clean master records.

One might view this as just a defensive measure in the name of compliance, but Joss argues that these measures could also help organizations build platforms for wider data management change that could lead to positive market and business opportunities.

"Because organizations hold all sorts of data on customers and employees, knowing what data assets you have, and putting those in the hands of the right people, is powerful, especially for marketing," he says.

"If the data pool is large enough, plus marketers have acquired the rights to use that data, that creates a powerful opportunity for CMOs."

– Andrew Joss, Informatica

“If the data pool is large enough, and if marketers have acquired the rights to use that data, that creates a powerful opportunity for CMOs. It makes them consider where the gaps are in their marketing strategy, so that their campaigns are more targeted, particularly for millennials who know that they’re being marketed to and are aware of the value of their data. So organizations could rethink what customers and employees are happy to share and why; it could change what data gets collected and how.”

The new controls over handling personal data and managing tricky compliance measures, such as information that needs to be isolated or retained for different periods for different regulations, will mandate new governance processes. But these measures will also create a far more reliable data platform that can be repurposed and analyzed for insight, opportunity and profit. “It’s not necessarily rocket science; just being really good at managing data,” Joss says.

Joss also believes that this is one area where the perception of start-ups benefiting from being able to start with a clean sheet of (virtual) paper is not necessarily significant. “Start-ups often have simpler products that are easy to automate, whereas incumbents have experience, data depth and compliance ‘nous’. The Chief Data Officer, as part of digital transformation, should play to these strengths.”

The sheer scale of potential penalties and reputational risks will be a wake-up call, he contends. “They’ve seen how big the pound/dollar/euro signs could be if things go wrong, and the fines are so big that it could be ‘game over’ for some. With the confluence of digital transformation and GDPR, companies have to ask ‘what else can we do with the money we’re investing?’ Managing data better is something they’ve always wanted to do - but never had the time or budget to do.”

“What is important is to view GDPR as a catalyst to a better understanding of the value of your data.”

– Srikant Kanthadai, Capgemini

Srikant Kanthadai, Head of MDM at Capgemini’s Insights and Data practice, adds: “Every organization will have its own way of looking at the challenges of GDPR, but we would recommend first understanding where all the in-scope data sits especially customer and employee data. So the initial step is likely to be an audit to understand the size of the problem, assessing the data with the highest risk first, then moving through to the lower risk data.

“What is important is to view GDPR as a catalyst to a better understanding of the value of your data: Take the GDPR ‘must do’ compliance exercise and use the disciplines to create a value-adding platform for creating data-driven insights that can impact directly business outcomes.”

Conclusion: It’s all about the balance

We live in an exceptional and momentous time where technology and sales and marketing practices are changing rapidly. The brutal force of that change may lead to many casualties where businesses fail to adapt to the new realities.

You could summarize these changes as follows:

- Business is becoming increasingly digital-savvy in terms of buyer awareness of products and services, the ability to serve up and compare offers, how transactions are conducted, and how value chains inter-operate
- Big Data is providing new opportunities to address markets more forensically and at an individual level
- Consumers are increasingly aware of the use of personal data, and are embracing numerous data privacy laws that require companies to gain explicit consent before using personal data
- IT provides the golden, insight-driven thread that enables sales, marketing, logistics and other aspects of the digital organization to work more effectively together and more importantly, provide a more compelling experience to customers
- And at the same time, IT holds the keys to navigating the tricky path of compliance regulations, helping organizations to stay onside with rules and avoid falling foul of increasing penalties for mishandling of customer, employee and other data.

Undoubtedly the challenges are significant, but those organizations that can balance risk with opportunity, and know how to innovate, while understanding the need for strong and agile information governance, will be best placed to thrive on their journeys to becoming truly insight-driven.





Want to find out more? Read on...

Contact us today to learn more about how Informatica and Capgemini Big Data and Data Management solutions can help your enterprise generate true business value from your Big Data initiatives.

Axel Toonen
Informatica

Global Alliance Director
axel@informatica.com

Sri Kanthadai
Capgemini Insights & Data

Global Head of Data Management
Srikant.kanthadai@capgemini.com

Steve Jones
Capgemini Insights & Data

Global Head of Big Data
Steve.g.jones@capgemini.com

About Informatica

Informatica, the Enterprise Cloud Data Management leader, accelerates data-driven digital transformation by enabling companies to become more agile, realize new growth opportunities and create new innovations. With 100 percent focus on everything data, the Informatica Intelligent Data Platform is the industry's most complete and modular solution, built on a microservices architecture. With over 7,000 customers worldwide, Informatica allows organizations to unleash the power of data.

For more information, call +1 650-385-5000 (1-800-653-3871 in the U.S.), or visit www.informatica.com.

Connect with Informatica at [in https://linkedin.com/company/informatica](https://linkedin.com/company/informatica), [🐦 https://twitter.com/Informatica](https://twitter.com/Informatica) and [f https://facebook.com/InformaticaLLC](https://facebook.com/InformaticaLLC).

Capgemini Insights & Data

In a world of connected people and connected things, organizations need a better view of what's happening on the outside and a faster view of what's happening on the inside. Data must be the foundation of every decision, but more data simply creates more questions. With over 11,000 professionals across 40 countries, Capgemini's Insights & Data global practice can help you find the answers, by combining technology excellence, data science and business expertise. Together we leverage the new data landscape to create deep insights where it matters most - at the point of action.

To find out more visit us online at www.capgemini.com/insights-data and [in www.linkedin.com/company/bi-big-data-&-analytics](https://linkedin.com/company/bi-big-data-&-analytics) or follow us on [🐦 Twitter @Capgemini](https://twitter.com/Capgemini)

Contact us at [✉ insights.global@capgemini.com](mailto:insights.global@capgemini.com)



About Capgemini

With more than 190,000 people, Capgemini is present in over 40 countries and celebrates its 50th Anniversary year in 2017. A global leader in consulting, technology and outsourcing services, the Group reported 2016 global revenues of EUR 12.5 billion. Together with its clients, Capgemini creates and delivers business, technology and digital solutions that fit their needs, enabling them to achieve innovation and competitiveness. A deeply multicultural organization, Capgemini has developed its own way of working, the Collaborative Business Experience™, and draws on Rightshore®, its worldwide delivery model.

Learn more about us at

www.capgemini.com

The information contained in this document is proprietary. ©2017 Capgemini. All rights reserved. Rightshore® is a trademark belonging to Capgemini.

©2017 Informatica LLC. All rights reserved. Informatica is a registered trademark of Informatica in the United States and in jurisdictions throughout the world.