# DATA STRATEGY

## HOW TO PROFIT FROM A WORLD OF BIG DATA, ANALYTICS AND THE INTERNET OF THINGS

BERNARD MARR

KoganPage

# Data Strategy

How to profit from a world of big data, analytics and the internet of things

Bernard Marr

KoganPage

**Publisher's note**
Every possible effort has been made to ensure that the information contained in this book is accurate at the time of going to press, and the publishers and authors cannot accept responsibility for any errors or omissions, however caused. No responsibility for loss or damage occasioned to any person acting, or refraining from action, as a result of the material in this publication can be accepted by the editor, the publisher or the author.

First published in Great Britain and the United States in 2017 by Kogan Page Limited



2nd Floor, 45 Gee Street
London
EC1V 3RS
United Kingdom

c/o Martin P Hill Consulting
122 W 27th St, 10th Floor
New York, NY 10001
USA

4737/23 Ansari Road
Daryaganj
New Delhi 110002
India

www.koganpage.com



ISBN 978 0 7494 7985 5
E-ISBN 978 0 7494 7986 2

**British Library Cataloguing-in-Publication Data**

A CIP record for this book is available from the British Library.

**Library of Congress Control Number**

2017001789

Typeset by Integra Software Services, Pondicherry
Print production managed by Jellyfish
Printed and bound by CPI Group (UK) Ltd, Croydon CR0 4YY

# CONTENTS

To accompany the book, the author has prepared an additional online resource, *Beyond the Big Data Buzz*, which can be found at: **www.koganpage.com/beyond-the-big-data-buzz**.

x

THIS PAGE IS INTENTIONALLY LEFT BLANK

# ABOUT THE AUTHOR

Bernard Marr is an internationally bestselling business author, keynote speaker and strategic adviser to companies and governments. He is one of the world's most highly respected voices when it comes to data in business and has been recognized by LinkedIn as one of the world's top five business influencers.

Bernard is a regular contributor to the World Economic Forum, writes regular columns for *Forbes* and LinkedIn Pulse and his expert comments have been featured on TV and radio (eg BBC News, Sky News and BBC World) as well as in high-profile publications including *The Times*, *Financial Times*, *CFO Magazine* and *The Wall Street Journal*.

He has written a number of seminal books and hundreds of high-profile reports and articles, including the international bestsellers *Big Data in Practice: How 45 successful companies used big data analytics to deliver extraordinary results*; *Big Data: Using SMART big data, analytics and metrics to make better decisions and improve performance*; *Key Business Analytics: The 60+ business analysis tools every manager needs to know*; *The Intelligent Company*; and *Big Data for Small Business* in the 'For Dummies' series.

Bernard has worked with and advised many of the world's best-known organizations including Accenture, Astra Zeneca, Bank of England, Barclays, BP, Cisco, DHL, Fujitsu, Gartner, HSBC, IBM, Mars, Ministry of Defence, Microsoft, NATO, Oracle, The Home Office, NHS, Orange, Tetley, T-Mobile, Toyota, Royal Air Force, SAP, Shell, the United Nations and Walmart, among many others.

If you would like to talk to Bernard about any advisory work, speaking engagements or training needs then you can contact him at www.ap-institute.com or via e-mail at bernard.marr@ap-institute.com.

You can also follow @bernardmarr on Twitter, where he regularly shares his ideas, or connect with him on LinkedIn where he writes a regular blog.

THIS PAGE IS INTENTIONALLY LEFT BLANK

# ACKNOWLEDGEMENTS

I feel extremely lucky to work in a field that is so innovative and fast moving and I feel privileged that I am able to work with companies and government organizations across all sectors and industries on new and better ways to use data to deliver real value – this work allows me to learn every day and a book like this wouldn't have been possible without it.

I would like to acknowledge the many people who have helped me get to where I am today; all the great individuals in the companies I have worked with who put their trust in me to help them and in return gave me so much new knowledge and experience. I must also thank everyone who has shared their thinking with me, either in person, in blog posts, in books or any other formats. Thank you for generously sharing all the material I absorb every day! I am also lucky enough to personally know many of the key thinkers and thought leaders in the field and I hope you all know how much I value your inputs and our exchanges.

Finally, I must thank my editorial and publishing team for all your help and support. Taking any book from idea to publication is a challenging process and I really appreciate your input and help – thank you Claire, David, Amy and Lucy.

# 01 Why every business is now a data business

Data is changing our world and the way we live and work at an unprecedented rate. Depending on your viewpoint, we're either at the start of something incredibly exciting or we're entering a terrifying Big Brother era where our every move can be tracked – and even predicted (both sides have a point). Business leaders and managers, however, have little time for data scepticism. Data is already revolutionizing the way companies operate and it will become increasingly critical to organizations in the coming years. Those companies that view data as a strategic asset are the ones that will survive and thrive. With the massive growth in big data and the Internet of Things, plus rapidly evolving methods for analysing data, the importance of data across every aspect of business will only increase.

## The astonishing growth of big data and the Internet of Things

Every two days we create as much data as we did from the beginning of time until 2003. Every two days. And the amount of data we're creating continues to increase rapidly; by 2020, the amount of digital information available will have grown from about 5 zettabytes today to 50 zettabytes. Almost every action we take leaves a digital trail – browsing online, shopping in a bricks-and-mortar store with a credit card, sending an e-mail, taking a photograph, reading an online article, even walking down the street if you're carrying a mobile phone or there are CCTV cameras in the vicinity.

The term 'big data' refers to the collection of all that data and our ability to use it to our advantage across a wide range of areas, including business. Data in itself isn't a new invention. Going back even before computers and databases, we still used data to track actions and simplify processes – think

of paper transaction records and archive files. Computers, and particularly spreadsheets and databases, gave us a way to store and organize data on a large scale, in an easily accessible way. Suddenly, information was available at the click of a mouse.

Until relatively recently, though, data was limited to spreadsheets or databases – and it was all very ordered and neat. Anything that wasn't easily organized into rows and columns was simply too difficult to work with and was ignored. Now, however, advances in storage and analytics mean that we can capture, store and work with many, many different types of data. As a result, data today can cover everything from spreadsheets to photos, videos, sound recordings, written text and sensor data.

There's no doubt that the sheer amount of data we're creating is, well, big. But, if I'm honest, I've never been entirely comfortable with the term 'big data'. It feels too simplistic to me, focusing on the volume of data rather than the incredible opportunities this data creates. I wish there was a better term to describe this huge shift in our technology, culture and world. That's why, in this book, I talk about 'data' in all senses, big and small – because it doesn't matter how much data you have, it's whether you use it successfully that counts.

## A brave new (data-driven) world

Big data knows a lot about you. It goes way beyond Google knowing what you've searched for online and Facebook knowing who you're friends with. Your Internet service provider knows every website you've ever visited. Ever. Even in private browsing. Google knows your age and gender (even if you've never told them) and you can be sure they have a comprehensive profile of you and your interests, so they can decide what ads to show you. Facebook clearly knows who you're friends with and who you're in a relationship with. But did you know Facebook can also predict whether your relationship is going to last or, if you're single, when you're about to be in a relationship (and with whom)? Facebook can also tell how intelligent you are, based on an analysis of your 'likes'. (In case you're wondering, liking Curly Fries, Science, Mozart, Thunderstorms or The Daily Show predicted high intelligence, while likes for Harley Davidson, Lady Antebellum, and I Love Being a Mom predicted low intelligence.)

The police know where you're driving, certainly in the UK, where they have access to thousands of networked CCTV cameras across the country that scan number plates and take pictures of cars and their drivers. In the United States, many cities make similar use of traffic cameras. Your phone

also knows how fast you're driving. For now, that information isn't shared with the police, but more and more insurance companies are starting to make use of smartphone data to deduce who is a safe driver and who's a riskier prospect.

Your grocery store loyalty card tracks the brands you like and collects mountains of information on your purchasing habits and preferences. Retailers use this data to personalize your shopping experience, but it can also be used to predict what else you might want to buy in future. In one well-known case, US retailer Target predicted a teenage girl was pregnant (based on her buying habits) and started sending her baby-related offers – the only problem was her own parents didn't yet know she was pregnant.

Big data powers much more than social media networks and coupon mailings, though. Its influence stretches to almost every aspect of modern life, from healthcare to space exploration, even to our political elections.

In an analytics-driven election campaign, for example, the focus is on targeting swing or undecided voters. After all, why waste time campaigning to those who are definitely going to vote for you, or those who never will in a million years? This technique was pioneered by the Obama campaign in 2012 when a team of over 100 data analysts were tasked with running over 66,000 computer simulations every day.

First, Obama's analysts collected and amalgamated all the data they could from voter registration data, donations, public records and bought-in third-party commercial data (including data mined from social media). Then everybody who had been identified was evaluated on their likelihood of voting for Obama, based on how well their data profile matched that of known supporters. Armed with their sophisticated demographic information, the team then launched targeted campaigns. These were aimed at increasing voter turnout and registration amongst sectors where the likelihood of backing their candidate was high, and influencing voter choice in sectors where the support metric indicated voters could go either way. This meant that targeted messages could be despatched – via e-mail, social media posts and browser display ads – depending on whether an individual needed to be convinced to register, vote, or pick the correct candidate.

In the years since then, all parties and most candidates have enthusiastically launched their own analytics strategies.

Big data is also helping to answer the question of whether there has ever been life on Mars. NASA's Jet Propulsion Laboratory, which runs the day-to-day mission planning for the Mars Rover spacecraft, is now using Elasticsearch technology (also used by companies like Netflix and Goldman Sachs) to process all of the data transmitted from the Rover craft during its four daily

scheduled uploads. While mission-planning decisions used to be based on the previous day's data, the move to real-time analytics vastly speeds up the time in which decisions can be taken by mission control. Patterns and anomalies in the data sets can be spotted far more quickly, and correlations which could provide mission-critical insights are more likely to become apparent, leading to a greater rate of scientific discovery and less danger of malfunction or failure.

Even healthcare has not escaped the big data touch. For years, the basis of most medical research and discovery has been the collection and analysis of data: who gets sick, how they get sick and why. But now, with sensors in every smartphone and doctors able to share information across disciplines, the quantity and quality of the data available are greater than ever before, which means that the potential for breakthroughs and change is growing just as exponentially. Smartphones and other popular smart devices including Jawbone, Fitbit and others, now have the capacity to help people track their progress towards a healthier lifestyle. Apps and devices to help track and monitor chronic ailments like diabetes, Parkinson's and heart disease are also being developed.

The medical industry already collects a huge amount of data, but it's often siloed in individual doctors' offices, hospitals and clinics. Unifying that data – and combining it with patient-collected data from smart devices – is the industry's next big hurdle to overcome. Healthcare providers are already focusing on digitizing patient records and ensuring access to one set of records across the healthcare system. Pattern recognition software is already being used to aid diagnostics. So far, certain algorithms are proving as effective or more effective than human diagnosticians in spotting cancers in test results. There is incredible potential here for catching more diseases at earlier stages, and thus increasing the likelihood of treatment success. Big data is also being used to track, analyse and treat epidemics across the world, including Ebola and Zika.

All this is just the tip of the iceberg, and data volumes will only continue to grow. More often than not, when we sign up to a new product or service, whether it's a fitness tracker or a store loyalty card, we're happily giving access to our personal data – in return for benefits like improving our fitness or collecting points towards a free coffee. As more companies tap into the possibilities of data, and as the technology advances to gather more and more information, the amount of data available is predicted to grow exponentially.

We will also get better at analysing these heaps of data, with new tools coming onto the market every week. In fact, Microsoft and Salesforce both recently announced tools to allow non-coders to create apps to view and analyse business data. And as we get better at analysing data, our ability to

make predictions improves, too. Market intelligence firm International Data Corporation predicts that half of all business analytics software will include prescriptive analytics abilities by 2020 – which means not only will the software be able to predict customer or user actions, it will be able to make specific recommendations based on those predictions. We're on the cusp of a very exciting time in terms of data and analytics, and the technology available in five or ten years' time could offer possibilities that we can't even imagine yet.

Part of the reason for this explosion in data is the Internet of Things (IoT), sometimes known as the Internet of Everything (IoE). The IoT refers to devices that collect and transmit data via the Internet, and covers everything from your smartphone, smartwatch, Fitbit band, even your TV and refrigerator. The IoT has seen enormous growth in recent years, and it's only just getting started. Today, there are about 13 billion devices that connect to the Internet. By 2020, that number is predicted to rise to anything between 50 and 70 billion. Smartphone users alone are predicted to number over 6 billion by 2020.

Smart devices are transforming our world, our cars, our homes and our businesses. By 2020, a quarter of a billion cars will be connected to the Internet, allowing scope for a whole host of in-vehicle services and automated driving. What was once science fiction is already becoming reality – Google's self-driving cars already clock up several thousand miles a week.

'Wearable' technology is a crucial part of the IoT, and the global market for wearable devices (things like smartphones, Fitbits, etc) grew 223 per cent in 2015. One in six consumers currently owns and uses wearable technology in one way or another. All of these devices create a wealth of data, and we're only just starting to realize the implications of this now.

Connected devices can not only connect to the Internet, they can also connect and share information with each other. In fact, machine-to-machine connections will grow to 27 billion by 2024. So, in the near future, it's not unreasonable to imagine your refrigerator knowing when your milk is out of date and automatically telling your smartphone to order more in the next online shop.

## Are we nearing artificial intelligence?

In computing terms, Artificial Intelligence (AI) has been the ultimate goal since the very first computers were invented. It's also been a tantalizing prospect for science fiction writers! But are we finally getting close to realizing AI? Cognitive computing definitely brings us a big step closer.

Combining cognitive science (the study of the human brain) and computer science, cognitive computing looks set to impact almost every area of our lives, from business to healthcare and even our private lives. The aim is to allow a computer to simulate human thought and mimic how our brains work. This allows computers to undertake things that we humans take for granted, like understanding natural language or recognizing objects in a picture.

IBM's Watson system is a prime example of cognitive computing. The system 'learns' as it processes information, so the more data the system is given, the more it learns, and the more accurate it becomes. In practical terms, this technology could be used in any field in which a large amount of complex data needs to be processed and analysed to solve problems, including healthcare, law, education, finance and, of course, business. The technology is already being used in the hospitality industry; Hilton Hotels recently unveiled the first concierge robot, Connie, which can understand natural language and respond to guests' questions about the hotel, local attractions, restaurants, etc.

As computers are more able to think like humans, they enhance our knowledge and capabilities. Just as the heroes of science fiction movies turn to their computers for analysis, predictions and conclusions on what to do next, in real life we're moving into an era where computers can enhance human knowledge in entirely new ways.

Cognitive computing is underpinned by machine learning and deep learning technology, which allows computers to autonomously learn from data. This technology means computers can change and improve their algorithms by themselves, without being explicitly programmed by humans. How does it work? Put simply, if we give the computer a picture of a cat and a picture of a ball, and show it which one is the cat, we can then ask it to decide if subsequent pictures contain cats. The computer compares other images to its training data set (ie the original cat image) and comes up with an answer. Today's machine learning algorithms can do this unsupervised, meaning they do not need their decisions to be pre-programmed. The same principle applies to even more complex tasks, albeit with a much larger training set. Google's voice recognition algorithms, for instance, work from a massive training set, but it's still not nearly big enough to predict every possible word, phrase or question.

But the technology is improving all the time, and machine and deep learning are responsible for advances in computer vision, audio and speech recognition and natural language processing. It is what allows computers to communicate with humans (not always 100 per cent successfully, as Microsoft's slightly crazy, racist Twitter bot proved), and makes Google's self-driving cars possible. It is also the reason Facebook is able to recognize individuals

in photos to the same level as humans can, automatically suggesting tags for individuals.

So, is artificial intelligence just around the corner? Probably not, at least not in the science fiction sense. Many scientists believe computers will never be able to 'think' like a human brain. Whichever way you look at it, computers' abilities to see, understand and interact with the world around them are growing at an incredible rate. And as the amount of data we have continues to increase, so too will computers' abilities to learn, understand and react.

The technology has advanced to such an extent that it is now possible for computers to recognize and respond to human emotions. Known as 'affective computing', this technology analyses facial expressions, posture, gesture, tone of voice, speech and even the rhythm and force of keystrokes to register changes in a user's emotional state.

Imagine the potential of this technology. Your computer could recognize when you are frustrated or struggling with a task and provide additional information to help you along. Your phone could tell you to take a break when your stress levels are high. Or, without being asked, your smart home could provide soothing music and lighting when you get in from a bad day at the office. If this all sounds a bit far-fetched, it's not. Leading organizations like Disney, the BBC and Coca-Cola are already partnering with Affectiva, a company specializing in facial recognition technology, to test the effectiveness of adverts and assess how viewers react to content. The same company is also working with a Japanese car company to create in-car technology that can detect when you're distracted or drowsy, and contact emergency services or your next of kin in the event of an emergency. Microsoft has even tested a bra that can sense stress levels in women.

Just as computers can never learn to 'think' in the same way as a human brain, these emotional machines will never really be emotional, but we're nearing the time when machines will at least appear to give us suitable emotional responses. The really exciting part is that we're only just starting to explore the possibilities of all this technology. In 20 years' time, cognitive and affective computing will be mainstream technology.

## How data is revolutionizing the world of business

I wholeheartedly believe that big data and its implications will affect every single business – from Fortune 500 enterprises to small businesses – and change how we do business, inside and out. It doesn't matter what field

you operate in or the size of your business; as data collection, analysis and interpretation become more readily accessible, they are having an impact on every business.

## *The key uses of data in business*

There are three core areas where data really matters to business: improving decision making, improving operations, and the monetization of data.

First, big data enables companies to collect better market and customer intelligence. With the ever-increasing amount of data available, companies are gaining much better insights into what customers want, what they use (and how), how they purchase goods, and what they think of those goods and services. And this information can be used to make better decisions across all areas of the business, from product and service design to sales and marketing and aftercare.

Second, big data helps companies gain efficiencies and improve their operations. From tracking machine performance to optimizing delivery routes to even recruiting the very best talent, big data can improve internal efficiency and operations for almost any type of business and in many different departments. Companies have even started using sensors to track employee movements, stress, health, and even who they converse with and the tone of voice they use, and using that data to improve employee satisfaction and productivity.

The IoT plays a huge role in improving operational performance. A big part of the IoT isn't so much about smart devices, but about sensors. These tiny innovations can be attached to everything from yogurt cups to the cement in bridges and then record and send data back into the cloud. This will allow businesses to collect more and more specific feedback on how products or equipment are used, when they break, and even what users might want in the future. For example, Rolls-Royce aircraft engines contain sensors that send real-time data on the engine's function back to monitoring stations on the ground. This information can be used to detect malfunctions before they become catastrophic, and possibly to investigate – and hopefully prevent – the causes of aircraft disasters.

Third, data also provides the opportunity for companies to build big data into their product offering – thereby monetizing the data itself. John Deere is an excellent example of a company that is not only using data to benefit its customers, but also as a new product offering. All new John Deere tractors are equipped with sensors that can help the company understand how the equipment is being used, and predict and diagnose breakdowns.

But they've also put the sensors to work for the farmers, offering access to data about when to plant, where, the best patterns for ploughing and reaping, and more. It's become an entirely new revenue stream for what was seen as quite a traditional company.

Let's look at some more examples that demonstrate the wide applicability of big data. For one, supply chain management is a field where big data and analytics have obvious applications. Of course, supply chains have for a long time now been driven by statistics and quantifiable performance indicators. But the sort of analytics that are really revolutionizing industry today – real-time analytics of huge, rapidly growing data sets – were largely absent. Many factors can clearly impact on supply chain management – from the weather to the condition of vehicles and machinery, and so recently leaders in the field have thought long and hard about how this could be harnessed to drive efficiencies.

Applications for data analysis have already been found in inventory management, forecasting, and transportation logistics. In warehouses, digital cameras are routinely used to monitor stock levels and the data provides alerts when restocking is needed. Forecasting takes this a step further – the same camera data can be fed through machine learning algorithms to teach an intelligent stock management system to predict when a resupply will be needed. Opportunities to create efficiency and savings through smart use of data are everywhere and concerted effort is being put into finding them. Eventually, the theory is, warehouses and distribution centres will effectively run themselves with very little need for human interaction.

In retail, both online and offline retailers who are embracing a data-first strategy towards understanding their customers, matching them to products and parting them from their cash are reaping dividends. Today, retailers are constantly finding innovative ways to draw insights from the ever-increasing amount of information available about their customers' behaviour. Big data analytics is now being applied at every stage of the retail process – working out what the popular products will be by predicting trends, forecasting where the demand will be for those products, optimizing pricing for a competitive edge, identifying the customers likely to be interested in them and working out the best way to approach them, taking their money and finally working out what to sell them next.

In banking, The Royal Bank of Scotland (RBS) has developed a big data strategy which it calls 'Personology' in an attempt to reconnect with customers. The bank, which is currently undergoing re-privatization seven years after it was bailed out to the tune of £45 billion by UK taxpayers during the financial crisis, is combining data analytics with a 'back to the

'70s' approach to customer service. The philosophy is one of the developments of the 800-person-strong analytics department, created as part of a £100 million investment in analytic skills and technology across the organization.

The move is about restoring a disconnect which developed between banks and customers after the 1970s. The theory goes that early attempts at data-driven marketing, such as audience segmentation and mass mailing, were too focused on what the banks wanted – usually making sales – and often ignored what customers wanted. The plan is to restore the trust and feeling of support that bank customers would have expected during the 1970s or before – when bank staff would know a customer by name, understand what their needs were on a personal level and attempt to offer services that supported those needs.

As an example of the new strategy in operation, analysts combed financial transaction data to pinpoint situations where customers may have been paying twice for services packaged with bank accounts – such as mobile phone insurance or breakdown assistance. Although at first there were worries that alerting customers to this situation could prompt them to cancel their RBS products in at least some cases, in practice, every single person who was alerted opted to cancel the duplicate third-party service and retain their RBS service.

Other services which fall under the Personology banner include wishing customers a happy birthday if they visit a branch on the day, and automated text messages to let them know that their cash is safe if they accidentally leave it behind after withdrawing it from an ATM.

Big data has even optimized the delivery of your Friday night pizza. Domino's, the world's largest pizza delivery chain, has consistently pushed its brand onto new and developing tech, and it is now possible to order pizzas using Twitter, Facebook, smart watches and TVs, and in-car entertainment systems such as Ford's Synch. So, while it may seem at first glance that pizza and big data are not well matched, the logistics of delivering close to a million pizzas a day across 70 countries throws up exactly the sort of problems that big data is good at overcoming.

Domino's multi-channel approach to interfacing with customers gives them the opportunity to generate and capture a lot of data – which Domino's capitalizes on by using it to improve the efficiency of their marketing. Data captured through all its channels – text message, Twitter, Android, Amazon Echo, to name just a fraction – is fed into the Domino's Information Management Framework. There it's combined with enrichment data from a large number of third-party sources, such as the United States Postal Service,

as well as geocode information, and demographic and competitor data, to allow in-depth customer segmentation. This means that individual customers or households can be presented with totally different presentation layers than others – different coupons and product offers – based on statistical modelling of customers fitting their profile. As well as customer segmentation, data is used to assess performance and drive growth at individual stores and franchise groups.

### *Smart factories and Industry 4.0*

First came steam and the early machines that mechanized some of the work our ancestors did. Next came electricity, the assembly line and the birth of mass production. The third era of industry came about with the advent of computers and the beginnings of automation, when robots and machines began to replace human workers on those assembly lines. And now we enter what is being called the fourth industrial revolution, Industry 4.0, in which computers and automation will come together in an entirely new way, with robotics connected remotely to computer systems equipped with machine learning algorithms that can learn and control the robotics with very little input from human operators.

Industry 4.0 introduces the notion of the 'smart factory', in which cyber-physical systems (a combination of computers, networks and physical actions) monitor the physical processes of the factory and make decentralized decisions. In the smart factory, the machines are augmented with web connectivity and connected to a system that can visualize the entire production chain and make decisions on its own. They essentially become IoT systems, communicating and cooperating both with each other and with humans in real time via the wireless web.

For a factory or system to be considered Industry 4.0, it must include the following four features: 1) interoperability, ie machines, devices, sensors and people that connect and communicate with each other; 2) information transparency, where the systems create a virtual copy of the physical world through sensor data in order to contextualize information; 3) technical assistance, both in terms of the ability of the systems to support humans in making decisions and solving problems, and the ability to assist humans with tasks that are too difficult or unsafe for humans; and 4) decentralized decision making, where the cyber-physical systems make simple decisions on their own and become as autonomous as possible.

As with any major shift in industry, there are challenges in adopting this approach. Data security issues are increased when you integrate new systems

and increase access to those systems. Systems need to be incredibly reliable and stable for successful cyber-physical operations, and this can be difficult to achieve and maintain, especially when you consider there is a systematic lack of experience and manpower to create and implement these systems. Likewise, avoiding technical problems that could cause expensive production outages is always a concern. In addition, there could be issues with maintaining the integrity and quality of the production process with less human oversight. Finally, whenever new automations are introduced, there is always a risk of losing valuable human jobs. All these issues combined with a general reluctance from stakeholders and investors to invest heavily in costly new technologies means Industry 4.0 has many hurdles to overcome before it becomes mainstream.

But the benefits of an Industry 4.0 model could outweigh the concerns for many production facilities. For example, in very dangerous working environments, the health and safety of human workers could be improved dramatically. Supply chains could be more readily controlled when there is data at every level of the manufacturing and delivery process. Computer control could produce much more reliable and consistent productivity and output. And the results for many businesses could be increased revenues, market share and profits.

In his book *The Fourth Industrial Revolution*,[1] Professor Klaus Schwab, Founder and Executive Chairman of the World Economic Forum, describes how this fourth revolution is fundamentally different from the previous three, which were characterized mainly by advances in technology. In this fourth revolution, we are facing a range of new technologies that combine the physical, digital and biological worlds. These new technologies will impact all disciplines, economies and industries, and even challenge our ideas about what it means to be human. These technologies have great potential to continue to connect billions more people to the web, drastically improve the efficiency of businesses and organizations and help regenerate the natural environment through better asset management, potentially even undoing all the damage previous industrial revolutions have caused.

Reports have also suggested that emerging markets like India could benefit tremendously from Industry 4.0 practices, and the city of Cincinnati, Ohio has declared itself an 'Industry 4.0 demonstration city' to encourage investment and innovation in the manufacturing sector there.

The question, then, is not if Industry 4.0 is coming, but how quickly. I suspect that the early adopters will be rewarded for their courage jumping into this new technology, and those who avoid change risk becoming irrelevant.

## *Automation and the (very real) threat to jobs*

As automation increases, computers and machines will replace workers across a vast spectrum of industries, from drivers to accountants and estate agents to insurance agents. By one estimate, as many as 47 per cent of US jobs are at risk from automation.

When you read or hear news stories about the imminent takeover of robots and algorithms that will eliminate jobs for human workers, many times the first examples given are blue-collar jobs like factory workers and taxi drivers. But plenty of professional jobs may also be at risk of being outsourced to computers. More and more, sophisticated algorithms and machine learning are proving that jobs previously thought to be the sole purview of humans can be done as well or better by machines. Boston Consulting Group has predicted that by 2025, as much as a quarter of jobs currently available will be replaced by either smart software or robots.[2] A University of Oxford study suggested that up to 35 per cent of existing jobs in the UK could be at risk of automation within the next 20 years.[3]

Let's look at some examples. In insurance, formulas have been used for decades to decide how much insurance a person is qualified for and at what rate. But much of what brokers and underwriters do today can be done by computers using big data and machine learning, and new tools are automating this decision-making process even further, reducing the need for human input.

Architects, too, could be under threat. Programs already exist to help individuals design their own homes, making architectural and design choices more automated. For now, these tools are mostly used as visualization tools or to replace architects on very small or simple projects. But as these programs become more sophisticated, the need for human architects and building designers will decrease.

In the financial industry, algorithms can now analyse financial data and prepare accounts (as well as do tax returns) – without the need for accountants. Bank staff have already been partially replaced by ATMs, but soon even higher-level bankers, including loan officers, could be easily replaced by automated systems. Even governments are now using big data and machine learning to check tax returns and identify potential fraud in tax matters. We know that computers are already being used to make stock trades faster than humans ever could and they're even used to predict how the market will react and make recommendations whether you should buy or sell.

Human resources, headhunting and hiring are already being affected by data mining as algorithms take on the job of sorting through resumes to find

the perfect candidates. Other traditional human resource tasks, including collecting and filing paperwork, advising employees about benefits, etc, can also be easily automated.

Marketing is all about that most human of skills, persuasion and manipulation. But even that is being successfully outsourced to computers. Persado, a natural language software firm, has put its computers to the task of writing compelling e-mail subject lines for large retail organizations that can as much as double open rates. Companies are also experimenting with automated ad buying – instead of having people choose which magazines to place ads in and on which pages, the computers take care of it, using billions of data points for reference.

Even lawyers, one of the quintessential professions, could be affected. In the discovery phase of a lawsuit, lawyers and paralegals can be required to sift through thousands, even tens of thousands of documents, depending on the case. Now, sophisticated databases can use big data techniques like syntactic analysis and keyword recognition to accomplish the same tasks in much less time. In fact, it's possible that a Watson-style machine learning system could be legally 'trained' to review precedent and case history and even draft legal briefs – traditionally the job of lower-level law firm associates. But don't think it's only the lowly junior associates whose jobs are at risk; lawyers are well paid now to predict the outcome of major cases, but a statistical model created by researchers at Michigan State University and South Texas College of Law was able to predict the outcome of almost 71 per cent of US Supreme Court cases.[4] That ability to predict outcomes is possibly the most valuable (and lucrative) service lawyers provide, and it was easily matched by a computer.

As computers become exponentially more sophisticated, it naturally follows that they will be able to perform more sophisticated work. The obvious downside is that these technological revolutions might not create as many jobs as they eliminate. Certainly, we will need more programmers, statisticians, engineers, data analysts and IT personnel to create and manage these sophisticated computers but not every factory line worker can easily shift gears and become a data analyst.

On the positive side, greater automation will be a boon in many industries, resulting in significantly increased accuracy and productivity. (Any lawyer, for instance, would agree that a faster, more comprehensive discovery phase is a benefit to the legal process.) Ultimately, I believe the possibilities for improving our business decision making, operational processes, products and customer experience are just too big to ignore.

## *Blockchain technology: the future of data and business?*

There's been a lot of hype lately about blockchain technology. In fact, a recent World Economic Forum report predicted that by 2025, 10 per cent of GDP will be stored on blockchains – making it something that every business leader should at least be aware of.[5]

What is blockchain technology? Well, we are used to sharing information through the Internet, but when it comes to transferring value (eg money), we usually revert to centralized financial institutions like banks. Even online payment systems like PayPal generally require a bank account or credit card to use them. Blockchain technology offers the chance to cut out the middle man by carrying out the tasks traditionally handled by financial services organizations, namely recording transactions, establishing identity and establishing contracts. It effectively enables peer-to-peer transactions, much like Bitcoin does (the virtual currency Bitcoin is underpinned by blockchain technology). A blockchain allows anyone to send value anywhere in the world where the blockchain file can be accessed. Each chain is essentially just an online database, stored in a distributed, peer-to-peer fashion among its users. Cryptography ensures that users can only edit the parts of the blockchain that they 'own' – by possessing the private keys necessary to edit the file. By giving private keys which you own to someone else, you effectively transfer the value of whatever is stored in that section of the blockchain.

Microsoft, IBM and many others have announced blockchain-based services, mainly aimed at financial services clients. However, there is vast scope to apply the technology across many other industries – after all, blockchains can be used to store any kind of digital information. The technology could be particularly useful in 'smart contracts', where contracts can be automatically signed off when stated conditions are met. For example, an invoice could be paid automatically when certain conditions are met, or a certain number of orders have been made. And payment could be made automatically using a blockchain payment system.

There's even a theory that blockchain technology could fuel the IoT. For example, devices in the home could automatically pay for the precise energy they use. 'Smart' local power grids could make use of blockchain technology to allow the distribution, metering and billing of electricity to be handled within communities themselves, which would be tremendously useful in remote communities.

I believe blockchain technology could be one of the most powerful data-related developments in the next few years and it's certainly worth business leaders staying abreast of this technology.

## Every business must become a data business

It's clear that data is becoming a key business asset, central to the success of every company. As the world becomes smarter and smarter, data becomes the key to competitive advantage, meaning a company's ability to compete will increasingly be driven by how well it can leverage data, apply analytics and implement new technologies. Data and the ability to turn data into business value will become increasingly important in every sector within a few very short years. In fact, according to the International Institute for Analytics, businesses using data will see US $430 billion in productivity benefits over competitors who are not using data by 2020.[6] In business, information is power, and big data is providing information we couldn't have dreamed of collecting or analysing just a few short years ago. Companies that don't evolve and embrace the data revolution will be left behind.

In addition to the growth of companies capturing and using their own data, we will see an explosion of the use of external data (from government sources, external providers, etc). Savvy companies are already predicting this, as with IBM's acquisition of The Weather Channel, mainly for its data.

The International Data Corporation predicts that over the next three to five years, companies will have to commit to digital transformation on a massive scale, including fundamental cultural and operational transformations.[7] Rather than using new technologies to complete old tasks, companies and IT departments will be looking at entirely new functions.

### *It all starts with a data strategy*

In order to thrive, business leaders will have to actively work to expand their thinking away from what has been traditionally done, and include ideas and systems that may never have been considered. Business leaders must begin questioning everything, starting with their strategy. I cannot stress this enough: if every business, regardless of size, is now a data business, every business therefore needs a robust data strategy.

perhaps with the help of an external data consultant. CDO or no CDO, it's clear that data needs to be a top priority for every business. And, as with every major business decision or investment, it all starts with a clear strategy – a roadmap for the journey ahead.

## Endnotes


1 Klaus Schwab (2017) *The Fourth Industrial Revolution*, Portfolio Penguin

2 Jane Wakefield (2015) Intelligent machines: the jobs robots will steal first, *BBC*, 14 September, available at: http://www.bbc.com/news/technology-33327659

3 Alan Tovey (2014) Ten million jobs at risk from advancing technology, *The Daily Telegraph*, 10 November, available at: http://www.telegraph.co.uk/finance/newsbysector/industry/11219688/Ten-million-jobs-at-risk-from-advancing-technology.html

4 Kim Ward and Daniel Martin Katz (2014) Using data to predict Supreme Court's decisions, *MSU Today*, 4 November, available at: http://msutoday.msu.edu/news/2014/using-data-to-predict-supreme-courts-decisions/

5 World Economic Forum (2015) Deep shift: technology tipping points and societal impact, available at: http://www3.weforum.org/docs/WEF_GAC15_Technological_Tipping_Points_report_2015.pdf

6 Bloomberg (2016) 6 predictions for big data analytics and cognitive computing in 2016, 6 January, available at: https://www.bloomberg.com/enterprise/blog/6-predictions-for-big-data-analytics-and-cognitive-computing-in-2016/

7 International Data Corporation (2015) IDC predicts the emergence of 'the DX Economy' in a critical period of widespread digital transformation and massive scale up of 3rd platform technologies in every industry, 4 November, available at: https://www.idc.com/getdoc.jsp?containerId=prUS40552015

8 Gartner (2016) Gartner predicts that 90 per cent of large organizations will have a chief data officer by 2019, 26 January, available at: http://www.gartner.com/newsroom/id/3190117

THIS PAGE IS INTENTIONALLY LEFT BLANK

# 02 Deciding your strategic data needs

Data is certainly exciting – revolutionary, even. But that doesn't always mean useful. To be truly useful, in a business sense, data must address a specific business need, help the organization reach its strategic goals, or generate real value.

I see too many businesses get so caught up in the big data buzz that they collect as much data as possible, without really considering what they want to do with that data. This isn't helpful (and, indeed, as we'll see in Chapter 10, it may land you in legal hot water in the future). Instead of starting with the data itself, it's vital every business starts with strategy. For now, it doesn't matter what data is out there, what data you're already collecting, what data your competitors are collecting, or what new forms of data are becoming available. For now, it doesn't matter whether your business has mountains of analysis-ready data at your disposal, or next to none. A good data strategy is not determined by what data is readily or potentially available – it's about what your business wants to achieve, and how data can help you get there.

The fact is there are many different types of data (as we'll see in Chapter 6). In order to find the right data for you, you must first define how you want to use data. You may need certain types of data for some goals and different types of data for others. Sensor data, for example, is extremely useful for increasing efficiencies in a manufacturing plant, but it's not going to help you predict demand for a new product, or understand how your customers feel about the service you provide.

There are countless ways data can help a business succeed but, broadly speaking, it comes down to the three categories outlined in Chapter 1: using data to improve your decision making, using data to drive operational improvements, and treating data as an asset in itself. In this chapter, I delve

deeper into these three categories to help you decide how best to use data in your organization. Chapters 3, 4 and 5 then set out the strategic process for each category.

In practice, even with huge resources, it's tricky to tackle all three categories at the same time. Guiding decision making is certainly the most prevalent way businesses use data today, and broadly speaking, in the majority of organizations, it's usually a good place to start. Therefore, most companies start with decision making and take it from there, building up to operational improvements and, potentially, data as an asset. However, for some companies – large manufacturers for instance – operational improvements may be the top priority. If that's the case in your organization, you can always skip the decision-making aspect for now and revisit it at a later date. And those companies with a mountain of customer data may well be motivated to start treating data as an asset right away. There are no hard and fast rules to making data work for your organization.

## Using data to make better business decisions

Making better business decisions is the goal for the majority of clients that I work with and I believe it's something that all businesses should work towards. Whether you want to better understand your market, develop a new product, increase revenue, or target new customers, it all comes down to making better, more informed business decisions. Data provides the insights needed to make those decisions.

Again, it helps to be as specific as possible about what data you want to use and how. In the context of making better decisions, you start by identifying your organization's priorities and unanswered business questions (such as 'how can we target this customer segment?' or 'how can we increase turnover by 10 per cent?'). You then source and analyse the right data to provide insights that help you answer those questions. In this way, having a clear data strategy helps you identify your key business questions and prioritize them, ensuring you use your time and resources in the most effective way. There's more on this in Chapter 3.

You may want to start by focusing on one specific area of the business, such as better understanding your customers, but the underlying idea of making more informed decisions, and building a culture of data-based decision making, should ultimately extend across the whole organization. This is a topic I'm particularly passionate about, and I explore it in more detail in Chapter 11.

## *Using data to better understand your customers and markets*

This is one of the most common (and most publicized) ways companies use data today, and social media in particular has made it easier than ever to build up a rich picture of customers and markets. With a deep understanding of customers and markets, the organization is able to make much smarter decisions – decisions rooted in data, rather than gut feelings or assumptions. There are three key strands to this: getting a full picture of your customers (who they are, where they are, their behaviour, their preferences, etc) so that you can better interact with them; identifying trends; and understanding the competition.

Building a complete picture of your customers can include what makes them tick, why they buy, how they prefer to shop, what they'll buy next, what makes them choose one company over another, and so on. Social media is an obvious and powerful source of this type of information. All of the main social media platforms, including Facebook and Twitter, offer targeted advertising, allowing you to target very precise age groups and geographical areas. Even without spending a penny, social media platforms can be used to see who is talking about what, and determine how that is likely to affect demand for products or services. Twitter – where pretty much all conversations play out in public – is easier to extract insights from than most platforms. In fact, IBM has now partnered with Twitter to offer a service allowing businesses to pull insights directly from tweets. Launching the service back in 2014, IBM gave some powerful examples of the insights that can be gleaned from tweets. These insights included a communications company which was able to reduce customer churn by 5 per cent by predicting where customers were most likely to be affected by loss of service due to bad weather. And a food and drink retailer discovered that high staff turnover was one of the factors which negatively affected the value of their most loyal customers.

Trend spotting is another popular use for data, whether it's industry-wide trends, customer behaviour trends, or indeed any kind of trends that could make a difference to the bottom line. Essentially, this comes down to spotting and monitoring patterns, and using that information to predict where things might go in the future so that you can make better decisions. Marketing is a great example of understanding and predicting trends and, again, social media and the Internet play a large role in this. As individuals, we're used to sharing vast amounts of data about ourselves, our interests, habits, likes and dislikes – whether knowingly or unknowingly – and savvy

(such as purchasing something in New York City at 2pm and in New Deli at 3pm using the same card in physical stores). The insurance industry has also made great strides in using data to detect fraud. By analysing the length of time taken to complete a claim online, or by analysing whether a client goes back and changes information on a previous page, they can flag up a potentially fraudulent claim.

Data can even help improve how you recruit and manage your staff – after all, your people are a vital part of your internal operations and processes. Sometimes, finding and keeping the right people can be the key to maintaining a competitive edge. Data can help you find the very best candidates, understand whether your current recruitment channels are effective, and help keep your existing employees happy. For example, one client of mine wanted to recruit self-driven people that were able to use their own initiative. By analysing different data sets from the type of people they wanted to recruit and those they wanted to avoid, the company found that candidates who filled out applications with browsers that were not pre-installed on their computers and instead had to be installed separately (such as Firefox or Chrome) tended to be better for that particular job. Measuring this simple indicator allowed the company to eliminate those that didn't meet the criteria before interview stage, thereby finding the right sort of people more easily. Another one of my clients, a retailer, analyses the social media profiles of candidates to (very accurately!) predict the level of intelligence and emotional stability of potential candidates.

The operational applications go way beyond recruitment. HR data, such as absenteeism figures, productivity data, personal development reviews, and staff satisfaction data, can all be analysed for insights. In addition to these somewhat traditional types of HR data, data can be captured in many new and exciting ways, such as capturing employees on CCTV, scanning social media data, analysing the content of e-mails, and even monitoring where staff are by using the data from geo-positioning sensors in corporate smartphones. Once again, the challenge is to establish which data is really going to make an impact on your company's performance, to avoid getting caught up in the overwhelming array of possibilities. You need to consider what is most useful from an operational perspective. It might, for example, be increasing employee satisfaction in order to reduce staff turnover. Plus, to avoid any sort of backlash, it's very important that staff are made aware of precisely what data is being gathered from them, and what it is being used for. Everyone needs to be aware that the purpose is to increase overall company efficiency, rather than assess or monitor individual members of staff in a Big Brother-type fashion.