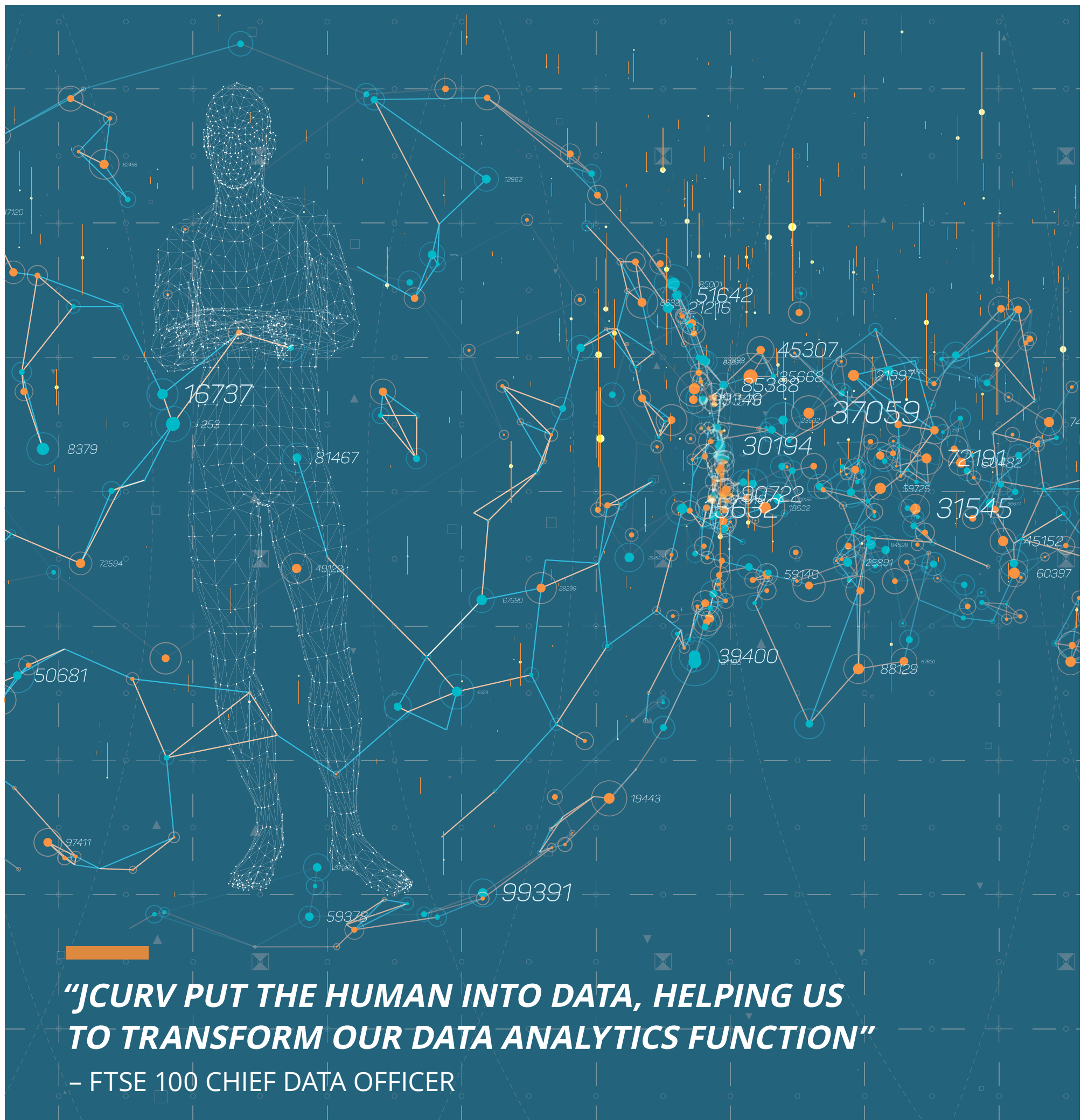


## 10 TREATING DATA LIKE A VALUED ASSET



## Connect, Prepare, Create and Automate your data analysis

**CLIC DATA**



***“JCURV PUT THE HUMAN INTO DATA, HELPING US  
TO TRANSFORM OUR DATA ANALYTICS FUNCTION”***

**– FTSE 100 CHIEF DATA OFFICER**

# **DATAGILITY**

JCURV brings culture, people and data together to help organisations maximise the value from data analytics

**[WWW.JCURV.COM](http://WWW.JCURV.COM)**





Contributors

- Cath Everett**  
Journalist specialising in workplace, leadership and organisational culture issues, she also writes about the impact of technology on business and society.
- Rupert Goodwins**  
Engineer-turned-journalist, he specialises in business and technology, and contributes to a range of national and international publications.
- Charles Orton-Jones**  
Award-winning journalist, he was editor-at-large of *LondonlovesBusiness.com* and editor of *EuroBusiness* magazine.
- Sooraj Shah**  
Contributing editor at *New Statesman* Tech, he also writes for Computing, IT Pro and a number of other business technology publications.
- Nafeez Ahmed**  
Investigative journalist and editor of *Insurge Intelligence*, he has contributed to *The Guardian*, *Independent*, *VICE* and *The Atlantic*.
- Suchandrika Chakrabarti**  
Freelance journalist and podcaster, her work has been published in *The Guardian*, *The Times* and the *New Statesman*.

Raconteur  
reports

- Publishing manager  
**Hannah Smallman**

Associate editor  
**Peter Archer**

Data editor  
**Michelle Ingham**

Managing editor  
**Benjamin Chiou**
- Head of production  
**Justyna O'Connell**

Digital content executive  
**Fran Cassidy**

Design  
**Joanna Bird**  
**Grant Chapman**  
**Sara Gelfgren**  
**Kellie Jerrard**  
**Harry Lewis-Irlam**  
**Celina Lucey**  
**Samuele Motta**

Head of design  
**Tim Whitlock**

Although this publication is funded through advertising and sponsorship, all editorial is without bias and sponsored features are clearly labelled. For an upcoming schedule, partnership inquiries or feedback, please call +44 (0)20 3877 3800 or email [info@raconteur.net](mailto:info@raconteur.net)

Raconteur is a leading publisher of special-interest content and research. Its publications and articles cover a wide range of topics, including business, finance, sustainability, healthcare, lifestyle and technology. Raconteur special reports are published exclusively in *The Times* and *The Sunday Times* as well as online at [raconteur.net](http://raconteur.net)

The information contained in this publication has been obtained from sources the Proprietors believe to be correct. However, no legal liability can be accepted for any errors. No part of this publication may be reproduced without the prior consent of the Publisher. © Raconteur Media

STRATEGY

# Uncovering the human side of data

The 'data versus humans' conflict is one that many organisations face, but companies must realise that software is no replacement for creative decision-making

Charles Orton-Jones

When Google wanted to discover the “stickiest” shade blue for a button it ran a multivariate test with 41 variations. Forget creativity – data provided a concrete answer. It's a classic example of how multivariate testing, or MVT, can be used to solve problems.

Whatever you want to know, MVT can offer insights. BGL Group, owner of ComparetheMarket.com and other online brands, recently faced a challenge when implementing cumbersome new GDPR and Insurance Distribution Directive legislation.

“We wanted to implement it in a customer-friendly way,” says Laura Mullaney, associate director of BGL Group. “Ahead of deadlines for implementation we were able to use MVT to test proposed customer journey designs with real customers using the website to determine the best approach. This enabled us to test customer response to some significant changes and understand the best approach to take for our customers, while adhering to our regulatory responsibilities.”

The company uses the same MVT logic to test all sorts of hunches. “We can test a journey or design even down to a demographic for insight on what works best to generate value. This can feed in to really significant initiatives, from rebranding through to development of adaptive or mobile-first design journeys,” Ms Mullaney reveals.

But here's the kicker. Data-centric thinking is a dominant force in shaping the customer experience. It is so powerful that less experienced brands are at risk of becoming dependent on it.

Gilles Boisselet, creative strategy officer of UNIT9, a consultancy working with brands such as the BBC, Huawei and Nike, says 'data versus humans' is a classic conflict. “Software salespeople will try to convince you that data testing and creative conceptualisation are the same thing. And companies looking to cut costs may think software can replace the people in charge of creative decision-making. But the reality is that data testing and creative decision-making are two different beasts.”

A lack of knowledge of the merits and limits of data methods like MVT are partly why, according to Gartner, 85 per cent of gig



Robin Lynne Gibson/Getty Images

data projects end in failure. In the worst case scenario, they forfeit creativity and become slaves to the data.

“Data testing more or less works from a mathematical base using dichotomy and statistical averages. So it trends towards value. But this just promotes uniformity and means it will become a struggle to differentiate yourself,” says Mr Boisselet.

“Genuinely creative people who can crack originality still need to be part of the process. Just look at how our new era of auto-completion is making our texts, emails and Google searches become homogenised. True creatives, like hip hop artists, know that you get

to rule culture by breaking the rules, not mathematically adhering to them.”

There's a second danger. The volume of data collected by companies can be overwhelming. Companies can be paralysed by the possibilities, and the chance to explore hundreds of avenues can result in years of procrastination.

“I worked with a FTSE 100 company on a project that they thought would take two years,” says Payal Jain, managing director of JCURV, an agile working consultancy in the data sector. “But you can't wait that long. In three months the market may have changed. The truth is that it doesn't matter whether you have

AI, robotics, or machine-learning, you still need to humanise the data process.”

Ms Jain's solution to the logjam is to create agile teams. These self-sufficient ensembles are able to come up with their own ideas, build models, and test them rapidly: “You may have 100 things on your list to look at. We say, break it down. Find the most important five. Then let an agile team take them on. This way you can see results as fast as two weeks.”

This approach to handling data also promotes creativity, without diminishing the contribution of data. “Give teams autonomy,” advises Ms Jain. “Those closest to the problem often know best. Another FTSE 100 client wanted to spend millions on a machine-learning and AI solution. But the team working on the job said we don't need to spend that. We have a batch process that does 80 per cent of the job. It worked and they save a huge amount of money. The truth is that you can spend a fortune on buying huge servers to store data, on buying the best data tools, but the money will be wasted unless you humanise your processes.”

As a happy dividend, giving data teams control over what they test and how leads to happier teams. “There is a war for talent,” says Ms Jain. “Data scientists are paid a lot. When leaders are prepared to embrace cultural change, and return power back to teams working closest to the data, they generate more loyalty and engagement, in addition to better results and top-line growth.”

The old cliché of “data is the new oil” was always a little tricky in practice. Oil is a commodity with a scientific refinement process. Data is altogether more slippery. Asking questions, drawing insights, and knowing how to react to evidence, is still a profoundly human endeavour.

In the end, there will always be a place for a creative leap. BGL Group increased market share by 76 per cent in a year when it gambled on a new creative direction with its bold and slightly bonkers Compare the Meerkat ad campaign. The talking aristocratic meerkat Aleksandr Orlov is still the face of BGL's most valuable brand. It's a reminder that although MVT is a powerful force for good, there is still space for old fashioned human ingenuity. ●

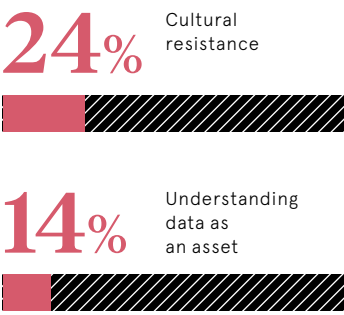
## BIGGEST CHALLENGES TO BUSINESS ADOPTION OF BIG DATA/AI

Survey of Fortune 1000 business and technology executives



Lack of organisational alignment or agility

NewVantage Partners 2018





## SOCIAL MEDIA

# Why the world needs trustworthy data from social media

People data from social media could further the understanding of economies and societies, but first we must explore why different motives for sharing data on various platforms may skew our insights

Suchandrika Chakrabarti

When the Cambridge Analytica data scandal broke, it damaged the relationship between consumer data and social platforms. Facebook chief executive and co-founder Mark Zuckerberg made his most recent attempt to address such security fears earlier this month, posting a 3,200-word essay to his platform, entitled 'A Privacy-Focused Vision for Social Networking'.

Mr Zuckerberg ends the lengthy post with: "We should be working towards a world where people can speak privately and live freely knowing that their information will only be seen by who they want to see it." Sadly, reputations are not repaired so easily. Drawing 10,000 comments, but only 3,600 shares, the post was "ratioed", a Twitter coinage meaning more people answered back to

a post, than liked it, or shared it, to express agreement.

"The timing of Mark Zuckerberg's essay is interesting, almost a year on since the Cambridge Analytica scandal," says Derek Roga, chief executive at encrypted communications firm Equiis Technologies. "But no matter what he says, essay or not, Facebook will never be known as a champion of data privacy. Consumers still remember how

Mark Zuckerberg testifying before Congress in the wake of the Cambridge Analytica data scandal

their personal data was commoditised and misused."

Facebook users gave up their data to the platform in a more innocent time, when signing up meant sharing personal content with friends and family. There was an expectation of privacy involved in the transaction. Compare this new Facebook user in 2005, or 2010, to someone setting up a LinkedIn account today. There are wildly differing levels of knowledge about data handling, as well as motives, at play in each scenario.

"For LinkedIn, the content of your profile has generally been something you want people to see; there are privacy settings available, but most of our members use LinkedIn to be found by colleagues, people they've met or potential employers," says James Upsher, LinkedIn's senior corporate communications manager, Europe, Middle East and Africa.

"One of the most popular features has been how it enables recruiters to find you and approach you with jobs that might be of interest. Our members are sharing information on the platform because they see direct value to themselves in doing so."

According to a survey of 4,000 people across the United States, conducted earlier this month by Jovi Umawing, senior content writer and researcher at cybersecurity firm Malwarebytes, just 12 per cent of respondents trust social media to protect their data.

The survey found that across the board there is a universal distrust of social media firms, among 95 per cent of respondents. Broken down by generation, baby boomers are the most distrustful of social media at 96 per cent, followed by gen Xers at 94 per cent. Gen Z and millennials showed distrust among 93 per cent and 92 per cent respectively, revealing a fundamental and comprehensive disconnect between social platforms and their consumers.

Why do these networks want this reluctantly surrendered data? Jim Cridlin, global head of innovation at media agency Mindshare, says: "Social data is incredibly powerful as it tends to be extensive, longitudinal and accurate. While consumers tend to post their best self, their interactions, likes, follows and so on tend to reveal their true self. Moreover, there is pressure to keep social data shared with friends roughly truthful, as friends will call out that which is untrue."

Beyond the veracity of the data, there's the sheer volume, which is a Holy Grail for marketers, according to Mark Brill, senior lecturer in future media at Birmingham City University. He says: "The potential power of data from social media is due not simply to its scale from billions of users, but also the frequency of their updates. That power to understand people has been frequently demonstrated. It would be possible to use this kind of data to identify the spread of infectious diseases, by marking the location and movement of

users. There is the potential to provide a powerful understanding of demographics, but it's harder to achieve in practice."

LinkedIn is using data for just that aim. The platform's economic graph is a digital map of the global economy, which individuals, businesses and governments can use to navigate the decisions they have to make, says Mr Upsher.

He adds: "Through the information that our 610 million-plus members share on their public profiles, properly anonymised and aggregated, we're able to help reveal trends. These insights are taken from all members, rather than a survey sample, and they are visible in real time. Insights from the economic graph, we hope, will remove personal

# 76%

of the UK public is concerned with sharing data on social media

Deloitte 2018

“There is the potential to provide a powerful understanding of demographics, but it's harder to achieve in practice

and organisational barriers to economic opportunity."

Recent projects powered by LinkedIn data include partnering with the World Economic Forum to look at gender gaps in artificial intelligence, working with the International Labour Organization, tracking how quickly women make it to director-level roles and an internal report on post-Brexit migration to the UK. The key takeaway from that last report is that fewer professionals from the rest of the world are moving to the UK.

User data could be the basis of more work that furthers understanding of the economy, society and our collective attempts at progress. Or it could be used to influence elections. Social networkers' data is out of their control. Platforms and governments need to work together to keep data in the right hands, for the right reasons. ●



JIM WATSON/AFP/Getty Images





# Data-driven businesses race ahead but change doesn't come easy

**Nick Bouch**, data and analytics in assurance leader and partner at PwC, discusses how the role of data has evolved in organisations, and why businesses must now prioritise technology and culture-driven change to succeed and grow in the digital age

For many years, issues relating to data were viewed in most organisations as the responsibility of the IT department. But as newly formed teams in areas such as customer insights and risk analytics took an interest in their own data-processing, analytics and reporting, a long-running battle between IT and the business ensued around who is accountable for data quality and accuracy, and who is to blame if something goes wrong.



**Nick Bouch**  
Data and analytics in assurance leader and partner, PwC

Data itself is not inherently wrong; it's the processes of capturing and transforming it that causes data quality issues, as well as how it is used by the end-user in the business. IT may provide the resources to build the data-processing and storage systems, but the business typically designs the logic and approach for how it works its way through the systems. Teams that sit in the business are usually the first to realise when data is wrong, but are unable to fix issues that have occurred during capture or processing.

Regulations such as the European Union's General Data Protection Regulation and revised Markets in Financial Instruments Directive, as well as costly cyberattacks, have now launched the data agenda to the very top of organisations. C-suite directors have also realised just how important data is to the business, including the risks of not securing it correctly, the implications of getting regulatory and statutory reporting incorrect, the ability to apply advanced analytics and machine-learning to improve the supply chain, and its role in developing new products, and driving customer insights and engagement.

Data is in everything we do and is at heart of the most successful businesses. It's behind how you watch TV, which products you buy online and what news you consume. Organisations operating in slow or cumbersome data environments are simply unable to adapt to the fast-changing needs and expectations of customers in the digital age.

New market entrants are disrupting established firms by being data led from day one. They may make it look easy, but companies with lots of legacy systems face far greater challenges and costs to reaching such a state of data enablement. Ripping out existing environments and putting a data-driven architecture in place is hard, expensive and scattered with failed programmes. Even in the world of big data, there are as many



**Organisations that successfully make the transition to being data enabled start small, fail fast and prove the value**

stories of failed and canned projects as there are of successful implementations.

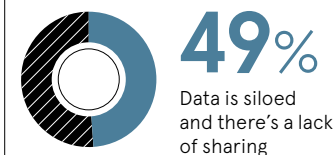
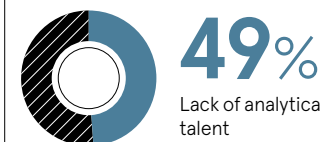
What's even more difficult than the implementation of a new data architecture is the cultural change organisations must go through to get everybody to understand the value of making decisions using data-driven insight rather than gut feeling. When people have been used to doing something in a certain way for a long time, it's a tough journey.

In PwC's 22nd Annual Global CEO Survey, 54 per cent of UK chief executives said there is a lack of analytical talent in the marketplace. They can't get the people even if they want to, making improving the data literacy of their existing staff absolutely crucial. Everybody can and should understand how data can be used to drive business value, the ethical implications of how data is exploited and how to secure it. By upskilling staff, companies won't need to recruit such huge amounts of highly expensive analytics resources.

Organisations that successfully make the transition to being data enabled start small, fail fast and prove the value. They have to be able to exploit and bring together their data faster and with greater agility, but they know they can't change overnight. They adopt an evidence-based approach to data and prioritise where to spend their efforts. However, many firms are not able to exploit their data in the most effective manner to drive the best insights when the data is stored across different silos, which over half of UK chief executives reported as being the case in the latest PwC Global CEO Survey.

Chief data officers (CDOs) are in a great place to build a programme of change. Historically, reporting to the chief information officer, the vast majority now fall under the chief operating officer, with some even on the board. With 96 per cent of UK chief executives in the latest PwC Global CEO Survey saying data on customer preferences and needs is important or critical, but only 13 per cent believing it to be reliable, it's clear somebody on the board, CDO or otherwise, must take full responsibility for data in the organisation.

## REASONS THAT DATA YOU RECEIVE IS NOT ADEQUATE OR THAT YOU DO NOT RECEIVE THE INFORMATION



UK data from PwC's 22nd Annual Global CEO Survey

Nobody can take a generic data strategy off the shelf and say that's how we should do it. A firm's culture drives how it adapts, uses its data and embeds insight into business processes. The organisations that have moved forward rapidly over the last few years have allowed the CDO to develop a strong data strategy and then exploited it with the right funding, leadership and sponsorship from the board.

Because organisations can't easily determine the value of their data, they find it hard to work out what the challenge would be if they didn't have a good data strategy in place. The problem is, while some CDOs are absolutely accountable for their organisation's data strategy, they're not always responsible. That means they can be on the hook if things go wrong despite, in many cases, the budget actually sitting inside the business.

Previously, the CDO was a very switched-on data expert who knew everything about data quality and governance. The next generation of CDOs, which has emerged in the last couple of years, is much more business led. They understand how the business works, can communicate effectively, build a strong team around them overseeing the more niche subsets like governance and quality, and drive the business forward around change.

As the role continues to evolve, we are likely to see more chief analytics officers looking after the exploitation of data through insight, while CDOs look after the data itself.

We need to think about data in the same way as other assets in an organisation. We all understand the importance of a secure building; doors are locked and if there is a water leak, a plumber is called immediately. Data should also be actively secured, with measures that ensure it is looked after, and enable it to be analysed across the firm and not in silos. It's fundamental to future success as data becomes synonymous with just doing business.

For more information please visit [www.pwc.co.uk/dataCI](http://www.pwc.co.uk/dataCI)



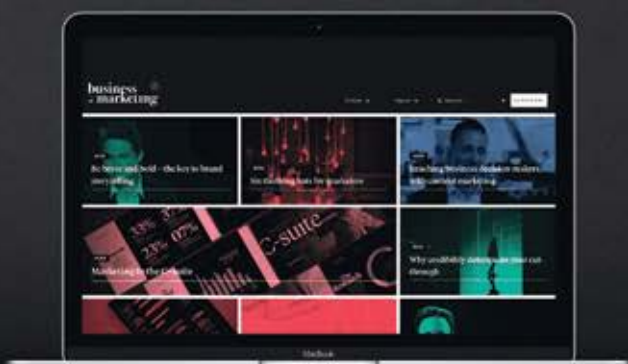


RACONTEUR

# Are you serious about business performance?

Then our new exclusive content hub, *Business of Marketing*, is for you.

This is not for the many, but for senior executives and marketers who are devoted to being in the top five per cent in their field.



[agency.raconteur.net/bom](https://agency.raconteur.net/bom)



UNSTRUCTURED DATA

## AI weaves value from unstructured data

Artificial intelligence is delivering benefits in the arena of unstructured data, helping companies to decipher insights and extract value from reams of unorganised information

Rupert Goodwins

**D**ata is pouring into companies in torrents, bearing unstructured information about markets, customers, resources and trends. Businesses know that it's something to be harnessed rather than feared, and are looking to artificial intelligence and machine-learning (AI/ML) to scry insights and value.

AI/ML is far from a fit-and-forget technology. For any business to embark on unstructured data-driven AI/ML, a lot of questions need answering. Starting with what is AI/ML in a business context?

"The best definition of AI I know is making machines do tasks that require human intelligence," says Fern Halper, vice president of research at data research and education company TWDI. "Companies have been using predictive analytics to do that in some ways for decades, but the uptake hasn't increased. Now we're in the early mainstream of AI/ML, that's going to change."

The most important thing, she says, is knowing what you want to do. Many of the companies TWDI survey are concerned with understanding customer sentiment, so are using natural-language analysis on unstructured sources such as email, customer reports and social media.

"Some 25 per cent are extracting sentiment and combining that with structured data from billing and service history to build a predictive model of who'll buy what," says Ms Halper. "The results are better than you can get from structured data alone."

**“The ability to link unstructured and structured data is where the value lies; you get new insights and reveal unknowns**

That sweet spot of identifying where AI/ML can extract meaning from unstructured data that then reinforces structured information is a productive one, says Nick Lynch, consultant at the Pistoia Alliance, a pharmaceutical industry not-for-profit organisation that identifies

and develops cross-company techniques and tools.

"Unstructured data is a field in transition," he says. "The health industry has lots of it, from early science to doctor's notes, images, histories, formulae, genetics. It's very valuable, but deriving interesting things from it is hard. Industry is making big improvements in unstructured data and the ability to link it to structured data is where the value lies; you get new insights and reveal unknowns."

Knowing whether you have the right data, what the right approach is and whether the results are working are all early and important stages on the path to AI/ML. That needs good people or good partners, says Ms Halper. "Talent is the number-one challenge," she says. "Then comes getting good quality data and integrating it. Here, all the vendors are trying to make it easier for anyone, not just developers, to use AI/ML and put it in apps. Regardless of how easy this is, explainability is very important. You need to explain how your model works for compliance and to foster confidence within your company. Is your model defensible? The market is moving this way."

Trust is important, both for buy-in within a company and when providing high-value, high-risk results such as diagnoses from medical images. The same is true where AI/ML is used to automate the most mundane, everyday tasks, says Marcos Jimenez, chief data scientist and co-founder of X.AI, which has developed an AI called Amy that arranges business meetings.

Amy is an example of another growing class of AI/ML, the agent or bot that replaces humans in an interactive task. "Amy is a task-oriented dialogue engine. It understands





commands like 'set up a meeting with Jim and Sally' and arranges it by email," says Mr Jimenez.

It's a precise example of automating a task that uses human intelligence and one that was created with a great deal of human input building another vital aspect of AI/ML, domain knowledge.

## MOST DIFFICULT GOALS AND OBJECTIVES IN DATA MANAGEMENT

Survey of IT leaders across North America

STRUCTURED DATA	UNSTRUCTURED DATA
Insight	Automating data workflows
Security	Governance
Governance	Insight
Mobility	Evaluating new technology
Protection	Accessibility

Igneous 2019

"We looked at thousands of situations, extracting elements for natural-language processing. Initially, humans in the loop called AI trainers labelled stuff to kick-start the process. We used ML to create datasets that characterise meetings from hundreds of thousands of hand-labelled emails. It was super-expensive to build," he says.

The AI then knew what a meeting was and how humans talked about it. After this, adjusting a meeting to what people wanted was a much easier task, as was creating a natural-language output system that could use email to get or send information.

Text is the ultimate unstructured data, says Mr Jimenez; you never know what you're going to get and it's never the same as what you've seen before.

"You have humans in the loop at first and, as your model gets better, you can move them out," he says. "Eventually, the customer becomes the human in the loop because human language always has ambiguity and incomplete

cases. Amy is trained to ask for clarification when this happens, but that's exactly what humans do. We measure Amy's performance against that of humans doing the same task and it's just as good, so we have a product."

Your company doesn't have to be so AI/ML focused, says Ms Halper, but as awareness grows there can be unexpected benefits. "I see a virtuous circle, a success cycle," she says. "You start to see how your infrastructure can work with all sorts of data, and how all techniques like data warehouses, data lakes, cloud and virtualisation can support AI/ML. The company becomes more data aware, more confident with analytics and new applications become apparent."

Expect some pain at the beginning, she advises, but pick early tasks that can show obvious success and grow from there. "Data is alive, organic; it's not structured. Evolving the tools and understanding to use it won't be optional for long," Ms Halper concludes. ●

# Be more efficient, profitable and de-risk your business by harnessing unstructured data

Access \$430 billion dollars worth of untapped opportunities by using our Skim Engine™ and Data Science expertise to make your data intelligent.

Enterprises are throwing away opportunities and insights, by not incorporating data effectively into their organisations. 16% of a business' data is unknown, but a larger 80% is unstructured which makes accessing and using that data difficult.

Our Skim Engine™ brings much needed structure to your data. Be it, internal documents or external sources so you can start unlocking the potential. We help you innovate by building Artificial Intelligence that scales, and delivers real business value.

Speak to one of our PhD data experts today about how we can help you become empowered by data.

✉ sales@skim.it  
☎ +44 (0) 207 129 7497  
🌐 skimtechnologies.com  
🐦 @skimit

# skim

## Examples of unstructured data

Companies already have vast and increasing amounts of unstructured data that can help them with customer and product development.

Report fields in customer relationship management systems, emails from customers and customer discussions of your company in social media contain huge amounts of information about how they see and use you. Plain text is still one of the most important communication methods and repositories of historical data inside companies and on the internet, and sentiment analysis through

natural-language processing can extract value in many ways.

Financial services companies that provide mortgages need to know the current value of properties and predict values over time. This comes from land registry records and price databases, but pictorial data from online street views and aerial maps, information in estate agents' sales literature and property reports, and even traffic flow and other infrastructure developments in the area can improve model quality.

Retail outlets can create and retain enormous amounts of CCTV data for security purposes, yet it also maps out customer flow in the store throughout the day. This can be correlated with

other information about weather and regular local cultural or sporting events that may generate different types of crowd. Reaction to changes in store layout, display, product mix and ambient temperature or sounds can also be gauged.

Exterior CCTV data at venues can provide insights into how visitors arrive, what mode of transport they use, providing demographic information and enabling better targeting of services and the use of spaces around the venue. The build-up of choke points or underutilisation of particular entrances can be automatically mapped, optimising retail exposure and improving customer experience.



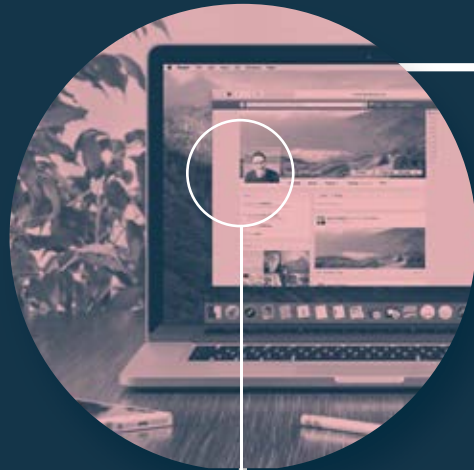
# A DAY IN DATA

The exponential growth of data is undisputed, but the numbers behind this explosion – fuelled by internet of things and the use of connected devcies – are hard to comprehend, particularly when looked at in the context of one day

**500m**

tweets are sent  
every day

Twitter



**4PB**

of data created by  
Facebook, including

**350m** photos

**100m** hours of video  
watch time

Facebook Research

**294bn**

billion emails are sent

Radicati Group

**320bn**

emails to be sent  
each day by 2021

**306bn**

emails to be sent  
each day by 2020

**3.9bn**

people use emails

**4TB**

of data produced by a connected car

Intel

## ACCUMULATED DIGITAL UNIVERSE OF DATA

**4.4ZB**

**44ZB**

PwC

2019

2020





DEMYSTIFYING DATA UNITS

From the more familiar 'bit' or 'megabyte', larger units of measurement are more frequently being used to explain the masses of data

Unit	Value	Size
<b>b</b> bit	0 or 1	1/8 of a byte
<b>B</b> byte	8 bits	1 byte
<b>KB</b> kilobyte	1,000 bytes	1,000 bytes
<b>MB</b> megabyte	1,000 <sup>2</sup> bytes	1,000,000 bytes
<b>GB</b> gigabyte	1,000 <sup>3</sup> bytes	1,000,000,000 bytes
<b>TB</b> terabyte	1,000 <sup>4</sup> bytes	1,000,000,000,000 bytes
<b>PB</b> petabyte	1,000 <sup>5</sup> bytes	1,000,000,000,000,000 bytes
<b>EB</b> exabyte	1,000 <sup>6</sup> bytes	1,000,000,000,000,000,000 bytes
<b>ZB</b> zettabyte	1,000 <sup>7</sup> bytes	1,000,000,000,000,000,000,000 bytes
<b>YB</b> yottabyte	1,000 <sup>8</sup> bytes	1,000,000,000,000,000,000,000,000 bytes

\*A lowercase "b" is used as an abbreviation for bits, while an uppercase "B" represents bytes.

463EB

of data will be created every day by 2025

IDC

95m

photos and videos are shared on Instagram

Instagram Business

65bn

messages sent over WhatsApp and two billion minutes of voice and video calls made

Facebook

28PB

to be generated from wearable devices by 2020

Statista

Searches made a day

5bn

Searches made a day from Google

3.5bn

Smart insights

Google



## INFONOMICS

# How to treat data as a valued asset

Focused on valuing data as a balance-sheet asset, "infonomics" is a nascent concept, but it could revolutionise how boards view and treat their corporate intelligence

Cath Everett

**D**espite the seemingly endless assertion that data is the new oil, it seems most organisations are treating it as neither a valuable nor a dangerous asset, even though it can be both.

Doug Laney, vice president and distinguished analyst at Gartner's data research business, says information-savvy companies with a chief data officer, enterprise data function and data science professionals currently have a book value twice the market average.

On the other hand, companies can lose millions if their data is stolen or subject to breaches, a situation that also affects both their credit score and reputation.

Nonetheless, fewer than half of all firms make any effort to measure the quality of their data, Mr Laney says, while only about 10 per cent gauge its value in any way other

than quality, for instance based on its utility or how it is used within the organisation.

"So most companies have a latent asset that isn't being measured or managed well, which means it's not offering an optimal level of economic benefit," he says. "While many organisations understand this, there's been no process available to help them do this."

As a result, most firms manage assets, such as office supplies and furniture, with more discipline than they do their information assets.

In a bid to inject some discipline, Mr Laney has come up with an economic concept he terms "infonomics" to measure, monitor and monetise information as an enterprise asset.

The idea came to him in the wake of the 9/11 terror attacks on the New York Twin Towers, in which many

“

**Most companies have a latent asset that isn't being measured or managed well, which means it's not offering an optimal level of economic benefit**

organisations lost not just their people and property, but also their data, only to find that their property and casualty insurance policies did not cover this kind of intangible electronic asset.

At that point it became clear there was also no mechanism to capitalise data as a balance-sheet asset and no defined method for valuing data. So to expound his theory on how to tackle such issues, Mr Laney published a book on the topic in 2017, although he acknowledges that it is still a "fairly nascent industry concept" that has yet to see widespread take-up.

Neil Walker, technical director of engineering consultancy Sweco, agrees. "Most people don't know what it means, and awareness and understanding of the term is low," he says. "But where we're seeing it used is in the context of systems thinking, which it links to by putting an agreed value on information assets and how they are communicated throughout the organisation."

Other early adopters include asset managers, who use the approach to help with investment decisions, equity analysts, who employ it to value information assets as part of a merger or acquisition, and

insurance companies, which are used to associating data with a given value.

Mr Laney also cites a US security systems firm, which used infonomics to identify underperforming information assets and come up with a number of ways to better exploit its data, leading to an increase in market value of \$300 million.

Meanwhile, an energy company discovered that some of its information was costing it more to collect, store and secure than the value it was generating. So the decision was taken to dispose of it, a move that saved the organisation more than £1 million a year on infrastructure costs in the process.

As Stijn Christiaens, co-founder and chief technology officer of data governance software provider Collibra, points out: "Conceptually and intellectually, people are intrigued by the idea, especially data specialists who believe that, if a dollar amount can be put to their data, the scales will fall from executives' eyes and they'll see the value of it. But while there's definitely activity, most organisations are still at an early stage."

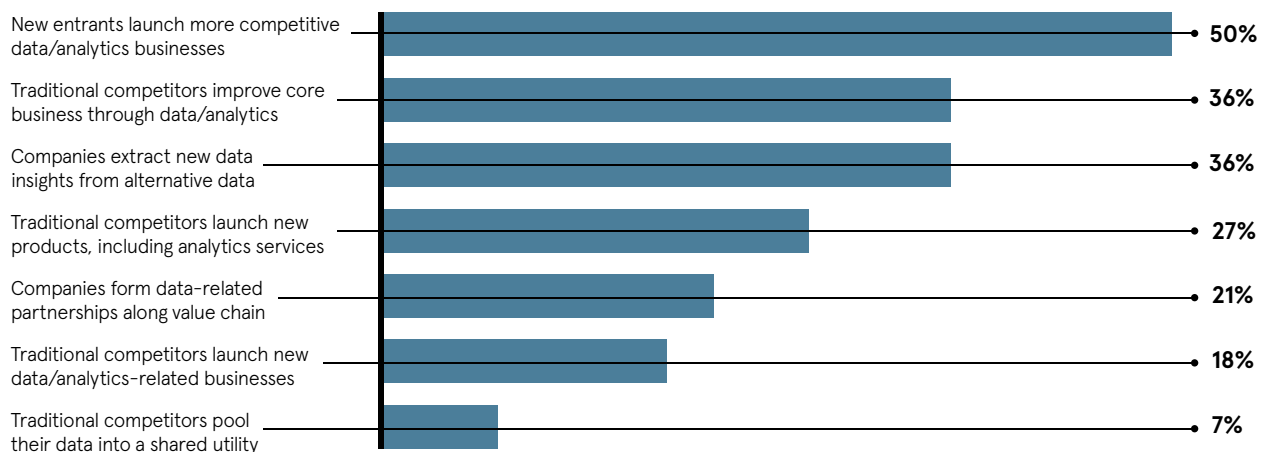
He also believes that, despite the concept's merits, it will take some time to be adopted by



Kim Steele/Getty Images

## HOW DATA AND ANALYTICS HAVE CHANGED INDUSTRY COMPETITION

Survey of C-level executives



McKinsey 2018



organisations' finance functions, not least because Mr Laney's formulas for valuing data as an asset are currently not accepted as industry standards.

"Doug's work is the best thing available, but it's still not sufficient as there's too much possibility for subjectivity in the formulas," says Mr Christiaens. "Bean counters need to count beans, not apples and oranges, so there has to be a standardised way to do it which is repeatable and auditable."

One of the issues relates to the intrinsic value of a dataset, such as a list of customer records. Under the infonomics approach, this intrinsic value is based, among other things, on the data's accuracy and completeness.

"But because the rules in the formula to determine the actual values are subjective, it means that if, say, PwC values my customer dataset, but I don't think it's valued highly enough, I could go to KPMG to see if I can get a higher valuation," Mr Christiaens explains.

There are also issues if that same customer dataset is sold on. For

example, if an organisation sells one million customer records to a given company for £1 million, the value of that data is £1 million to both the seller and the purchaser. But if the same data is also sold to nine other rivals for a further £1 million each, the value to the seller becomes £10 million.

This scenario creates problems in terms of infonomics-based calculations though. "The data amounts to intelligence, so if you're the only one that has this intelligence, it's worth a lot. But it's worth a lot less if everyone has it, so the value goes down if all the rivals buy it," says Mr Christiaens.

As a result of such ambiguities, he believes that while Mr Laney has done great work in developing a platform and baseline for discussion, the focus now should be on how it can be developed and improved upon.

Mr Laney concludes: "A big challenge is that the accounting standards bodies have not yet come round to allowing for information to be categorised as an asset on the balance sheet." ●



## Beginners' guide to implementing infonomics

When going down the infonomics route, a good starting point is for the chief information or data officer to create an inventory of company data. This includes understanding how and why such information is used on an organisation-wide basis and classifying it appropriately.

As infonomics inventor Doug Laney says: "You can't manage what you don't measure and you can't monetise what you don't manage, so by measuring it, you have to manage it, which creates more opportunities to generate benefits from your assets."

The next step is to assign the data a value based on its actual or potential value. To do so involves applying a range of defined infonomics

formulas, such as measuring the information's quality, business relevancy or impact on key performance indicators where relevant. Other possible approaches include adapting traditional approaches for valuing such assets as software licences.

Once the data has been valued, it is then helpful to list and prioritise possible use cases before acting on them. These might include employing machine-learning technology to predict customer churn rates or better target certain demographics to boost sales. They might also include establishing whether it makes economic sense to buy more storage infrastructure or simply delete non-essential information. Doing so could not only save the organisation money in administration and equipment, but also reduce its exposure to data breaches.

# Discover your data

With increasing regulation of data, organisations must ensure compliance or face big fines



**F**or most of us, January 21 was a day like any other. But for those tasked with safeguarding our data, it was a watershed moment for privacy. For this was the day when Google, a multinational technology company, was fined £44 million for breaching the European Union's General Data Protection Regulation (GDPR). And with California, which is home to the largest concentration of tech companies in the world framing its own data protections laws, many believe the fine meted out to Google is just the beginning.

Take Oksana Sokolovsky, an entrepreneur and smart data discovery expert, for example. She believes that both GDPR and the California Consumer Privacy Act (CCPA), which comes into effect next January, are game-changers for regulatory compliance and not just for the Silicon Valley tech titans, but for any large enterprise.

"We all live and work in a data-centric world," she says. "The internet of things, which has enabled everyday objects to collect and exchange real-time data, has created an abundance of information. With technology moving at such a fast pace, companies receiving the information cannot process it fast enough. The result? Unfathomable and unnavigable 'lakes' of data or 'dark data' as many in the industry refer to it."

For Ms Sokolovsky, who is also chief executive and co-founder of Io-Tahoe, a firm specialising in smart data discovery, the consequences of companies failing to act could run into millions of dollars.

"Firms need to take a proactive stance or data protection regulators will," she warns. "For a chief data officer, there are only three questions that matter.

They are: 'what sensitive data does my enterprise possess?', 'where is it stored?' and 'why do we have it?' By addressing these questions, companies can begin to implement the necessary foundational steps. Without them in place, it is impossible to implement the correct regulatory policy and controls needed to safeguard personal data."

**“**  
**Firms need to take a proactive stance or data protection regulators will**

One firm addressing GDPR compliance is Centrica. With more than 25 million customer accounts, the British multinational energy and services company handles a significant amount of data. The problem was that it didn't know what data it had and where it was housed.

But Centrica did know one thing: if it was to conform to GDPR, it would need forensically to demonstrate to the Information Commissioner's Office that it had the answers. It therefore decided to embark on a data discovery exercise. However, with 30 billion records, 1,200 databases and 1,500 applications to process, it soon became clear it needed a partner skilled in smart data discovery to enable regulatory compliance.

And that's when it turned to Io-Tahoe for help. Over a four-month period, working in partnership with Centrica, Io-Tahoe not only helped to enable Centrica's GDPR compliance, but in shining a spotlight on its dark data, it meant the data relationships Centrica had previously been unable to identify or reconcile were suddenly revealed.

Mike Young, Centrica's group chief information officer, explains: "Thanks to Io-Tahoe, we have a far better understanding of where all our customers' personal and sensitive data resides. Doing change programmes, for example, gets you thinking about privacy up front instead of an afterthought. You realise that privacy is not merely a bolt-on and that more and more of what we're doing is embedded in our technology. This has helped us to address emerging issues like ePR, and has even enabled us to start up a programme for data privacy by design."

Back at Io-Tahoe's New York headquarters, Ms Sokolovsky is expecting the phone to keep ringing. "Our product helps enterprises to become GDPR or CCPA compliant, ahead of next year, quickly and efficiently, but with legislation becoming ever stricter, it also demonstrates that while manual data initiatives may have been an option for companies five years ago, they have no place in today's fast-moving data-driven world."

**For more information please visit:**  
**[www.io-tahoe.com](http://www.io-tahoe.com)**

**Io-Tahoe**





sanjari/Getty Images

## DATA SCIENTISTS

# Specialist data scientists in high demand

As the need for specialised expertise increases in the role of extracting knowledge and insights from data, have we seen the end of the generalist data scientist?

Sooraj Shah

**D**ata scientists have for many years been considered a rarity which large enterprises have been fighting to get hold of.

Karl Hoods, chief digital and information officer at the Department for Business, Energy and Industrial Strategy, explains that as organisations have increasingly realised the importance and value of data, their need for the right expertise has also increased.

"In trying to make sense of the vast amount of data gathered on a daily basis and use it to solve business problems, spot trends and make decisions to support new ideas, we need people with a mix of statistics, databases, data visualisation, machine-learning, coding and data preparation skills," he says.

That mix of skills is exactly why data scientists are a rare breed. And the demand for them does not seem to be disappearing. According to a recent *Data Surge Report* by MHR Analytics, 80 per cent of UK companies are planning to hire a data scientist or seek data consultancy this year.

Bernard Marr, data expert and author, suggests that while there is currently a massive talent gap in data science, as the number of data professionals increases over the coming years, companies will be looking for people with more specialised skillsets. This includes data science leaders, data translators, who can provide the link

between the data function and the business, and domain-specialist data scientists.

The separation of different types of data scientists may have occurred because current data professionals have found themselves having to cover too much ground to manage what's being demanded of them.

"This can sometimes be a case of trying to master both the technical and the business understanding and communication aspects of their role," says Iain Brown, head of data science at SAS UK and Ireland.

SAS carried out research that found more than half of data scientists were reportedly suffering from some level of work-related stress due to such a broad remit.

"The shift towards sector-specific and less general data scientists is probably one that needed to happen anyway as the role was becoming too stretched," Mr Brown says, adding that some organisations

may need to consider skill-specific data scientists too.

However, Caroline Carruthers, a former data leader at Network Rail and Lowell, says that even in the future, data scientists won't necessarily only be hired if they have domain knowledge.

"Data is still data whether it sits in a retailer or a charity. It's more important for employers to find data scientists who are able to harness their unending curiosity coupled with a logical and creative mind, and a background in making a quantifiable impact on their business," she says. Mr Hoods adds: "Domain experience is useful, but it can be learnt."

Michelle Wren, human resources director at Freedom Finance, recently had a requirement for a data scientist with a broad skill set encompassing data, analytics and soft skills, without the need for domain expertise.

"A recent hire took five months to recruit, and we found many applicants could be more selective with where and what they settle for," she says.

The power, therefore, is still very much with data scientists. But with new technologies on the horizon, there is a view that the data professionals could automate themselves out of a job.

Matt Cockbill, head of the IT and digital leadership practice at Berwick Partners, isn't convinced. "It may have some hint of truth for lower-value data skills, but

80%

of UK companies are planning to hire a data scientist or seek data consultancy in 2019

Deloitte 2018

high-value, full-stack data science roles will continue to reform and reshape the world around them, creating continued heavy demand," he says.

"Data scientist is a broad catch-all title. While we will see more specific career paths evolve, the bubble around data science and data engineering skills isn't set to burst."

Technology will give these professionals more capabilities, but it won't mean their roles will be made redundant in the years to come. As Ms Carruthers says, it just means the overall level of data literacy will improve across the workforce, with other employees gaining a better understanding of how to use data more widely.

"The future data scientist can therefore be even more specialised, tackling the most business-critical and complex challenges that will help their businesses make revolutionary leaps forward," she concludes. ●

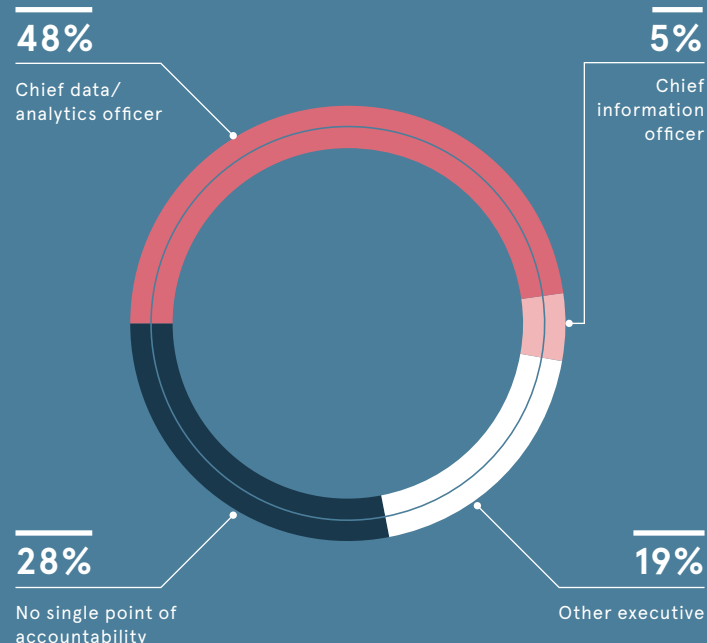
“

The future data scientist can tackle the most business-critical and complex challenges that will help their businesses make revolutionary leaps forward



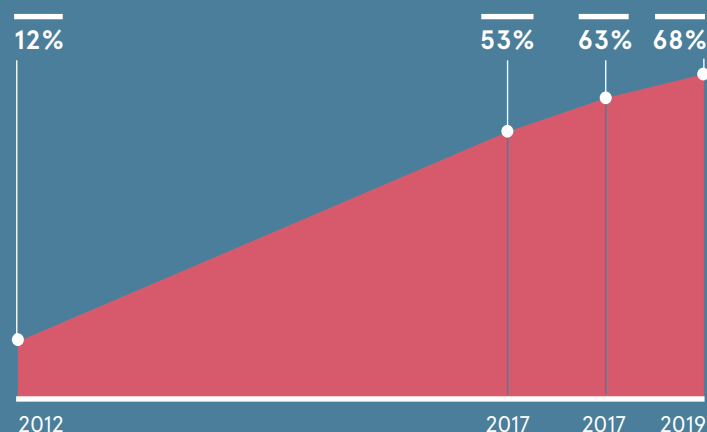
# DATA PEOPLE

PERSON PRIMARILY RESPONSIBLE FOR COMPANY DATA\*



Percentages do not equal 100 due to rounding

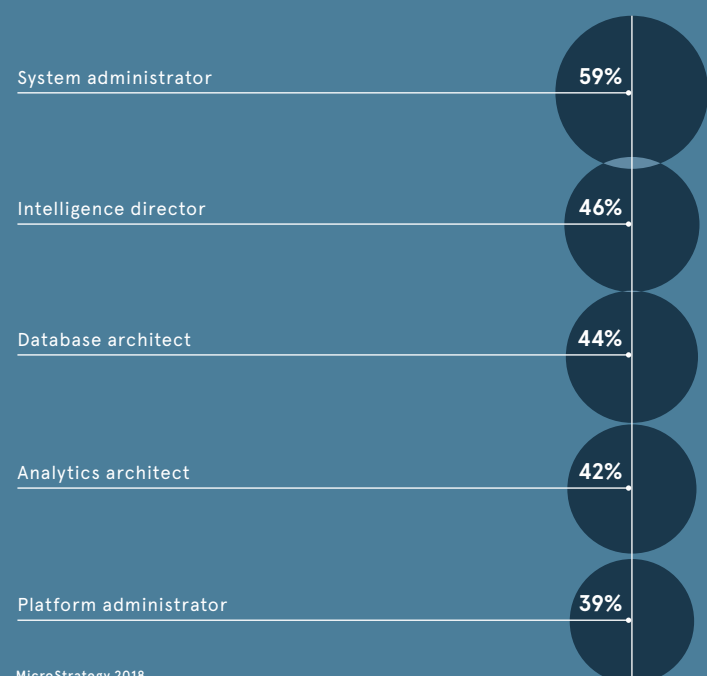
PERCENTAGE OF FIRMS WITH A CHIEF DATA OFFICER\*



\*Survey of Fortune 1000 business and technology executives

NewVantage Partners 2018

COMPANIES WITH THE FOLLOWING ROLES IN DATA AND ANALYTICS



MicroStrategy 2018

## Clean data drives better decision-making

High-quality data and its full understanding across an enterprise is increasingly crucial to ensure decision-making is free from bias and driving innovation in the digital economy

**D**ata is no longer an issue confined to the IT team. Its quality, trust and understanding within an organisation is paramount in decision-making. The real value of information is not in its technical nature or how it supports applications, but rather how knowledge workers within the business can comprehend and leverage it in new ways.

Truly understanding data means thinking about the environment it lives in. Businesses need to know what systems are involved and how the data is manipulated on its path between applications, data lakes and reports. It also increasingly requires knowledge of third-party data sources.

"When you ask people if their data quality is good, they'll almost certainly say no," notes Rex Ahlstrom, chief strategy and technology officer at BackOffice Associates, which provides solutions for improving data quality across the enterprise. "The narrow way people used to look at data has really expanded into a multi-dimensional problem of trying to get a better handle on the trust and understanding of data."

Data quality has been an issue facing companies for many years. Previously, if data was unclear or inaccurate, the whole process broke down and required manual intervention. With the emergence of analytics and machine-intelligence tools, however, data quality has become even more crucial because, if the data is biased, it will lead to bad decisions.

When utilised successfully, data-driven decision-making can help grow revenues and serve other core business objectives. Its value may be as simple, yet effective, as improving



internal processes to boost efficiency and minimise the cost of execution.

But a robotic process automation tool making automated decisions about supply or inventory tools, for example, could prove to be very costly if it has been fed poor or incomplete data. Biases in the data itself can seriously hinder business decision-making.

As companies are increasingly influenced by automated methods of decision-making, it's vital to ensure the data flowing into systems is not only complete, but can also be backed up for compliance purposes when reports are delivered to governing bodies.

"The challenge becomes much more sophisticated," says Frank Schuler, BackOffice Associates' vice president SAP technical architecture. "Knowing the underlying data is good quality keeps companies out of court and reduces compliance failures, as well as optimising processes, reducing costs and driving innovation. Data could look perfectly fine, but is driving biased decisions so data quality becomes enormously more important."

Enterprises need policies and rules that ensure data is managed in an effective way, and staff can see how it flows across the organisation and impacts decision-making. Only a complete understanding of where the data came from, how it changed on its journey and its current state can enable them to comprehend all its source and lineage.

Often companies don't understand what can be archived or decommissioned and the impact of doing so across the enterprise. By cataloguing

all their systems, BackOffice Associates provides companies with a Google-like experience of searching and immediately discerning how data is being distributed throughout the organisation.

"Few companies know all the potential locations where their data resides," says Mr Ahlstrom. "There are so many applications, on-premise and in the cloud, structured and non-structured, that it becomes very hard to quantify. Cataloguing those environments enables everybody to understand what lives within the applications."

There are many tools focused on only one thing, such as analysing data quality, but companies need to empower their knowledge workers to contribute to understanding data. The challenge is getting line-of-business leaders to care about data quality.

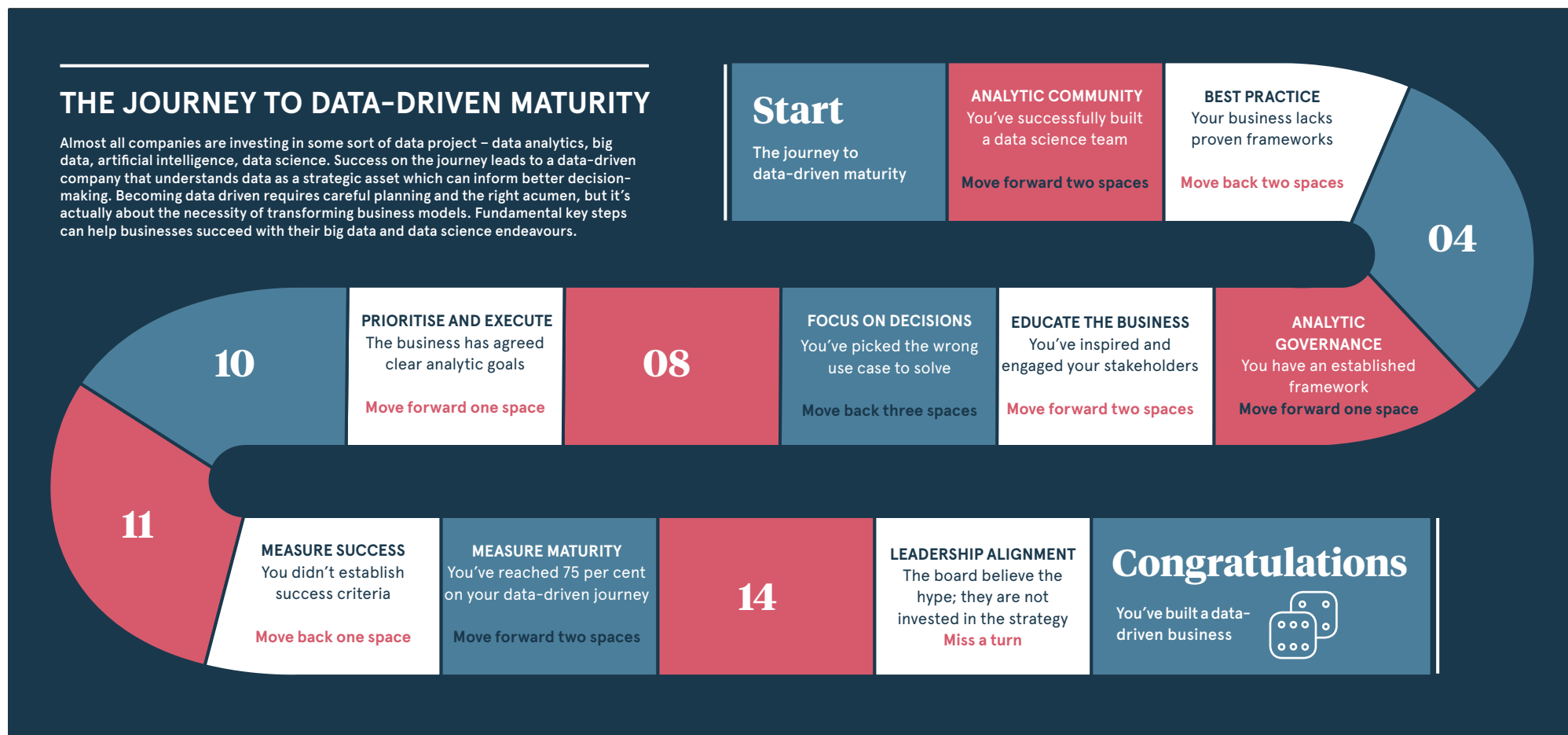
"You have to orchestrate the processes by which they participate and provide ways to become data contributors in an effort to crowdsource the improvement of information at the enterprise level; that orchestration of resources is crucial," says Mr Ahlstrom. "BackOffice Associates creates new user experiences that transform a logical or physical understanding of data into a business-level understanding of information."

For more information please visit [boaweb.com](http://boaweb.com)

**BackOffice**  
ASSOCIATES

**“The narrow way people used to look at data has really expanded into a multi-dimensional problem of trying to get a better handle on the trust and understanding of data**





# Data: a cultural transformation and not a quick fix

Amid stronger business competition than ever before, companies need to do more than simply embrace buzzwords or trends. Emphasis instead should be to instil transformation into their very being, especially when it comes to the role of data

**Q** uick fixes are not the order of the day and, while the utilisation of tools such as artificial intelligence (AI) and machine-learning (ML) may reap initial rewards, focus needs to switch to a much more all-encompassing cultural shift surrounding data analytics.

Bypassing the AI-ML gamble, and opting for a more comprehensive adoption of data science, understanding and realising its true value, enterprises will be much better equipped to deal with the fluctuating trends and emerging competition that awaits them.

This has been the view of advanced data analytics specialists Mango Solutions for more than 16 years now. While the company and its chief data scientist and co-founder Rich Pugh urges businesses to embrace methodical and pragmatic data processes before they dive in at AI-ML level, these tools have at least engaged people in the data conversation.

"The notion that ideas like AI or ML can just be plugged in and the company then watches as money pours out of their servers is dangerous, but at least it's finally opened the door

to have the conversation about how companies can become data driven," says Mr Pugh. "Our organisation is focused on facilitating these conversations that we believe should have been occurring 16 years ago when we were inceptioned, so we can help companies avoid quick buzzword-led reactions and instead strive for a cultural transformation based on data."

Much of this facilitation revolves around encouraging businesses to ask the right questions pertinent to their strategic goals. Rather than jumping on buzzword bandwagons for short-term gains, by ingraining a climate of data reliance within a company, a more sustainable resistance to challenges is formed and a more streamlined route to tailored business success is established.

Mr Pugh continues: "Some organisations we work with are just starting out and skilling up, trying to make sense of data science and indeed the tools that contribute to the conversation. Others are more experienced and established, but have more of a challenge in terms of building capabilities."

"The question for all reverts to 'where are you on your data-driven journey

and what's the best way forward for your company?'"

Mango Solutions has put together a nine-step process to assist clients in this regard. It's a practical checklist that helps companies not only learn about data science, but how to enact its benefits long after Mango's team has left the building.

Initiated by building an analytic community, analysing best practices and analytic governance, and educating the business to ensure consistency of analytics across the organisation, businesses are then encouraged to focus on asking the right questions, before prioritising and executing goals, and subsequently identifying success parameters around these goals. Finally, an assessment of



**The rallying cry throughout has been for companies to become data junkies, just like Mango**

overall data-driven maturity is conducted, before alignment across the entire organisation, and especially leadership positions, is confirmed.

"What has never been more prevalent is the conversation at C-suite level and the burden of making the most of data. These nine steps aren't a linear, standardised service, but a methodical, customised, healthily cynical and, above all, pragmatic way to get companies to a point of self-sufficiency," Mr Pugh explains. "The key is to find experience, look for knowledge and skills transfer, and to develop autonomy and support."

"Ultimately, Mango believes this nine-step approach is a proven effective way to facilitate a company's data-driven journey."

Entities to have reaped the benefit of this service over the years range from financial and insurance institutions, to pharmaceutical firms, retail and transport companies, all the way up to government level, reflecting how widespread the data conversation now is and how vital it is for organisations to be saying the right things in this conversation.

Mango Solutions played a pivotal role in creating a blueprint for how data science can be implemented across UK government, instilling a sustainable culture of methodical data analytics, before slowly removing personnel from the operation once the education and transformational shift was complete.

"This is key," says Mr Pugh. "Yes, we deliver value and guide them through their data journey. More than that though, we help them build something sustainable beyond our input; a repeatable process ingrained into the culture of their organisation."

"Most businesses are enthusiastic about improving, but we channel that enthusiasm in a more pragmatic way. Getting them from A to B, from B to C, C to D, avoiding the pitfalls and essentially laying out a map for their data journey."

Since 2002, Mango Solutions has been championing the role that data science can play for organisations and

while other competitors, large management consultancies and startups alike, have looked to provide a similar service, the UK-based provider has remained at the forefront; its longevity now a differentiator in itself.

The rallying cry throughout has been for companies to become data junkies, just like Mango. The sooner all businesses can see data as a cultural shift rather than a short-term technical project, the sooner the goal of a data-driven future can be realised.

Chief executive Matt Aldridge concludes: "Companies might well have been making millions off individual projects, classing that as success in analytics. However, these are anecdotal stories of isolated successes as opposed to examples of an ongoing 'turn-the-wheel' process to generate value. They still weren't seeing the opportunity for data to improve every decision and that made data an unsustainable aspect of their business."

"It is now vitally significant that companies realise every decision can be improved with data and this can only be done by creating an entire transformational and cultural shift in how they see data and how they apply it to their operations."

"Things like AI, big data, ML, they're all enablers, but they're not a business culture or strategy. Separating this hype from actual strategy is what we've been renowned for since inception. We've witnessed the evolution of data and other companies now need to embark on their journeys too."

"By doing so in an open, honest and outcome-focused way, they too will find out that there's massive value to be realised from the future of data."

For more information please visit [www.mango-solutions.com](http://www.mango-solutions.com)







Artur Widak/NurPhoto via Getty Images

PUBLIC SERVICES

# Are we sleepwalking into an AI police state?

Predictive analytics enabling law enforcement to identify “high-risk” areas has highlighted ethical and legal quandaries

Nafeez Ahmed

**S**cience fiction has long speculated on the danger of a dystopian future and machines powered by artificial intelligence (AI). But with the advent of big data, we no longer need to speculate: the future has arrived.

By the end of March, West Midlands Police is due to finish a prototype for the National Data Analytics Solution (NDAS), an AI system designed to predict the risk

of where crime will be committed and by whom. NDAS could eventually be rolled out by every police force in the UK.

Fourteen police forces around the UK have used or planned to use such tools. But a report published in February by human rights group Liberty warns that far from being objective, police crime-mapping software reinforces pre-existing biases about who commits crime.

Current mapping tools use past crime data to identify so-called high-risk areas, leading to more intensive patrolling. Yet these areas are often already subject to disproportionate over-policing. By relying on data from police practices, according to Liberty’s advocacy director Corey Stoughton, these tools might simply “entrench discrimination against black and minority-ethnic people”.

Ben Hayes, a data protection and ethics adviser to the European Union, United Nations and other international organisations, warns that the increased use of such mapping tools is increasingly turning ordinary citizens into suspects.

“People can be categorised as vulnerable, at risk, threatening, deserving or undeserving,” says Dr Hayes, noting that this tends to target those already marginalised. “Services such as border control, policing and social welfare are all subject to inherent bias. Machine-learning doesn’t eliminate those biases, it amplifies and reinforces them.”

Predictive-mapping tools have nevertheless become ubiquitous among national and local governments around the world. Among the most widely used is Mosaic, a geo-demographic segmentation tool that profiles every single person in the UK using 850 million pieces of data.

Created by marketing company Experian, Mosaic is used extensively by councils and political

parties in the UK. It draws on GCSE results, welfare, child tax credits, the ratio of gardens to buildings in different areas and even internet scraping of public websites to build its profiles.

According to Silkie Carlo, director of privacy campaign group Big Brother Watch, Mosaic “perpetuates crude stereotypes, such as the category of ‘Asian heritage’, or ‘dependent greys’, or the use of postcodes to link people living in certain areas to alleged risks of certain behaviours”.

Councils and police often combine such categories with other highly sensitive data, including children’s school records, use of social services and family issues. These can then be used to “predict” whether a person might be at greater risk of committing benefit fraud, engaging in violence or even being sexually abused.

“There are already real risks of discrimination against minorities or poor people due to mistaking correlation for causality,” says Ms Carlo. “When you add into the mix a complex AI tool, this simply covers those biased decisions with a veneer of scientific credibility.”

As cities around the world from London, to Dubai, to Shanghai aim to become “smarter” in using data technologies to improve government services, the scope for abuse is magnified.

The main problem with smart cities is that sensors are on all the time, says Ann Cavoukian, a former privacy commissioner in the Canadian province of Ontario, who now heads up the Privacy by Design Centre at Toronto’s Ryerson University. Not only is there no opportunity to consent to the use of personal data, that data can then be used in ways outside your control.

Citizens might be penalised for jaywalking or not meeting recycling targets, or have their credit rating affected due to social media usage.

Meanwhile, private data companies contracted to help run smart city processes have tremendous access to people’s entire lives, blurring the checks and balances essential to democracies.

Dr Cavoukian’s solution is simple: data should be anonymised.

By scrubbing personal identifiers from the data, it can still be used to improve public services, but avoids the risk of wrongfully persecuting individuals.

For Ms Carlo, the key is transparency. “No one wants to be a Luddite, but local authorities are running before they can walk,” she says. “The starting point before they use such technology is to consult local people in advance, which hasn’t happened. People are being assessed through these systems without their knowledge or consent.”

Lack of transparency on what technology is being used and where is compounded by a complete lack of legal oversight. According to Dr Hayes, this urgently needs to change to avoid creeping into more authoritarian societies.

“Ultimately, we need to be able to choose as a society how we use these technologies and what kind of society we really want to be

“Right now there is a legal vacuum which has given carte blanche to governments and companies,” he says. “This has led to a race to the bottom in standards as we compete with countries like China. Ultimately, we need to be able to choose as a society how we use these technologies and what kind of society we really want to be. Regulation need not be viewed as a barrier to innovation. A robust legal framework can provide the oversight to make technology work for people, rather than against them.” ●



98%

of automated facial recognition matches by the Metropolitan Police during the 2017 Notting Hill Carnival wrongly identified innocent people

Big Brother Watch 2018

64%



of US citizens would approve of facial recognition being used by law enforcement

Morning Consult 2018

28%



of the UK public believe the rise of artificial intelligence will be a positive for the UK economy

YouGov 2018





# Leverage the power of your data with BI

Drive your business efficiently  
with real-time data in live, and  
automated dashboards.