Board Briefing ibe

Corporate Ethics in a Digital Age

By Peter Montagnon

Published by
ibe Institute of Business Ethics

What we do

The Institute of Business Ethics, whose purpose is to promote high standards of business behaviour based on ethical values, is an important partner to any business wanting to preserve its long-term reputation by doing business in the right way.

All organisations need to demonstrate they are trustworthy in order to operate effectively and sustainably. Reputations are not based solely on the delivery of products and services, but on how an organisation values its stakeholders. Having a reputation for acting with honesty and ethics will not only differentiate an organisation, it will make it more successful.

For over 30 years, the IBE has advised organisations on how to strengthen their ethical culture by sharing knowledge and good practice, resulting in relationships with employees and stakeholders that are based on trust.

We achieve this by:

- Acting as a critical friend to organisations we work with
- Advising senior business leaders and those with responsibility for developing and embedding corporate ethics policies
- Supporting the development of these policies through networking events, regular publications, research and benchmarking as well as training
- Providing guidance to staff through bespoke training and decision-making tools
- Educating the next generation of business leaders in schools and universities.

The IBE is a registered charity funded by corporate and individual supporters.

Donate today and be part of a network sharing good practice in business ethics.

www.ibe.org.uk

Advisory Services

Events

Research & Publications

Ethics Training & Tools

Business Ethics in Education

Assurance

Advocacy

ibe
Institute of Business Ethics

**“Doing business ethically ...
.... makes for better business”**

Board Briefing

Corporate Ethics in a Digital Age

By Peter Montagnon

Corporate Ethics in a Digital Age

ISBN 978-1-908534-42-2

First published June 2019
by the Institute of Business Ethics
24 Greencoat Place
London SW1P 1BE

Company No. 11594672
Registered Charity No. 1180741

Contents

Author

Peter Montagnon joined the IBE as an Associate Director in September 2013. Previously he was Senior Investment Adviser at the Financial Reporting Council, which he joined after almost ten years as Director of Investment Affairs of the Association of British Insurers.

For two decades from 1980 Peter was a senior journalist on the Financial Times (FT), including spells as Head of the Lex Column and in charge of coverage of the international capital markets. His last assignment, from 1994 to 2000, was as Asia Editor, responsible for the FT's coverage of a region stretching from Pakistan to New Zealand.

Peter served on the European Commission's Corporate Governance Forum from 2005 – 2011. He is past Chairman of the Board of the International Corporate Governance Network. He is a member of the Board of the Hawkamah Institute for Corporate Governance, Dubai and a former member of the Corporate Governance Advisory Board of the Norges Bank Investment Management.

Acknowledgements

Writing this Board Briefing has put me on a steep learning curve and I could not have completed it without the help of large numbers of people, including the many who gave their time to be interviewed and, later, to comment on the finished text. Many of these helped in a personal and off-the-record capacity, but I would like to record my thanks to them anyway.

By name, I should like to thank Lord Tim Clement-Jones CBE both for reading and commenting on the text and, along with his colleagues at DLA Piper, for generously agreeing to host the launch event. Those who read and commented on the text include Carl Wiper and Simon Reader at the Information Commissioner's Office and, in a personal capacity, Jocelyn Brown, Paul Brown, Fiona Ellard and Dominic Hall.

At the IBE I am grateful to many colleagues for their help and support, including Philippa Foster Back CBE, Director, Dan Johnson and Sophie Hooper Lea, who edited the text in ways that greatly enhanced it. Thanks also to Neil Pafford and Graham Martin, who prepared the text for publication. They are all entitled to share in the credit for the result, and any deficiency is entirely down to me.

IBE Foreword

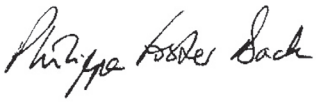
Whether we like it or not, we live in a world that is increasingly dominated by technology. The opportunities this creates to improve all our lives are huge, not only in obvious areas like healthcare and financial services, but in a whole range of services and products. Yet we cannot ignore the risk that the development of machine learning will lead to excessive concentration of power in those that control data. This raises genuine concerns about loss of security and abuse – for example through intrusion into privacy, exploitation of vulnerability and unfair treatment of individuals when systems are biased.

The introduction of Artificial Intelligence (AI) therefore needs to be accompanied by strong and carefully considered ethical principles. This is not a question of restraining or limiting its adoption. As Ginni Rometty of IBM says, companies are judged not just by how they use data, but whether they are trusted stewards of other people's data. Those who consider and respond to the ethical challenges of AI, and are true to their values, are more likely to be trusted. And those who are trusted are more likely to survive and prosper in the long run.

There is thus a clear link between taking an ethical approach and competitiveness. This, in turn, suggests a business or economic model that is quite distinct from that commonly found in large countries that are pushing hard on the technology button, like the US and China.

This Board Briefing does not claim to offer all the answers. Rather, it seeks to set out the issues in a practical way in the hope that this will help boards to engage and cope with what confronts them. Although they originate in the technology, most of these challenges are more values-based and philosophical than technical. An important conclusion is that they belong not in an IT silo, but in the general debate about business judgement and risk appetite.

If some boards have tended to put AI to one side on the basis that it is too technical and difficult, we hope this briefing will persuade them to take a second look.



Philippa Foster Back CBE
Director
Institute of Business Ethics

Executive Summary

Growing reliance on data and the integration of AI into business activity has thrown up some large challenges for governance. Boards not only have to manage a new set of risks and opportunities – they have to do so in a world that is rapidly changing in ways that make it harder for them to exercise control.

This Board Briefing sets out to help boards adapt. It rests on the premise that, while directors have to take account of AI and understand the role it is playing in their business, they do not need to be experts in technology to tackle the relevant questions. Indeed, most of the challenges facing directors are more ethical and philosophical than technical.

What boards do need, however, is a reliable source of advice either from within the company through a strong Chief Information Officer (CIO) or Chief Technology Officer (CTO) and/or through the appointment of independent outside advisers.

> *Boards need to consider the application of technology as integral to their discussions on risk appetite and risk management*

One of the core challenges is the way in which access to data creates information asymmetry. This gives power to those who enjoy such access, which may be used to the detriment of those who do not. Much of the challenge facing boards is around deciding where and how to draw the line between what is appropriate and what is inappropriate.

Thus this Board Briefing examines a series of challenges. It starts with the need to make sure the board remains in charge in a world where power and knowledge can become concentrated in the hands of the IT experts – both at the group support and at the divisional business level – even though, formally, they may be of junior rank. It moves on to the need to share the benefits of technology fairly, establish human accountability, avoid bias in the development and operation of algorithms, and treat customers and employees fairly. The remaining challenges include protecting data entrusted to the company, what to do in the event of a cyberattack and the desirability of incorporating the issues into codes of ethics.

One conclusion is that boards need to consider the application of technology as integral to their discussions on risk appetite and risk management. Ultimately, some of the issues are only tangentially about technology and more concerned with the business model. No system can be perfectly secure and, if it were, a business would find it almost impossible to function.

It is thus a key objective to manage risk in such a way that the business can grow within a sustainable framework while enjoying public trust. In this regard, technology is not a fundamentally different challenge from many of the others facing boards today.

Introduction

The corporate journey into the world of AI is only just beginning. Business leaders perceive enormous change on the horizon, but they are uncertain about where AI will take them and how they will cope with something whose dimensions keep changing just when they think they have begun to understand them.

For boards, there is a temptation to prevaricate. Changes may never happen as currently expected. Indeed, past experience of technological advances shows that a first wave of excitement is followed by uncertainty, as it turns out that change is slower than the enthusiasts expected. Only later does the technology begin to grow and build from a firmer base. The Gartner Hype Cycle in Figure 1 provides a graphic representation of how acceptance of a technology or application will evolve over time.

Thus, the experts say that general AI – by which we mean the development and exploitation of machines that can make qualitative judgements – is still decades away. At present we are in a position where the primary role of AI is to augment human activity, making it more reliable, efficient and productive. Many companies are already using AI in this way and, however uncertain the end destination, the prospect is that the use of AI will grow. Meanwhile, even with the present use of AI there is still a need for human accountability, for example to enable bias to be challenged.

Figure 1 The Gartner Hype Cycle shows the phases of acceptance of a new technology

Source: Gartner Methodologies, Gartner Hype Cycle – www.gartner.com/en/research/methodologies/gartner-hype-cycle

This is, in itself, a major challenge for boards and corporate leadership. The requirement to manage the consequences of AI cannot be ducked and the principal aim of this Board Briefing is to encourage boards to put the issue firmly on their agenda.

Chapter 1 of this Board Briefing explains the current situation. Chapter 2 presents nine challenges around the use of AI, offering what are hopefully some practical thoughts about how they can be addressed. Finally, Chapter 3 looks at the expertise that is required in the boardroom.

Perhaps contrary to intuitive expectations, the skills needed to address these challenges require less of a technical mastery of the inner workings of AI than a philosophical and ethical approach to resolving the issues thrown up.

Of course, boards need to continuously inform themselves about developments in AI, including what it does, what its limits are and where it is heading. For that, they need sound advice from a trusted source, which ought to be a strong CIO or CTO who is backed up, as appropriate, by independent expertise. Yet, armed with this help, most board decisions will be about where to draw the line.

The skills needed to address these challenges require less of a technical mastery of the inner workings of AI than a philosophical and ethical approach to resolving the issues thrown up

How far, for example, should one go in monitoring an individual's behaviour in an attempt to persuade a potential customer to spend more money or to check on an employee's reliability? Do permissions need to be granted repeatedly and, if so, who should keep track of the process? Does it matter if the individuals concerned do not realise they are being monitored, even though this is no longer supposed to happen under the European Union's General Data Protection Regulation (GDPR)? At what stage should a decision made by a machine be overridden because it is clearly wrong, perhaps because the machine itself is confronting new circumstances that its creators had not anticipated? What do we do when we cannot be sure how the AI has reached its conclusions? Who is accountable for errors – the programmer, the manufacturer of the product that incorporated the system or the person who bought that product, perhaps a self-driving car? One important answer at the outset is that there is always a need to keep a human in the loop.

All these questions are very difficult, but they are less about the technology itself than how it is applied. In that sense, the decisions that boards must take fit naturally into their general view of risk appetite, risk management and oversight. The decision that boards have to make about restoring their systems after hacking is related to business risk, given that closure threatens the continuity of the business. It is similar in quality to the decision taken by racecourse owners in February 2019 to resume racing fixtures after an outbreak of equine flu. They had to weigh up the lack of absolute certainty that the episode was over with the certain impact on their business of continuing closure. This is primarily a business judgement, not a technical one.

Both decisions involve addressing uncertainty, because safety can never be 100 percent secured. Yet, if no decision is taken, the future of the business cannot be guaranteed either.

With AI, there is a strong ethical dimension for two particular reasons. First, personal data has acquired an economic value, as the accompanying inset on Facebook reminds us. While, in principle, the value ought to reside with the subject of the data, it is usually others who are best placed to exploit it. Second, access to the data creates an information asymmetry that confers power on those who have it and vulnerability on those who do not.

What happens if Facebook collapses?

In a provocative article in the New York Times, John Herrman discusses the possibility that Facebook's customers might eventually move on, leaving the company to wither. Even at that stage, he concludes, Facebook will still control some valuable commercial assets.

"The advertising data exposed in a user's personal Facebook archive is, of course, just a sliver of what is available to the company. Facebook's real profile of who you are – the one that it uses to fill your needs and show you ads – is far more comprehensive. The company's relentless accumulation of user data isn't just a grab for power or a default behaviour. It's a long-term investment. You may forget Facebook; it could happen sooner than you expect. But it's not likely to forget you." [1]

> *Access to data creates an information asymmetry that confers power on those who have it and vulnerability on those who do not*

It is not surprising, therefore, that most of those who look closely at these issues tend to emphasise the importance of ethics and codes of ethics.

Ethics matter because an ethical approach inspires trust, and trust is needed to build public confidence in organisations that control data with such power over people's lives. This is not a reason for seeking to curtail the adoption of new technology. It is instead an opportunity for adopting it in a way that delivers clear benefits within a trusted framework. Companies and economies that can do this will set themselves apart, as well as finding it easier to comply with data protection requirements. That is where competitive advantage lies, and is indeed the real opportunity.

[1] New York Times (12 December 2018) *What happens when Facebook goes the way of Myspace?*

What are we Dealing With?

The two major developments addressed in this Board Briefing are the rapidly growing use of AI throughout business and the tightening up of rules on data protection through the European Union's GDPR and the UK Data Protection Act 2018. This section briefly describes what is at stake.

Artificial Intelligence

Computers have long been able to perform complex calculations more quickly and efficiently than humans, but at the outset they depended entirely on data input by humans. With the advent of AI, that has changed. Computers still rely on humans for the insertion of raw data, but they can also acquire knowledge by continuing to process the results of their initial calculations and/or by combining one set of data with another or by connecting with other networks.

This enables them to identify choices more efficiently than humans, for example in medical diagnosis. It also enables them to track, understand and influence human behaviour using a wealth of data that they have acquired on individuals. Computers are also able to connect the data in ways that humans might not have been able to work out for themselves. The targeted way in which individuals were nudged to vote in both the UK Brexit referendum and the last US Presidential election is a case in point. We now call this AI, and it can be seen to present enormous opportunities and risks for business and society at large.

The Engineering and Physical Sciences Research Council (EPSRC) defines AI as follows:

> *"Artificial Intelligence technologies aim to reproduce or surpass abilities (in computational systems) that would require 'intelligence' if humans were to perform them. These include: learning and adaptation; sensory understanding and interaction; reasoning and planning; optimisation of procedures and parameters; autonomy; creativity; and extracting knowledge and predictions from large, diverse digital data."* [2]

An important ethical issue arises because, once machines are autonomously able to learn, adapt, reason and optimise procedures, they could eventually fall outside human control and no longer be susceptible to challenge. In exploiting AI, both governments and business need to be clear about where to draw the line. For all sorts of reasons, including conscious or unconscious bias in programming, machines may come up with sub-optimal decisions.

So far, computers can beat world champions at chess or score more highly than anyone at computer games, but they are not yet capable of conceptual or moral thought. Speaking on the BBC 4 documentary *The Joy of AI* in September 2018, Professor Jim Al-Khalili [3] showed that, from a picture, a computer could identify that the subject was a dog and even what sort of dog it was. However, when the pixels in the picture were altered, it turned out that, as far as the computer was concerned, one of the dogs was not a dog at all, but a trombone. Common sense would tell us that this was wrong, but AI is not invested with common sense.

[2] The EPSRC's definition of AI is available on the *Artificial Intelligence Technologies* section of their website www.epsrc.ukri.org

[3] BBC (2018) *The Joy of AI*

> *What if we reach the stage where we can no longer control or manage the decisions made by machines?*

Professor Al-Khalili put it this way:

"Machine learning powers most AI today. They learn from data, the soundness of data and, potentially, the solution to problems. The next step is neural networks, which starts to tackle abstract thought.

"The network is in some sense intelligent, but at the same time there's no understanding of concepts there. It doesn't actually know what a dog is, let alone anything else, which is why it can be fooled by just a few pixels."

So far so good, but what if we reach the stage where we can no longer control or manage the decisions made by machines? It is very easy to say that those responsible must at least be able to explain why the machine made a particular decision. However, with the processes becoming so complex and autonomous, this may already be easier said than done in some cases.

This is how two academics, Bryce Goodman and Seth Flaxman, put it:

> *"Neural networks, especially with the rise of deep learning, pose perhaps the biggest challenge – what hope is there of explaining the weights learned in a multilayer neural net with complex architecture?"* [4]

Conclusions for boards

AI offers enormous opportunities, but it is still limited in its applications and those responsible for it must hold themselves accountable for what it does. Boards need to make sure that their organisations are properly interpreting the conclusions it reaches. They need to understand how AI has affected decision-making and be clear about where the machine's capability to make decisions ends.

Machines themselves are amoral. They essentially rely on rule-based analysis and cannot deal with uncertainty and the unforeseen. That said, they add greatly to the power of those that use them. The tasks to which AI is applied and the nature of its decisions will reflect the values of those who employ it. This suggests that, however difficult it is to organise, it should always be possible to override an AI decision when it is clearly wrong.

> *However difficult it is to organise, it should always be possible to override an AI decision when it is clearly wrong*

[4] Bryce Goodman and Seth Flaxman (AI Magazine, June 2016) *EU regulations on algorithmic decision-making and a "right to explanation"*

GDPR and NIS

GDPR is a European Union (EU) regulation that aims to give citizens and residents control over their personal data in an age where data flows have increased significantly and where it is increasingly recognised that data has an economic value in its own right.

GDPR seeks to improve an individual's right both to privacy and control over their personal data. It also aims to simplify and standardise the regulatory environment for international business. It became effective on 25 May 2018 and is being applied in the UK through the new Data Protection Act.

The UK is implementing GDPR both to align its own rules with those prevailing in the EU and because, whatever its location, any entity that processes the data of EU citizens is subject to GDPR where the processing activities are related to the offering of goods or services or the monitoring of an individual's behaviour. All such entities may face heavy fines of up to four percent of worldwide turnover for breaches of the GDPR legislation.

GDPR applies to both controllers and processors of data. A controller is defined as an entity that determines how and why data is processed and a processor is an entity that does the processing. [5] Individuals are most likely to be aware of it because of requests from organisations asking for permission to continue to stay in touch.

Those involved in holding and processing personal data face a raft of requirements. Controllers must ensure that data is processed according to certain key principles, including lawfulness, fairness and transparency. The data can only be collected for specific and legitimate purposes, it must be processed securely and, once it is no longer required, it must be deleted.

Controllers must have a lawful basis for processing any personal data, and an additional basis for processing special categories of data relating to issues such as revealing racial or ethnic origin, or concerning health. Having the consent of the data subject remains one basis for processing personal data. However the regulations require that it must be affirmative and freely given rather than involving passive acceptance – for example through pre-ticked boxes or opt-outs. Consent is not the only basis on which an organisation can process personal data; for example, other lawful bases for processing include when it is necessary for compliance with a legal obligation or the execution of a contract. [6]

> *Accountability is a key principle under GDPR*

Accountability is a key principle under GDPR, which makes controllers responsible for complying with the regulation and, furthermore, being able to demonstrate their compliance. This involves putting in place appropriate technical and organisational measures, and also adopting data protection by design and default. If a project involving the processing of personal data is likely to result in a high risk to individuals, it is a requirement that a data protection impact assessment is carried out to identify and try to mitigate those risks.

5 See Regulation EU 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, Article 4

6 See the Information Commissioner's Office (ICO) Guidance: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing

In the event of a breach, the data controller is under a legal obligation to notify the supervisory authority within 72 hours, unless it is unlikely to result in a risk to the rights and freedoms of individuals. The latter have to be notified if adverse impact is determined.

Each person has a right of access to the personal data that organisations hold about them, rights of erasure and rectification of that information, as well as a right to data portability. GDPR also gives citizens a right to question significant decisions that have been made solely by automated means and without human involvement.

According to ICSA: The Governance Institute, decision-makers at the highest levels of organisations will need clear, reliable updates from those who are closely involved in the management of data throughout the organisation. [7] Input will be required from multiple functions, including legal, HR, IT and other departments such as customer services and marketing.

Eight individual rights under GDPR

Most of these rights existed previously (with the exception of the right to portability), but have been strengthened. They are the rights for individuals to:

1. Be informed
2. Access their own data (within one month of a request and, ordinarily, free of charge)
3. Rectification of inaccurate or incomplete information
4. Erasure, including when data is held by third parties
5. Restrict processing
6. Data portability
7. Object to processing
8. Not be subject to automated decision making, including profiling.

In addition to GDPR, the EU Directive on Network and Information Systems (NIS) has also been applied in the UK. This is designed to promote collaboration among and between governments on cyber security and ensure a prompt reaction to cyberattacks. It requires governments to be more prepared and to co-operate to share information about risks. It also requires a culture of security across key sectors that are seen as vulnerable. This includes firms in the energy, water, banking, financial market infrastructure, healthcare and digital infrastructure sectors.

Relevant businesses will need to take additional security measures and to notify the relevant national authorities about serious incidents. This gives them a greater reporting obligation than exists on firms that are only subject to GDPR where the reporting requirements focus on leakage of personal data rather than cyberattack more generally.

[7] ICSA: The Governance Institute (2017) *Guidance Note: EU General Data Protection Regulation*

Conclusions for boards

The impact of GDPR will depend not only on the level of compliance, but also on the spirit in which it is approached. The regulation is based on an important principle that individuals have rights over their own data and that these rights should not be abused. Firms that acknowledge these rights will gain reputational and competitive advantage over those that merely seek a minimalist approach to compliance, especially where individual rights are concerned.

Individuals have rights over their own data and these rights should not be abused

However they approach GDPR, boards will nonetheless need to ensure the right disciplines and governance oversight are in place. This includes ensuring that the company is aware of what data it actually controls. Boards will also need to decide:

- Which committees will have responsibility for reviewing the detail and implementation of data protection measures
- How often and in what way this information will be communicated to the board
- What escalation procedures will be in place for non-routine updates
- What criteria will be used to evaluate the effectiveness of data protection measures, perhaps through crisis simulation to identify how improvements will be made.

Nine Challenges

This section outlines nine challenges facing boards and suggests some practical steps that will help to address them.

Challenge 1 – Making sure the board remains in charge

One of the core problems thrown up by AI is information asymmetry. This may exist between a firm and its customers, for example where the firm has used machine learning to tell it things about its customers that even they do not know about themselves. It may exist between a firm and its employees, when monitoring employee behaviour leads the firm to manage differently and often in its own interest, rather than in that of the employees themselves.

> *In today's world of algorithms, it is not always easy for boards to understand or monitor what is going on in the company*

We shall address these points later, but for the time being it is important to recognise that information asymmetry can exist *within* the firm as well. In today's world of algorithms, it is not always easy for boards to understand or monitor what is going on in the company. This means that highly, but narrowly educated data scientists can wield enormous power in the 'engine' room. At the extreme, this changes the hierarchy of governance to the detriment of a board's ability to deliver understanding and strategic judgement at the apex of the organisation.

The example of UBS and derivatives trading during the global financial crisis of 2008 is a case in point. It would have been rash to presume that the board of the Swiss bank – or indeed any board – would have understood the algorithms on which the trading was based. Yet it is right that the directors remain accountable for what happened. A first task for boards is therefore to make sure that they do remain in charge.

A good starting point is to ensure that the firm's values are properly articulated and embedded in the workforce at all levels. Highly skilled employees, who might still be quite junior in rank, need to be encouraged to think about whether their behaviour matches up to the expectations of the company and about the implications of their decisions for the company's stakeholders.

Much has been written about the 2015 emissions scandal at Volkswagen. On the one hand, it is tempting for the corporate leadership to dismiss the incident as the actions of one or more rogue employees and paint the company and the board itself as victims of fraud. Yet there was also a failure of values. It must have appeared to the employees concerned that it was acceptable, or even encouraged, to behave as they did. They were there to deliver results, and no questions were asked as long as they did so.

Four ethical principles to consider in the way algorithms are used

1. **Reliability**. Do we keep to our promises?
2. **Honesty**. Do we deceive and lie to people?
3. **Transparency**. Do we operate in secret and can we explain our decisions?
4. **Respect**. Do we trample over the interests of others to get what we want?

The inference for boards is that values must be embedded and incentives established throughout the company to reward good behaviour and punish bad. As part of the embedding process, boards constantly need to be alert to the way in which the values are being communicated and how adherence to them is being tested. [8]

Beyond that, employees need to be helped to understand that this is not simply an issue of compliance. In the fast-moving world of modern technology, employees need to be constantly aware of the consequences of their actions. This means developing an ability to think 'out of the box'.

One way that businesses in Silicon Valley have approached this problem is by bringing people with philosophy training into the software teams. [9] After some initial resistance, the experiment appears to have worked in so far as it encouraged the engineers to think more broadly about the implications of their work. In this way, the understanding of the firm's values and behaviourial expectations can be better understood. Reverse mentoring – whereby younger staff mentor older, more senior colleagues – can also help break down barriers and introduce a healthier mindset.

> *"In the fast-moving world of modern technology, employees need to be constantly aware of the consequences of their actions*

As to oversight, boards need to equip themselves to ask the right questions of the relevant executives and weigh up the answers they receive. This does not necessarily require specialist knowledge, but it does require directors to think through the issues and press for answers that make sense. This is essentially an extension of their work on risk oversight. The mandate for the Audit Committee and Internal Control may need to be adjusted accordingly. Having robust audit systems in place for AI and the way it is used will make a great difference to the ability of boards to exercise effective oversight.

In understanding AI and in the design of such audit systems, boards may also find it helpful to appoint an expert advisory committee, although problems can arise if the experts are paid more than the directors themselves.

The rest of this chapter explores these issues in more detail.

[8] See IBE (2018) *Culture indicators: understanding corporate behaviour*

[9] Forbes.com (9 March 2018) *Why your board needs a chief philosophy officer*

Conclusions for boards

- It is important that boards remain in charge and do not allow technology teams to take over simply because the directors do not understand the technologies used
- Embedding sound values throughout the company, and monitoring the results of doing so, can help to ensure that specialist employees behave appropriately
- Boards need to seek advice where necessary, ask the right questions and demand answers they can understand. The Audit Committee or the Ethics Committee may be well placed to monitor this area and should be encouraged to develop effective internal audit processes
- Boards need to understand how decisions involving AI are taken and the role of executive oversight. This could be helped by the adoption of a decision-making model or framework.

Challenge 2 – Sharing the benefits

New technology is expected to be disruptive and, for many, constitutes a threat to employment. Dealing with the labour market consequences is, of course, a major task for government, but companies will face a public backlash and the risk of intrusive regulation if they keep the efficiency benefits of AI to themselves.

The UK controversy around zero-hours contracts could even look trivial if whole classes of employee are made redundant. Even if they are not, a fierce backlash could arise if there is a perception that new technologies are adding to inequality by driving down wages for low-skilled workers and piling on bonuses for top management.

One answer is to provide training that will help employees to keep their skills up to date and provide, perhaps in collaboration with government, outplacement services that will help them find new and different employment opportunities where necessary. A key principle here is that companies should share the benefits of the new technology with all stakeholders – customers as well as employees. This may require them to work with universities to ensure that the right skills are being developed, and to think how existing skills can be adapted. For example, insurers could encourage actuaries to develop AI skills.

> *Companies should share the benefits of the new technology with all stakeholders – customers as well as employees*

The idea of shared benefits is not new. The US agrochemical firm Monsanto provides a good example of how not to introduce a new technology. Though controversial, its pioneering work in genetically modified seeds created products that were potentially useful in helping farmers to increase their crop yields. Admittedly, some people will never accept that such products are safe, even if they have been declared so by the relevant authorities. Yet Monsanto made its own case harder by the rigorous way in which it enforced patents in order to maximise financial returns.

Farmers were not allowed to collect and save seeds for planting the following year. Instead, the terms of their contract required them to buy more. Activists claimed that the company's approach contributed to a major problem of farmer suicides in India. [10] The company was subsequently bought by Bayer, which decided to drop the name Monsanto.

Arguably, Monsanto could have helped its case by more careful consideration of its impact on stakeholders. A US study of consumer attitudes showed that people were more relaxed about eating genetically modified soybeans if there was an obvious perceived consumer benefit to doing so. [11]

Similar arguments can apply to AI. Amazon has often been roundly condemned for paying too little tax and for its working conditions, including low wages. Yet it also commands customer loyalty because of the ease of ordering and delivery. In 2018, the company responded to widespread criticism by raising wages for workers in the UK and US. [12] Amazon's board must continue to judge how far customer benefits will outweigh public concern over the way in which it does business. This is especially the case as it diversifies more and more into areas such as groceries, thereby displacing more traditional operators.

One other way in which companies can seek to keep the benefits to themselves is by using their technological advantage to squeeze out competition. This was, at least in part, what Monsanto was seeking to do through its aggressive use of patents. Companies certainly cannot be expected to give away their competitive edge, and innovative use of technology forms part of that. However, there is always a temptation to go too far, as demonstrated by cases brought by the EU against some of the large US technology companies.

> *Boards must be very sensitive in working out for themselves where the boundaries of acceptability lie*

For example, the EU announced it was fining Google €4.34 billion in July 2018 for what it described as illegal practices regarding Android mobile devices to strengthen dominance of Google's internet search engine. Margrethe Vestager, the EU Commissioner responsible for competition policy, said that Google's practices had denied rivals the chance to innovate and compete on merit. They had also denied European consumers the benefits of effective competition in the important mobile sphere. [13]

This is not just a question of legal compliance, but also a question of mindset. Vestager talks in terms of damage to others – competitors and consumers. Companies that inflict damage, whether intentionally or not, will find their franchise weaker in the end. Boards must be very sensitive in working out for themselves where the boundaries of acceptability lie and the need to draw a clear line.

[10] See Business & Human Rights Resource Centre (6 February 2012) *India: Activist Vandana Shiva links Monsanto's genetically modified seeds to farmers' suicides* and Monsanto's response

[11] J. Lynne Brown, Yanchao Ping (Journal of the American Dietetic Association, Vol.103, Issue 2, February 2003) *Consumer perception of risk associated with eating genetically engineered soybeans in the presence of a perceived consumer benefit*

[12] The Telegraph (2 October 2018) *Amazon raises wages for UK and US workers following widespread criticism*

[13] European Commission (18 July 2018) *Press Release: Antitrust: Commission fines Google €4.34 billion for illegal practices regarding Android mobile devices to strengthen dominance of Google's search engine*

Questions for boards

- How far are we sharing the efficiency gains of AI with our stakeholders, especially our customers?
- What are we doing to help our workforce adjust, either by upgrading their skills and rewarding them appropriately or by helping them to rebuild their skills and find new employment if we have to let them go?
- Are we using the new technology to gain unfair advantage through predatory pricing or other strategies that damage consumers and competitors?

Sage Group plc – a principled approach

Sage is a market leader in providing a comprehensive range of services to businesses of all sizes, ranging through accounting, financials, enterprise management, people and payroll, payments and banking. AI is part of its world and it operates under five core principles:

1. AI should reflect the diversity of the users it serves
2. AI must be held to account – and so must users
3. Reward AI for 'showing its workings'
4. AI should level the playing field
5. AI will replace, but it must also create.

Source: Sage Group (2017) *Optimism and Ethics: an AI reality check*

IBM – working for trust and transparency

"Every company that develops or uses AI or hosts or processes data, must do so responsibly and transparently. Companies are being judged not just by how we use data, but by whether we are trusted stewards of other people's data. Society will decide which companies it trusts."

Ginni Rometty, IBM Chair, President and CEO

At the World Economic Forum in Davos in 2017, Rometty presented three guiding principles for the use of AI: [14]

1. **The purpose of AI is to augment human intelligence:** AI should augment, not replace, human intelligence and the benefits of the AI era should touch the many, not just the elite few. Global workforces should have the skills needed to work in partnership with these technologies.
2. **Data and insights belong to their creator:** Clients are not required to relinquish rights to their data in order to benefit from IBM services. IBM is fully committed to protecting the privacy of client data and has not provided client data to any government agency.
3. **AI systems must be transparent and explainable:** Technology companies must be clear about who trains their AI systems, what data was used in that training and, most importantly, what went into their algorithm's recommendations. While bias can never be fully eliminated, we and all companies advancing AI have an obligation to address it proactively.

[14] See *IBM's Principles for Trust and Transparency at www.ibm.com*

Challenge 3 – Ensuring accountability

Most people would agree that someone at the human level must be responsible for decisions made by machines, but the nature of AI makes this principle of keeping a human in the loop hard to deliver. Corporate boards are at the epicentre of the discussion on accountability and, for them, the issue adds a new dimension to their consideration of risk and risk appetite.

Financial regulators generally take a robust view. They say they do not want firms to allow machines to make decisions that cannot subsequently be explained. However, this is easier said than done when technological advances now allow machines to learn for themselves and develop processes that are beyond the capability of humans.

So, should boards simply refuse to adopt technologies that make decisions that cannot be easily explained? There is no absolute answer to this question, which also needs to be considered in the context of risk appetite. In practice, there is a trade-off between *explainability* and *accuracy*.

Boards need to draw the line between the desire to deliver good outcomes for most people and their willingness to accept the need to explain how and why poor decisions are made for the few. They will find the task easier if, at the outset, they insist on explainability being written into algorithm design. This is not as difficult as some developers suggest.

> *Should boards simply refuse to adopt technologies that make decisions that cannot be easily explained?*

Risk appetite

One of the most difficult issues is around risk appetite. AI can bring large net benefits, both commercial and social, even though there is an accompanying risk of a spectacular accident. Even at this embryonic stage, for example, driverless cars may be safer with fewer road deaths than conventional vehicles. However, the public will very quickly forget this when crashes are blamed on technology failure. See case study on self-driving cars on page 22.

The power of machines to diagnose medical conditions accurately is a great human benefit. The more sophisticated the algorithm, the more accurate (in theory) the diagnosis will become. However, AI decisions will also become harder and harder to explain, and, in the case of bad ones, to defend. Machines, generally, have no inbuilt sense check. We are still some way from the stage where they are capable of conceptual thought and moral judgements. Human judgement is needed to draw the line, even if it is less ‘perfect’ than a machine-made decision.

In taking on new technology, boards therefore have to have conscious regard to both their risk appetite and their risk management. They must ask themselves how much risk they are prepared to take, be aware of the risks they are taking and, as with all types of risk, develop mitigation strategies.

The buck does stop with directors, even when work is outsourced. There may be legal arguments about liability, for example about who is responsible for a software design fault, but the reputational damage from declining to accept responsibility for damaging decisions would be enormous.

There are various ways in which boards can mitigate risk, even when dealing with very complex algorithms. It may mean:

- Testing the decision-making capacity of the algorithm by feeding in different data, comparing the results to see whether they are consistent and reliable, and then tweaking the algorithm to correct any bias
- Keeping a data provenance record or audit trail of the data that was used in the model creation, together with ad hoc and periodic reviews of AI deployments and decisions [15]
- Deciding not to accept new technology if they or the company's executives cannot vouch for its reliability, creating risks that outweigh the benefits
- Ensuring that there is capacity to challenge or override machine-made decisions when they are clearly wrong in terms of common sense.

> *"It is not acceptable to deploy any artificial intelligence system which could have a substantial impact on an individual's life, unless it can generate a full and satisfactory explanation for the decisions it will take."*
>
> Source: House of Lords Select Committee on Artificial Intelligence (2018) *AI in the UK: ready, willing and able?*

Conclusions for boards

- Ultimately, there must always be human accountability. In the corporate world this resides with boards, whose members are responsible for all that happens. They cannot hide behind an inability to understand or a supplier's failure
- Risk appetite is as important as risk management. How much risk are boards willing to take on? And what steps have been taken to mitigate these risks?
- Even complex algorithms can be subject to risk oversight, but there will be occasions where boards need to draw the line. At the very least, boards need to ensure that patently wrong decisions can be reversed
- It may help to have a governance framework for AI in place to include checks and controls.

[15] See Personal Data Protection Commission Singapore (2018) *Discussion Paper on Artificial Intelligence (AI) and Personal Data – fostering responsible development and adoption of AI*

Case study

When a self-driving car crashes…

In December 2018, traffic police in Redwood City near San Francisco faced an unusual problem. A Tesla Model 30 was cruising down the freeway at 70mph with its human driver asleep. The car did not respond to the police sirens and flashing lights, so the policemen wracked their brains to find a way of stopping it. Eventually another police car got in front of the Tesla and, as it slowed down to an eventual stop, the Tesla's sensors told the self-driving car to do the same. But it still took an old-fashioned knock on the window to persuade the 'driver' to wake up. [16]

This is just one example of how automated and, in future, fully autonomous cars present one of the biggest legal, ethical and insurance challenges in the AI sphere. It is not always clear how these types of vehicles would react in an emergency. This might be because of unfamiliar circumstances, for example if snow or fog has obstructed a car's ability to scan the road ahead. Or it might be because a car faces a split-second choice between two dire alternatives, say driving into a river or running over a child.

> *Automated and, in future, fully autonomous cars present one of the biggest legal, ethical and insurance challenges in the AI sphere*

With time, it should be possible to design cars that are capable of addressing pretty much every eventuality. Governments will also come to a conclusion on who is legally responsible for what happens, and the insurance industry will sort out the liability. Yet, from a corporate point of view, there will always be questions of both actual and perceived accountability. Does the blame lie with the owner, the manufacturer, the software programmer or a sub-contractor?

As with other areas of business, manufacturers who wish to preserve their public reputation have a responsibility that goes beyond mere compliance. Even if the overall safety record of automated vehicles were better than conventional cars, public confidence could be severely shaken by accidents in which the technology is seen to be at fault.

Take, for example, the way in which a new automated car is sold to the public. Drivers need to know the limitations of a system. For example: can it steer around a parked vehicle? Does the automation involve different features that might cancel each other out if used at the same time? For example, the car might try to accelerate and slow down if it encounters an obstacle while overtaking. According to the Law Commission, it is not necessarily sufficient simply to explain all this in the user manual. There may need to be proactive warnings and even user training. [17]

continues >

[16] The Times (5 December 2018) *Police put the brakes on Tesla self-driving car*

[17] Law Commission and Scottish Law Commission (2018) *Automated Vehicles: a joint preliminary consultation paper*

Case study *continued*

Thatcham Research, the UK motor insurer funded automotive research centre, and the Association of British Insurers (ABI) fear that drivers may not be sufficiently sophisticated to understand the limitations of their machines. They think that more needs to be done to regulate how automated functions are described to consumers. In particular, it is unacceptable to use words that suggest a higher level of automation than actually offered. [18]

The Law Commission publication cited above further points out that research into the effect of automation within the airline industry shows that pilots may over-rely on the autopilot system and lose skills in the process. People often find it more difficult to passively monitor a task rather than actually engage in it. They may not know what they are looking out for, or they may be over-confident after using the vehicle for a long time without incident. Also, people tend to ignore warnings that are simply seen as legal disclaimers. There is a risk, therefore, that drivers may fail to override the automated system at precisely the moment when they need to do so.

Confronted by the risk of an accident, a human driver might decide that the appropriate course of action is to do something that is normally considered illegal. For example, he or she may drive the car on to the pavement to avoid a head-on collision with an oncoming emergency vehicle or may accelerate beyond the speed limit to get out of trouble. It is currently difficult to see automated vehicles having the discretion to do something like this.

Manufacturers might feel that they will fulfil their duty if they programme their cars to obey the law at all times. Yet if they take such a narrow view, they may find themselves blamed for accidents where the public believes the technology is at fault.

Conclusions for boards

- Manufacturers need to be clear about the extent to which they are accountable for what happens when a vehicle is operating under automated systems that they have installed. Reputational issues mean that this goes beyond legal compliance
- Marketing literature should not exaggerate the extent of the automation. Indeed, customers should be clearly warned about its limitations
- It is not a sufficient defence to programme cars slavishly to obey all traffic regulations. Sometimes drivers need to use discretion to avoid an accident.

[18] ABI and Thatcham Research (2017) *Regulating Automated Driving – a UK insurer view*

Challenge 4 – Avoiding bias

When people make decisions, their choices are consciously or subconsciously affected by their particular view of the world. This is not as obviously apparent as prejudice, but simply that people are influenced by all the emotional baggage that goes with their upbringing, gender identity and so on. One might assume that machines would be ruthlessly objective and free of bias, but this is not the case. As the AI Now 2017 Report put it:

> *"AI does not exist in a vacuum. We must also ask how broader phenomena like widening inequality, an intensification of concentrated geopolitical power and populist political movements will shape and be shaped by the development and application of AI technologies."* [19]

Inherent bias is thus an important issue for users of AI. It can creep in for a number of reasons, often to do with the nature of the data and the way it is collected. The result can be damaging, for example:

- Ethnic origin or postcode may become a factor in machine-made decisions with regard to predicting credit status or mortgage risk, leading to grotesquely unfair and discriminatory decisions
- West Midlands Police are leading a predictive policing project that uses AI to sift nearly 1,400 indicators to identify individuals who are likely to commit violent crime. The intention is not pre-emptively to arrest but to provide counselling through social services in an effort to prevent crime. [20] However, the project has raised a number of ethical concerns around whether the predictions will be accurate. These are echoed elsewhere in predictive policing initiatives. The American Civil Liberties Union and the Brennan Center for Justice have raised concerns about the risk of a feedback loop whereby the use of arrest data will reinforce expectations about 'bad' neighbourhoods [21]
- Amazon supposedly abandoned development of an AI recruitment programme that had been found to be biased against female jobseekers. The company said it had never used the system to evaluate candidates. [22]

The risk is both that unconscious bias in those who develop algorithms will become embedded in the machine-learning process and that the bias will then become self-reinforcing, as a result of the decisions made by the programmes themselves. Companies that use AI have a responsibility to ensure that they are aware of the risk of bias and take steps to mitigate it. These might include:

- Making sure that software engineers are trained and incentivised to avoid the risk
- Promoting diversity in the software team
- Monitoring outcomes to check for bias and adjusting programmes when it is found [23]
- Providing for redress when machines make biased decisions against individuals, a requirement that would certainly need highlighting in the risk register.

[19] AI Now (2017) *AI Now 2017 Report*

[20] New Scientist (26 November 2018) *UK Police wants AI to stop violent crime before it happens*

[21] Smithsonian.com (5 March 2018) *Artificial Intelligence is now used to predict crime. But is it biased?*

[22] Fortune Magazine (10 October 2018) *Amazon reportedly killed an AI recruitment system because it couldn't stop the tool from discriminating against women*

[23] *IBM Principles for Trust and Transparency* (see Challenge 2 above) state: *"We continually test our systems to find new data sets to better align their output with human values and expectations"*

Challenge 5 – Treating customers fairly

The starting point for any discussion about AI and customers is recognition that, correctly used, AI can add greatly to customer experience and outcomes. In that sense, it should be seen as a significant business opportunity. However, the risk remains that companies trip over into a world where they are using AI to extract value from their customers rather than delivering value to them. This would ultimately lead to loss of trust and damage to the franchise.

To ward against this, it is useful to remember that one of the core principles applying to AI, which is picked up in GDPR, is that individuals have rights over data processed about them and that organisations processing the data have obligations. This is something that many companies fail to recognise.

This principle therefore needs to be built in to the way data is used, even though, in practical terms, individuals may find it difficult to exercise their GDPR rights. Blockchain technology [24] may eventually give them a means of doing so because of its ability to record individual transactions securely across different computers. However, for the time being, it may be too complicated. Individuals may not be aware of how their data is being used or they may simply not be interested. [25] Also, given the way information is traded and amalgamated through different controllers and processors, it may be hard for individuals to determine how it was originally obtained and by whom.

The situation is complicated by the way different strands of information may be combined to enhance predictive quality. Therefore, the context in which information is used matters. Austrian research institute Cracked Labs lists a number of datasets, almost all in the public domain, that can be brought together and used by companies to predict customer behaviour. [26]

As reported in the Financial Times newspaper, when Belgian privacy campaigner Paul-Olivier Dehaye requested his data from advertising technology company Amobee, he discovered that the company had used weather conditions to predict that he was *"likely to suffer from overactive bladder"* on a particular day in June 2018. [27]

The gathering and bundling of data is now commonplace. *"A network of major online platforms, publishers, app providers, data brokers and advertising networks is now able to recognise, profile and judge people at nearly every moment of their lives,"* write Wolfie Christl and Sarah Spiekermann in a report published by Facultas of Vienna in 2016. [28]

[24] According to SAP, blockchain technology is *"a reliable, difficult-to-hack record of transactions – and of who owns what. Blockchain is based on distributed ledger technology which securely records information across a peer-to-peer network"*

[25] According to the Pew Research Center and Berkman (2013) *Teens, Social Media and Privacy,* 60 percent of teenagers report that they are either *"not too concerned"* or *"not at all concerned"* that some of the information they share on social networking sites might be accessed by third parties like advertisers or businesses without their knowledge

[26] See Cracked Labs (2017) *Corporate Surveillance in Everyday Life: how companies collect, combine, analyze, trade, and use personal data on billions*. Datasets mentioned include age, gender, education, employment, relationship status, number of children, purchases, loans, income, new credit granted, religion, health indicators, alcohol and tobacco habits, gambling record, size of home and socio-economic status

[27] Financial Times (8 January 2019) *Data brokers: regulators try to rein in the 'privacy deathstars'*

[28] Facultas (2016) *Networks of Control: a report on corporate surveillance, digital tracking, big data and privacy*

> *The risks begin when data is used for purposes other than the ones stated at the time of its collection*

The risks begin when data is used for purposes other than the ones stated at the time of its collection. For example, data collected in the context of online fraud prevention, credit scoring or payment processing might then be used for customer relationship management, online targeting and other marketing purposes.

Customers of social media sites may not be aware of how their data is being accessed and used by data brokers. This could be at least partly because the site hides, disguises or omits some of the privacy choices available to them. Companies may use misleading language to describe how data is treated. For example, by indicating that data will be anonymised or de-identified when, in fact, they are using disguised identifiers to track, match, profile and target individuals. [29] It is not yet clear that GDPR will be effective in addressing this issue.

The process of blending and aggregating data can lead to a number of damaging conclusions, especially since some of the incorporated data may be out of date. Another potential problem is that a prediction of financial distress may become a self-fulfilling prophesy, if the individual concerned is denied credit as a consequence. Or companies may use data to understand the minimum they have to do to secure customer loyalty, or to single out lucrative customers to the detriment of others. For example, AI can be used to determine the price of air tickets based on general supply and demand. However, it should not be used to determine what a particular customer is willing or able to pay, without making it clear what other customers are paying for the same thing.

Meanwhile, some targeting of advertising remains socially acceptable. Brewers traditionally like to advertise their wares around sporting events, for example, even though the public are becoming more sceptical around targeting soft drink and confectionery advertisements at children. The use of data analytics, however, has exposed new limits. In 2017, The Australian newspaper revealed how Facebook had mined user data to reveal teenagers' emotional state to advertisers, specifically targeting depressed teens. [30] The now defunct Cambridge Analytica has been reported to have had individual profiles on 220 million adult Americans. [31]

For corporate leaders, the overall message is clear. They need to be transparent in the way they use data in connection with customers, take a clear view of what is acceptable and ensure that the line they draw is not crossed. The Monetary Authority of Singapore (MAS) has developed a set of principles that include fairness, accuracy and avoidance of bias; alignment with the firm's ethical standards; accountability and transparency. [32]

[29] ibid

[30] The Australian (1 May 2017) Facebook targets insecure young people to sell ads

[31] Motherboard (28 January 2017) *The data that turned the world upside down*

[32] MAS (2019) *Principles to Promote Fairness, Ethics, Accountability and Transparency (FEAT) in the Use of Artificial Intelligence and Data Analytics in Singapore's Financial Sector*

Questions for boards

- Do we use AI to add value to our customers or to extract value from them?
- Are we open with our customers about the way their data is being used? Do we give them clear and transparent choices with regard to the protection of their privacy?
- How do we use data brokers? Is the context in which the data is used fair to customers? Are we compliant with rules around the development of personal profiles and do we have clear internal guidelines on this?
- Are we using data in a way that exploits the vulnerabilities of our customers?
- Do our customers have effective right of redress if our use of data means they have been treated unfairly?

Case study

Machine learning and insurance risk

AI is both a boon and a torment for the insurance industry. On the one hand, it can be used to predict risk more accurately and thereby reduce it. Thus, for example, the use of 'black boxes' in cars can monitor the way people drive. As this technology makes drivers safer, insurance premiums can be reduced, claims are likely to fall and the roads should become generally less hazardous. [33] Similarly, as it becomes easier to predict the durability of cancer remission, the health aspect of travel insurance should be easier for sufferers to obtain. Another healthcare example is that predictive technology should make it easier to offer life insurance to diabetes patients who are disciplined in the way that they handle the condition. In the process, the industry can also help patients to manage their condition better.

> *AI is both a boon and a torment for the insurance industry*

On the other hand, the greater certainty also creates some big problems which are, ultimately, a threat to the insurance business model. People who know they are low risk may opt out of insurance altogether. This means that the pool of those insured will be slewed towards high risk clients and the most risky individuals may become uninsurable. As the industry collects more data about individuals, there is a danger of insurers becoming more intrusive and prescriptive about how they expect customers to behave. Taken to extremes, this could be infringement of individual freedoms.

While the industry and its regulators grapple with the long-term implications of this, it is clear that the genie is out of the bottle. The availability of data that will help insurance companies refine their understanding of risk cannot be ignored.

continues >

[33] Black-boxes are not, however, necessarily foolproof. In his 16 January 2019 Turing Lecture on *Information Manipulation*, Craig Silverman of BuzzFeed News displayed a swaying cradle for a mobile phone that is designed to fool health insurers' pedometer apps into thinking that their customer is completing the requisite number of steps per day

Case study *continued*

Nowadays insurers can use postcode data to set health premiums. They could even analyse social media postings, which could tell them a lot about a customer's lifestyle choices, such as drinking habits. This information could have an impact on the customer's health or driving insurance premiums. [34]

The problem is further complicated by the development of data brokers, who will prepare and sell profiles using social media and other public data. Insurers may feel under competitive pressure to use such profiles, even though they cannot always be sure how the profiles have been compiled or how reliable they are.

Swiss Re is one company that has responded to these challenges by the introduction of what it calls a *"comprehensive, global data protection compliance framework".*

Compliance at Swiss Re

Swiss Re's *Code of Conduct* says: *"We handle personal data with the greatest care and use it only for legitimate and specified business purposes."*

- Principles include respect for privacy and protection of personal data. Also, *"we obtain personal data fairly"*
- Swiss Re says that it uses a variety of information, including online, health and financial information
- In most cases, the information comes from third parties such as corporate clients. It may be shared with service providers and agents, professional advisers and the client who provided the data
- Uses include underwriting, the management of claims and *"enhancing our knowledge of risk and insurance markets in general".*

Source: Swiss Re – www.swissre.com/about-us/data-protection-brochure

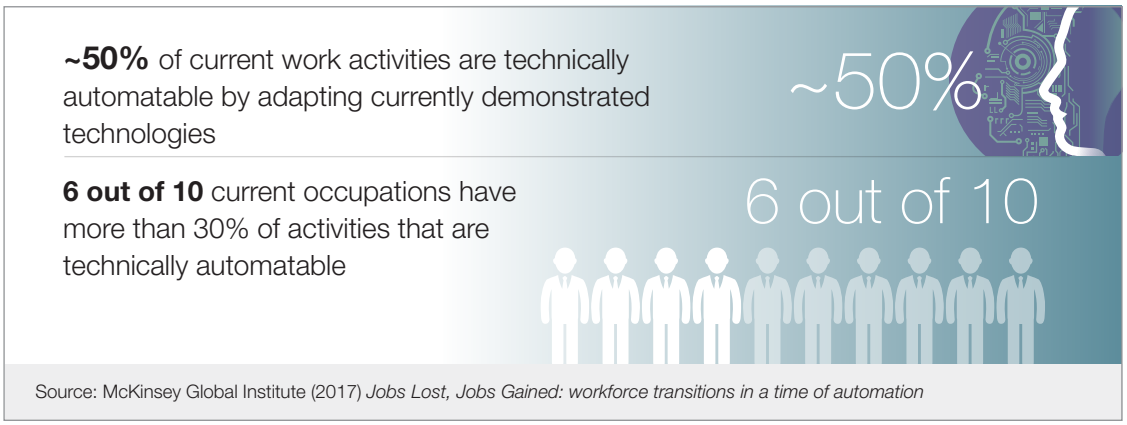
Other executives talk about the importance of codes of ethics, data accuracy and the need to authenticate the data they use, particularly at the individual level. But as the AI revolution takes hold, there is an increasing need for a new understanding between insurers, government and regulators. They need to decide on the role that AI should play and its implications for the pooling of risk, which has been one of the fundamental tenets of the insurance industry.

[34] Insurers already use social media to investigate cases of suspected fraud. For example, see Which? Magazine (5 October 2018) *Could your Facebook posts affect your car insurance quotes?*

Challenge 6 – Treating employees and contractors fairly

The introduction of AI seems likely to have a profound effect on the nature of work and the job market. The McKinsey Global Institute, for example, has forecast that automation could eliminate between 400 and 800 million jobs around the world by 2030 and that up to 375 million workers may need to switch job categories entirely. [35]

Figure 2 The impact of AI on jobs



Source: McKinsey Global Institute (2017) *Jobs Lost, Jobs Gained: workforce transitions in a time of automation*

This raises profound social and political questions that go well beyond the normal scope of corporate leadership. [36] Admittedly, the forecasts have yet to be proved right. McKinsey itself notes that, in some industries, the jobs lost will be replaced by new ones as workers perform new tasks. At the same time, job losses will be lower in areas that require social interaction or that are difficult to replicate, like gardening. The House of Lords meanwhile notes that many AI systems currently aim to supplement, rather than fully replace, human labour, thereby making the workforce more productive. [37]

Nonetheless the context is delicate. Public nervousness of social change, as well as the sense of alienation increasingly expressed by general populations towards the elite, means the introduction of AI needs to be handled sensitively or it may not be generally accepted. Already, some leaders in the financial services industry report that fear of technology-related job losses is adding to stress at work, which could impair the performance of individuals. [38] Well-being at work is rising up the agenda for managers and corporate leaders, and the impact of automation is part of this.

One important conclusion is that AI systems cannot simply be imposed from on high. Companies need to be sensitive to the impact on their workforces and be willing to mitigate this where appropriate through redeployment and retraining. The new *UK Corporate Governance Code* calls on boards to strengthen engagement with their workforce. [39] Clearly, the introduction of AI, and how it is operated, should be a subject for such engagement.

[35] McKinsey Global Institute (2017) *Jobs Lost, Jobs Gained: workforce transitions in a time of automation*

[36] For example, Darrell M West of the Brookings Institution has suggested that some Western democracies might resort to authoritarianism in order to keep restive populations in check. See Brookings (18 April 2018) *Will robots and AI take your job? The economic and political consequences of automation*

[37] House of Lords Select Committee on Artificial Intelligence (2018) *AI in the UK: ready, willing and able?*

[38] Source: various IBE conversations with senior leaders in banking

[39] Financial Reporting Council (2018) *UK Corporate Governance Code*

The AI Now Institute of New York University says in its 2018 report that:

> *"Technology companies need to protect workers' ability to organise, whistleblow, and make ethical choices about what projects they work on. This should include clear policies accommodating and protecting conscientious objectors, ensuring workers the right to know what they are working on, and the ability to abstain from such work without retaliation or retribution."* [40]

AI and related algorithmic systems may already be changing the balance of workplace power

Admittedly, the AI Now Institute's very demanding set of principles is designed for specialised US technology companies. However, there is some read-across to the rest of the corporate world. Employees working in AI should be aware of, and comfortable with, the purpose of what they are doing. All employees should be able to raise concerns, especially if systems are being used to introduce unfair working practices.

This is a question that needs addressing now, not at some time in the future. In its previous report published at the end of 2017, the AI Now Institute argued that AI and related algorithmic systems are already changing the balance of workplace power. Machine learning techniques are quickly being integrated into management and hiring decisions. New systems, the report stated, make promises of flexibility and efficiency. However, they also intensify the surveillance of workers, who often do not know when they are being tracked or evaluated, or why they have been hired or fired.

The institute cites a study by Luke Stark and Alex Rosenblat, arguing that Uber's drivers are put at a disadvantage by the platform through which their assignments are awarded. This, it claims, can force drivers to accept short, unprofitable fares. [41] For Uber, this is part of their effort to provide near-instantaneous service to all prospective riders. As the study states:

> *"Because Uber designs the platforms and can change it at will, conflicts of interest between worker and platform owner are systematically settled in favour of Uber via the platform itself, not collective bargaining or other processes that allow for worker participation."*

The paper points out that such asymmetries are not new, but AI is different because it normalises workplace surveillance: *"As AI-driven management becomes more common, so will the data collection and worker surveillance practices on which it relies."*

Another feature of the platform, the paper continued, is that it enables Uber to nudge drivers into staying on the road at times when it might otherwise be short of drivers. It does this by working out their earnings targets and reminding them that they are close to reaching the target, when they might otherwise be inclined to give up because the pace of business has slowed. This practice was exposed by the New York Times, but the paper surmises that there may be similar practices that workers and the public may never know about. [42]

[40] AI Now Institute (2018) *AI Now Report 2018*

[41] Alex Rosenblat and Luke Stark (International Journal of Communication, Volume 10, 2016) *Algorithmic Labor and Information Asymmetries: a case study of Uber's drivers*

[42] New York Times (2 April 2017) *How Uber uses psychological tricks to push its drivers' buttons*

Another firm, Veriato, helps companies to capture information from employee computers and other devices. This allows it to infer whether individual employees are 'a productivity risk', because of the way in which they use their devices.

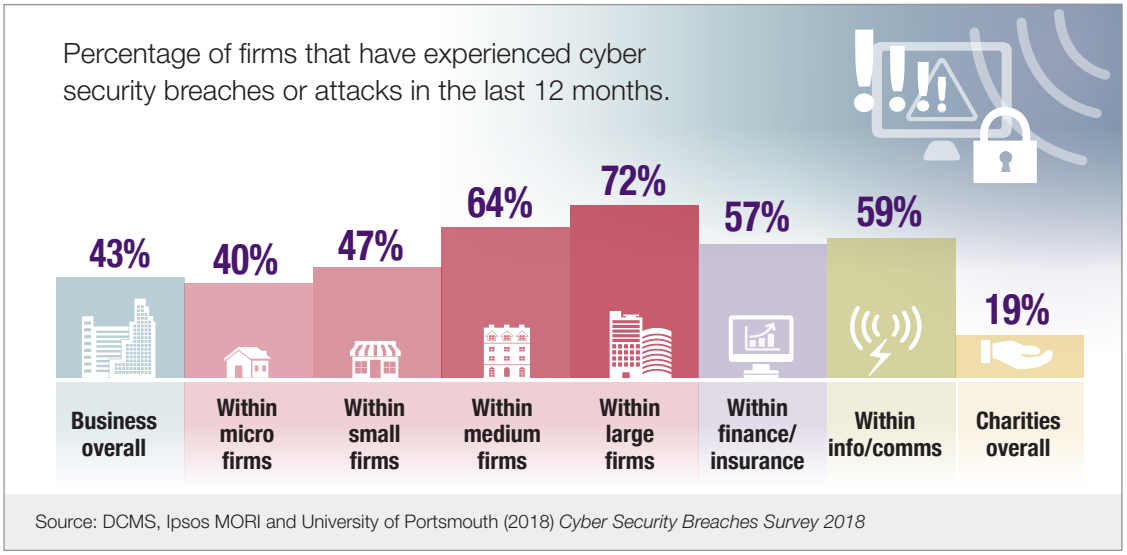
Conclusions for boards

- Boards need to be aware of the sensitivities of the workforce around the introduction of AI. This is consistent with the new UK Corporate Governance Code
- Senior corporate leaders need to maintain dialogue with governments and other authorities to help promote a framework for dealing with the broader social consequences of AI
- Boards need to be aware that the asymmetry of information between management and workforce can create conflicts of interest that are hidden from workers. Companies thus need to be careful about the use of AI as a tool of employee surveillance or being used to induce employees or contractors to adopt working practices they would otherwise reject.

Challenge 7 – Keeping data secure

Cyber security breaches have become commonplace. According to the UK Government's *Cyber Security Breaches Survey 2018*, more than one in four businesses were hacked over the previous 12 months with nearly three quarters of large firms affected. Breaches were more often identified among organisations that hold personal data, where staff use personal devices for work or that use cloud computing.[43]

Figure 3 Cyber security breaches by size and sector



Source: DCMS, Ipsos MORI and University of Portsmouth (2018) *Cyber Security Breaches Survey 2018*

Even though these breaches are increasingly likely to attract the attention of the data protection authorities, most organisations are not well prepared to defend themselves. Many will already have been hacked, but sometimes they will not even know that this has happened. Nowadays hackers tend to sit and watch for some time before they start to take money, and they sometimes target several organisations to see which are worth stealing from.

[43] Department for Digital, Culture, Media & Sport (DCMS), Ipsos MORI and University of Portsmouth (2018) *Cyber Security Breaches Survey 2018*

> *Most organisations are not well prepared to defend themselves against cyber security breaches*

It is better for companies to react before, rather than after, they find out that they have been breached. Having to admit to a cyberattack can affect an organisation's reputation and franchise. According to professional advisers, boards are now challenging more on cyber security, but they are not necessarily getting enough information. The breaches survey found that senior management teams are not always sufficiently engaged. One in five businesses and two in five charities never update their senior managers on cyber security issues, although this result was an improvement on the 2017 survey.

An essential starting point is that companies should know what data they hold and where it is. Sensitive data may be held in a number of places, not just in customer records or HR. For example, internal email trails created in the normal course of work may turn out to be sensitive. If companies are not fully on top of where the data is, there is a risk of initially under-reporting breaches, which compounds the impact on trust.

One thing to watch for is whether data is encrypted when 'at rest' as well as 'in transit'. If it is, it will be more secure. Algorithms can also be used to provide warning signs – for example, they can tell whether people are typing faster than normal or making a lot of typos. This could be taken to indicate that an individual is stressed and there is therefore a heightened risk of security lapses. This approach, however, is an intrusion into an employee's privacy.

Two watchwords in relation to cyber security are resilience (the ability to get up and running again quickly) and forensic (knowing what has been lost). Boards need to understand the risk, ensure they have a full and up-to-date inventory of data in their control, make controls relevant and demand relevant metrics.

Five essential controls

1. Apply software updates when available
2. Keep malware protection up to date
3. Maintain firewalls with appropriate configurations
4. Restrict IT admin and access rights to specific users
5. Install security controls on company-owned devices

Source: National Cyber Security Centre *Cyber Essentials* scheme - www.cyberessentials.ncsc.gov.uk

The main issues are less about technology than about culture. One of most common ways in which malware can be introduced is when employees, even senior ones, fail to follow appropriate procedures. In its *2016 Cyber Security Intelligence Index*, IBM found that 60 percent of all attacks came from within. [44]

[44] Cited in Harvard Business Review (19 September 2016) *The biggest cybersecurity threats are inside your company*

Of these, three quarters were malicious and the remainder inadvertent – such as leaked passwords, confidential material sent to personal devices, misaddressed emails and unauthorised downloads.

The damage can be large and companies need to protect themselves from within as well as from outside.

Questions for boards

- Do we know what data we hold?
- What are our policies with regard to encryption?
- What safeguards do we have in place? Are they covered by internal audit and how effectively are they monitored?
- How well is access to data segregated within the organisation?
- Do senior management receive regular reports on data security and are lapses reported to the board?
- Do all staff, including senior staff, receive training in data security?

Case study

Inside jobs at Morrisons and The Body Shop

Two cases from the retail trade show how cyber security is not just about defending companies from outside attack. Companies can be vulnerable because of actions by their own employees, and these actions themselves can make the company more susceptible to outside attack.

The Body Shop

In July 2008, the Financial Services Authority (FSA) – which was the UK financial regulator at that time – imposed a fine of £85,000 on John Shevlin, who had previously been employed as an IT support engineer at The Body Shop, the beauty and cosmetics products company.

The FSA found Shevlin guilty of insider trading around a disappointing Christmas trading update in 2006. The penalty notice described how Shevlin had acquired the passwords of key executives in the run-up to the update and was therefore able to access their emails. [45] From this, he learned that the announcement was likely to disappoint the market and took out a market position that would generate a profit for him if the share price fell.

According to the FSA, Shevlin was able to acquire this information because his possession of the passwords of certain senior executives gave him *"full access with their log-in identities to their individual email accounts"*. He was able to log in to any physical desktop and/or laptop and create a user profile that would enable him to access emails. Had the company's IT security been stronger, he would not have been able to do this.

continues >

[45] FSA (1 July 2008) *Final Notice* concerning Mr John Shevlin

Case study *continued*

Morrisons

The Morrisons case has had more damaging consequences for the company. It concerned disgruntled employee Andrew Skelton, who in 2014 leaked the payroll data of 100,000 employees, revealing names, addresses, bank account details and salaries.

Skelton was jailed for eight years in 2015 on the charge of fraud, securing unauthorised access to computer material and disclosing personal data. The company, which saw itself as a victim of this crime, received £170,000 in compensation.

However, a group of the affected employees sued Morrisons on the basis that they had been exposed to the risk of identity theft and potential financial loss. The employees claimed that Morrisons was responsible for breaches of privacy and confidence, as well as data protection loss.

An initial High Court judgement against the company in December 2017 was upheld by the Court of Appeal in October 2018, even though Morrisons said it had worked to take the data down quickly, provide protection for staff and reassure them they would not be financially disadvantaged. [46]

> *It is important to maintain strict rules and discipline about the way employees treat data*

Morrisons has said it will appeal to the Supreme Court. But the incident shows how companies can be vulnerable to the actions of their employees and how important it is to maintain strict rules and discipline about the way employees treat data.

Challenge 8 – Dealing with attacks

All the evidence shows that cyberattacks are now commonplace, but most organisations have yet to learn to deal effectively with them. Dealing with an attack for the first time is probably the hardest part, but companies can already learn from the experience of others and the simple recommendation is the old adage: 'Be prepared.'

This means that companies need to look at the experience of others, develop crisis scenarios and work through them at every level including the board. They need to understand why as well as how they may be at risk of attack. Planning how to respond is also important, so that there is a 'runbook' similar to that compiled by airlines. This would provide a set of procedures for routine and exceptional circumstances, which would help companies to avoid and also deal with incidents. Defence processes should be in place that can be audited.

[46] See The Independent (1 December 2017) *Morrisons data leak: thousands of staff to receive payout in landmark judgement over personal details posted online* and the Law Society Gazette (23 October 2018) *Court of Appeal upholds 'surprising' Morrisons data leak*

In cyber breaches, the victim may not actually know who is doing the attacking and it can also take a long time to work out exactly what has gone wrong. A key point to remember is that data is actually an asset in itself. Nowadays, organised crime groups are selling the malware, rather than trying to mine the data themselves, because that is more lucrative for them. Serious organised crime groups also want to hack into the core systems of organisations like banks, because it seems that a number of them will pay a ransom. Hostile states want to do the same to apply geo-political pressure.

Wannacry (see Figure 4) was a worldwide attack in 2017 on companies using older Microsoft technology. Computers were encrypted and victims were asked for a ransom to restore their files.

Figure 4 Wannacry ransom note

Source: Wikipedia

Once an attack has taken place, it is important to undertake the following five tasks:

1. Find out as quickly as possible exactly how much data has been compromised
2. Ensure that all regulations and laws regarding disclosure are complied with. Consider the need for additional disclosure [47]
3. Take steps to mitigate the damage and restore the trust of customers and other stakeholders
4. Ensure effective communication with those who are affected so they know where they stand
5. If systems have been closed down, decide when it is appropriate to restore them.

[47] GDPR requires organisations to disclose within 72 hours when personal data has been compromised. It does not require disclosure of other forms of attack, for example, when a company's systems have been hacked and money stolen from it

These tasks require some difficult judgements, which is one reason why boards need to have planned in advance how they would react in the event of an incident. For example, whether and when to disclose an attack is not always an easy decision. Having a crisis management framework may help.

> *Whether and when to disclose an attack is not always an easy decision*

When it was attacked in 2015, the telecoms company TalkTalk decided to disclose quickly, even though that meant it was not in a position to be sure how much data had been affected. The company made the decision because it wanted to warn its customers – especially the most vulnerable ones – that their bank accounts might have been compromised. Its decision to disclose was made against the advice of the police, who felt this would make it harder to catch the attacker. The company was also embarrassed by having to admit that the initial estimates of data loss were understated, and that the situation was worse than originally portrayed. Nonetheless, the company took the view that this was the right thing to do in the interest of its customers, a decision that ultimately appears to have been vindicated. [48]

Once the disclosure is made, companies need to anticipate how their stakeholders are likely to react. Too often, aggrieved customers complain to the media that – despite promises made to them – they have not been contacted and that it is impossible to get through to the company via their helplines or website. Stress testing the means of communication in advance of an attack is important, but rarely seems to have happened or been effective.

Finally, it can be difficult to make a decision about when to restore systems following a breach. It is natural for boards, senior executives and those with operational responsibility for the technology to become highly risk averse in the wake of a cyberattack. Yet the real need may be for them to weigh up the risks between trying to make the system absolutely watertight and losing current or future business. This is a business judgement, not just a technical one. It requires the technology teams to be able to talk to the business marketing teams in ways that each of them can understand. It requires boards to be aware that weighing up the risks is a truly challenging task.

Questions for boards

- Are we aware of what data we hold and where it is stored?
- Have we planned, drawing on the experience of others, what we would do in the event of a cyberattack?
- Are our defence mechanisms and our state of readiness subject to regular audit?
- Have we tested the crisis management process or business continuity plan?
- Have we considered who might attack us and why?
- Do we have a considered policy on communications and disclosure beyond the requirements set out in the law?

[48] See *TalkTalk – Customer Communication in a Crisis* case study in FRC (2016) *Corporate Culture and the Role of Boards*

- Are our systems for mitigating the damage caused by an attack sufficiently robust and reliable?
- Have we drawn an appropriate line between the desire for absolute security and the need to keep our business going?

Challenge 9 – Can codes of ethics help?

Most experts agree that codes of ethics can be an important tool in the safe development of AI. There are two levels at which these types of code would work. First, outside bodies could develop overarching industry codes. Second, elements of these could be written into the ethical conduct codes of individual companies.

The House of Lords Select Committee report raised the possibility of an AI Code. [49] It noted that companies have started to develop their own principles, including IBM and Sage (see Challenge 2). The Lords quoted the Market Research Society as stating that companies using boards, committees and processes within a self-regulatory framework will generate trust and confidence among their clients:

"AI companies or companies employing AI technology, to the extent they demonstrate they have ethics boards, review their policies and understand their principles, will be the ones to attract the clients, the customers, the partners and the consumers more readily than others that do not or are not as transparent about that," the Society said in evidence.

The Lords warned, however, of the risk that the trend for ethical principles might turn into *"a meaningless box ticking exercise"*. A code should have five overarching principles, as follows:

1. AI should be developed for the common good and the benefit of humanity
2. AI should operate on principles of intelligibility and fairness
3. AI should not diminish the data rights or privacy of individuals, families or communities
4. All citizens have the right to education allowing them to flourish mentally, emotionally and economically alongside AI
5. Autonomous power to hurt, destroy or deceive human beings should never be vested in AI.

> *AI should be developed for the common good and the benefit of humanity*

For the time being, the development of codes and best practice standards is in its infancy. Among the organisations looking at the issue is the newly formed Centre for Data Ethics and Innovation, which was established by the UK Government in 2018 and is due to become an independent statutory body under its Chair, Roger Taylor. Further work is being done by the Ada Lovelace Institute, an independent body set up by the Nuffield Foundation, which describes one of its roles as being *"to define and inform good practice in the design and deployment of AI"*. TechUK, an industry body, has a Digital Ethics Working Group designed to keep its members up to date on digital and data ethics.

[49] House of Lords Select Committee on Artificial Intelligence (2018) Op cit

One of the most important challenges is how to move from the sort of high overarching principles described by the House of Lords to more granular principles that can be made to apply at the coal-face. The challenge is to ensure that the more granular elements do not become so specific that they are silo-bound and simply reflect the nature of the organisation – or even the division – that has introduced them. Issues can and should extend across different sectors and activities.

The technology teams of banks, for example, are working in a world that is startlingly different from colleagues who are actually delivering banking services. Teams are hired from pools of technology specialists, who might equally be working in another industry. The behavioural expectations imposed on them by their employer need to reflect both what is generally expected of colleagues, but also specific expectations related to the nature of their activity. Ideally, the latter should be benchmarked against what goes on in technology departments elsewhere.

It follows from this that those designing overarching codes need to refine the process to include more granular elements that would still apply across different industries. In addition, firms need to anchor any specialised provisions for technology activities – such as software design – in their overall code of ethics or conduct. The code may still have some general provisions relevant to all employees, for example around respecting data privacy and data security.

> *Codes of ethics are an important means of encouraging appropriate use of technology*

Finally, it is very important that codes of ethics or conduct are supported by effective Speak Up or whistleblowing arrangements. Those responsible for these arrangements need to be alert to the possible issues around AI.

Conclusions for boards

- Codes of ethics are an important means of encouraging appropriate use of technology and can provide a secure framework for companies to make the most of opportunities
- While overarching principles that take account of societal expectations are a good and necessary starting point, codes need to be sufficiently granular to be of practical help to employees
- This must be done without creating silos in which technology staff are treated differently from other employees. Provisions relating to their work should be anchored in the firm's general code, and specific expectations should be benchmarked against other organisations and sectors
- Codes need to be backed up by effective Speak Up and whistleblowing arrangements.

Expertise and the Boardroom

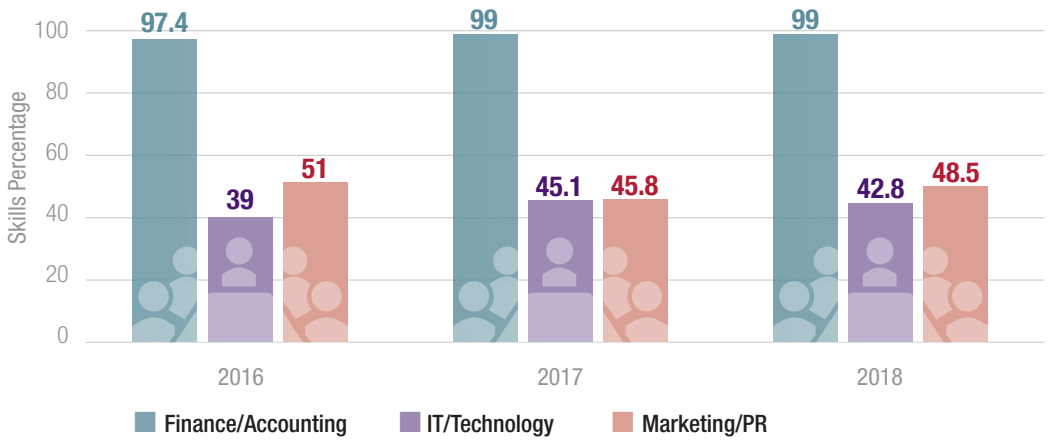
Dealing with AI and cyber risk raises an important question about board composition. Boards must be ultimately accountable for what happens, but does this mean that companies need to bring formal expertise on to the board and, if so, how much expertise is required?

Most of those who have grappled with this question say that common sense, plus the ability to ask good questions and to obtain high quality advice, are more important than recruiting specific technological expertise.

This may seem counterintuitive. For example Grant Thornton, in its most recent annual review of UK corporate governance, registers some surprise that boards lack expertise in technology. Its 2018 survey showed that only 43 percent of FTSE350 companies disclosed that they had technology expertise on their boards. IT and technology ranks way below financial expertise and is roughly on a par with marketing and PR (see Figure 5 below). This is, admittedly, way ahead of HR and law at 18 and 24 percent respectively. [50]

Figure 5 What skills dominate boards?

Percentage of companies with one or more board members with experience in the areas below.



Source: Grant Thornton (2018) *Corporate Governance Review 2018*

Yet there are two main arguments against rushing to acquire technology expertise. First, there is always a risk when a board appoints just one or even two directors to deliver a critical skill. The board as a whole may come to rely too heavily on such people, forgetting collective responsibility in the process and losing its interest in challenge.

The second is that technology experts may not be able to contribute actively to all the other tasks that boards have to undertake such as decisions on strategy and capital allocation, financial controls, reputation management, health and safety, and remuneration. Or they may have limited expertise across the whole bandwidth of technology.

[50] Grant Thornton (2018) *Corporate Governance Review 2018*

Directors who are hired for their technology expertise must still be able to participate in discussions on these other issues. One difficulty appears to be generational: those with specific technological knowledge are usually younger and have not yet accumulated sufficient practical, commercial experience to address the rest of the board agenda. The most experienced board members are likely to have spent the bulk of their careers in businesses where technology was less central. They may therefore be inclined to shelve or downplay the issue.

> *Directors who are hired for their technology expertise must still be able to participate in discussions on other business issues*

Admittedly, the need for specific expertise depends heavily on the nature of the business. In sectors where AI is a core product or critical to the business model – as in some branches of finance – then the board is likely to need expertise that goes beyond the presence of one knowledgeable director. Curiously, though, even in this area Grant Thornton notes that only 26 percent of financial services companies disclosed that they had board members who were experts on IT.

It is worth noting that only five out of nine current board directors of Facebook Inc have an IT background. The others come from a range of disciplines, including economics, corporate finance, law, pharmaceutical chemistry, philosophy and political science. Many of these have, nonetheless, worked in technology-oriented industries and, at least in theory, have some practical experience of the interface between business and technology risk.

> *Directors need to be able to ask the right questions, insist on answers they can understand and set limits*

Most companies use AI, not as an end in itself, but as a tool to make their business more competitive. Directors need to be tech-aware, to keep up to date with the way AI is developing and to be able to relate this to the company's strategy, business model and its inherent risks. That means, above all, being able to ask the right questions, insist on answers they can understand and set limits.

For this, boards need reliable information. The personality of the CIO, CTO or Chief Risk Officer is critical. Are they able to explain what is going on in plain English? Can they be trusted?

Consider the sort of judgements that a board may need to make. One might be a situation where a company's systems have been hacked and the board needs to know whether to bring them back up (see Challenge 8). This involves a delicate judgement.

Bringing systems back up too early might mean the problem recurs and the damage to the business will multiply. Waiting too long might mean such a serious loss of business that the company will collapse anyway.

A technology expert who wants to give an absolute answer would perhaps be inclined to wait. A marketing person wanting to resume sales might be inclined to plunge in. Boards need to have the mix of skills and experience that will help them strike the best balance, based on an understanding of their risk appetite. This is as much about business risk as about technology risk.

This is not to say that boards should ignore technology issues, which should be on the risk oversight agenda. According to the Grant Thornton *Corporate Governance Review 2018*, all FTSE350 utilities, telecommunications and technology companies and all but one consumer services company disclosed technology risk as a key threat.

Boards should be aware of the questions they need to ask and confident of the answers they are getting:

"Boards are asking the wrong question when they want to know if we are OK," says one Chief Executive. *"The answer to that is always no. But you should be aware of the risks and how they are mitigated."*

This is not, however, just a defensive matter. Evolving technology offers great business opportunities. Boards need to keep abreast of developments so that they can make the most of them. This puts an onus on directors to keep themselves up to date and watch for opportunities to adapt their business model.

Questions for boards

- How does the use of AI sit with our values?
- Who is in charge of technology? Do they make sense and answer questions intelligently?
- Who is likely to attack us and why?
- Do we know exactly what data we own and where it is so we can tell what we've lost in an attack?
- Are our firewalls up to date and are the staff properly trained to take precautions?
- Has the company done a stress test against the impact of leakage/hacking? What was the result?
- Has the CIO/CTO worked in crisis management and shared experience of risks with others?
- What is the CIO/CTO budget?
- Who are the company's suppliers?
- Could we see the control room?
- Are we confident that our algorithms are free of bias?

Conclusions for boards

- Boards are accountable for what happens as a result of the application of technology to the business. They cannot hide behind experts whose answers they don't understand or who they do not trust
- Business experience, common sense and sound advice are more important for directors than technology expertise. *"You do not have to be able write software to ask the right questions,"* said one experienced director in discussions with IBE
- The appointment of the CIO/CTO is very important. Boards must ensure that they are happy with the successful candidate
- Boards may be helped by the appointment of an advisory committee of experts, even though this can raise problems if the experts end up being paid more than the directors themselves
- The use of technology presents opportunities as well as risks. Boards should factor it into their strategic planning, both for the short and long term.

> *Business experience, common sense and sound advice are more important for directors than technology expertise*

Conclusion

This Board Briefing started from the premise that AI offers huge opportunities to companies, the economy and broader society. It will, however, be hard to realise these opportunities if AI is not introduced in a climate of trust. Ethical considerations are thus paramount and, indeed, competitive advantage may well accrue to those that take the trouble to develop their ethical understanding alongside the introduction of new technology itself.

Up till now, many boards have been reluctant to confront the issues or have adopted primarily a defensive approach centred around the need to develop defences against the threat of hacking and loss of data privacy. Yet directors should not be put off by lack of technical expertise. Of course, they need to understand what technology delivers and keep themselves up to date with the way it is changing and developing, but most of the questions they then need to ask themselves are philosophical and ethical. This requires them to draw on the company's and their own values for answers. These questions may be challenging, but they are not difficult in a technical sense.

This Board Briefing will have succeeded if it encourages boards to mainstream their thinking about AI. Many of the questions that boards will have to ask are about where to draw the line, for example in the use of potentially biased algorithms in recruitment or the use of personal data to target advertising. These types of question sit quite naturally within the board's regular discussions about risk appetite, risk management and oversight. It is better that AI issues are handled in that context, rather than sidelined and dealt with in a separate silo. Responsibility for oversight and management of AI should also be embedded in executive remuneration schemes.

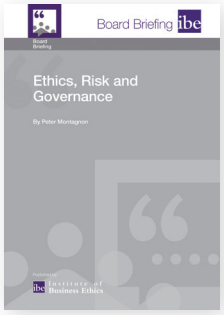
For effective oversight, it is very important that different parts of the business are able to talk to each other in language they can all understand. The tech world tends to have its own culture and own way of thinking. Yet companies will find it difficult to adapt to the world of AI if they end up with two different and entirely separate cultures within the same organisation.

Similarly, there is a need for a bigger discussion between the corporate world that is starting to use AI, civil society, government and regulators about how to handle change. Companies cannot handle all the employment implications on their own. They need a legal and regulatory framework that enables them to deliver the benefits of AI while enjoying the trust of the public.

The best chances of success lie with an inclusive approach, rather than one that is exclusive and repeats the mistakes made around globalisation, which left the public resentful at the impression of benefits being siphoned off by the elite. While building the right governance in their own organisations, companies need to engage proactively in public dialogue. In so doing, they should listen carefully to the views of others, but also not be afraid to put their own case.

Related IBE Publications

IBE publications provide thought leadership and practical guidance to those involved in developing and promoting business ethics, including senior business people, corporate governance professionals and ethics and compliance practitioners. Some recent publications related to this topic which you might be interested in include:



Ethics, Risk and Governance

Peter Montagnon

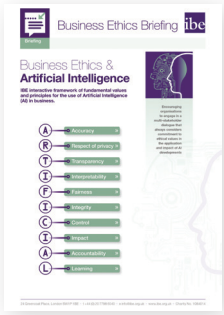
Setting the right values and culture is integral to a company's success and its ability to generate value over the longer term. The challenge for business is how to develop and embed real values. This requires leadership and is a core task for boards. Many boards acknowledge the importance of a healthy corporate culture, both because of the role this plays in mitigating risk and because of the value to their franchise of a sound reputation. This IBE Board Briefing sets out why directors need to be actively involved in setting and maintaining a company's ethical values and suggests some ways to approach it. It aims to help directors define their contribution to the maintenance of sound values and culture.

Culture Indicators: understanding corporate behaviour

Peter Montagnon

Boards are increasingly focused on corporate culture, yet they often struggle to understand the forces that drive behaviour in their business. Culture cannot easily be measured, but boards can and do have access to a range of information that will shed light on the culture of their organisations.

Culture Indicators: understanding corporate behaviour analyses survey data and draws on interviews with directors and those who advise them to provide practical and tangible assistance for boards in how to understand the corporate culture of their organisations. It examines a wide range of relevant indicators and how to interpret them in order to produce a useful and authentic picture of the culture of a business.



Business Ethics and Artificial Intelligence

What is artificial intelligence (AI) and what is its impact on our society? What are the biggest risks that new technologies can pose? How will we seek to control the way it affects our daily lives? Are we preparing ourselves sufficiently? This Business Ethics Briefing looks at these questions and sets forth a framework of fundamental values and principles for the use of AI in business. The IBE encourages organisations to engage in a multi-stakeholder dialogue that always considers commitment to ethical values in the application and impact of AI developments.

Other IBE Resources

E-learning

The IBE's e-learning package *Understanding Business Ethics* is available in English, French, German and Spanish.

This short introductory online training course is designed to raise awareness of business ethics and provide an understanding of why ethical standards in the workplace matter.

The course is designed to support employees at all levels, in organisations of any size and in any sector to 'do the right thing'. The programme takes approximately 35 minutes to complete.

The IBE's e-learning package is used by professional bodies to develop ethical sensitivity in their members and for CPD (Continuing Professional Development) accreditation.

If you would like to offer this to multiple users, please contact us.

Say No Toolkit

The IBE's *Say No Toolkit* is a decision-making tool to help organisations encourage employees to make the right decision in difficult situations. The *Say No Toolkit* delivers immediate guidance to employees on a wide range of common business issues, especially those that could lead to accusations of bribery.

Employees tap through a series of questions about the situation they face and the tool will provide the right decision to take: Say No, Say Yes or Ask. The answer also makes it clear why it is important to make that decision so your employees can have the confidence and the knowledge to respond correctly.

Organisations can use both the IBE *Say No Toolkit* app and website for free. The app can be downloaded on to any smartphone or tablet.

Simply go to www.saynotoolkit.net

The *Say No Toolkit* can be customised and branded to suit your organisation's needs and detailed procedures. For more information email info@ibe.org.uk or call the IBE office on +44 20 7798 6040.

For details of all IBE publications and resources visit our website www.ibe.org.uk

Corporate Ethics in a Digital Age

IBE Board Briefings aim to support board members and those who advise them by drawing their attention to particular ethical issues and offering practical ways to approach them.

Growing reliance on data and the integration of AI into business activity has thrown up some large challenges for governance. Boards not only have to manage a new set of risks and opportunities – they do so in a world that is rapidly changing in ways that make it harder for them to exercise control.

This Board Briefing presents nine challenges around the use of AI, offering practical thoughts about how they can be addressed, and looks at the expertise that is required in the boardroom. These challenges are less about the technology itself than how it is applied, requiring a philosophical and ethical approach to resolving the dilemmas that AI provokes. The decisions that boards must take will fit naturally, therefore, into their general view of risk appetite, risk management and oversight.

The requirement to manage the consequences of AI is a major challenge that boards must pluck up the courage to address, even though they may still be at the learning stage. The principal aim of this Board Briefing is to encourage boards to put this issue firmly on their agenda.

ibe Institute of Business Ethics

ISBN 978-1-908534-42-2 Price: £30