

C|B|I|



AI: Ethics into practice

Steps to navigate emerging ethical issues

August 2019

Exclusive practical guide for members

Ethics can help the UK surge ahead in the AI race

Jump straight to a section of our practical guide and get started:

Three E's for practical ethics	4
Embed: Put ethics at the core of your approach to AI	5
Engage: Empower your workforce when it comes to AI	11
Explain: Engage customers and the public on AI	18

AI is transforming the way we live and work

Artificial Intelligence (AI) – technology that exhibits behaviours that appear intelligent and autonomous – is having a seismic effect on society. CBI research shows that it is leading the charge for business investment and **is set to impact most sectors by 2021**, transforming the workplace to boot.¹

AI will help to tackle some of the biggest issues we face. But the changing world of work – indeed, a changing world – has led to new ethical challenges. From identifying job candidates to detecting financial fraud, AI is being used to make decisions about people that could be lifechanging. Businesses are already taking a more ethical approach, but sometimes the conversation focuses more on principles than practice.

That's why the CBI has created this guide. Our aim isn't to 'solve' the ethical questions societies have been debating for millennia. But, whether your business is an AI pioneer or is yet to adopt it, this guide offers real life advice on putting ethics into practice at this point in time.

Ethical AI isn't just a business challenge – it's an opportunity for the whole of the UK

At home and abroad, the spotlight is being turned on AI, with countries around the world forging their own strategies to leverage this game-changing technology.



10%

UK GDP could be up to 10% higher in 2030 due to AI...



£232

... which is equivalent to an additional £232 billion for the economy.²

Building trust in AI technology by taking an ethical approach to AI adoption will be a key enabler for the UK and UK businesses to lead.

2021
AI is set to impact
most sectors by
2021

1. CBI, Disrupting the Future (2017)

2. PwC, The economic impact of artificial intelligence on the UK economy (2017)

“The UK has a unique opportunity to forge a distinctive role for itself as a pioneer in ethical AI, which would play to our particular blend of national assets.”

House of Lords Select Committee on Artificial Intelligence³



With outstanding universities, renowned civic institutions, and a thriving tech sector, ethical AI is an opportunity for the UK to distinguish itself with ideas and standards that can be exported across the world.

Ethics can support business to answer some of the difficult questions raised by AI

How do you explain a complicated computer process to a customer? How should you introduce new technologies into the workplace? And, most importantly: not *'could we do this?'*, but *'should we do this?'* – and, if so, *'how should we do this?'* An ethical approach to AI requires openness about the challenges and considerations which could undermine public trust. Businesses are grappling with these questions today – and getting it right will help ensure that the social and economic benefits of AI are felt by everyone.

“It is society’s decision to trust or not to trust AI, and the companies that deliver it, that will determine its success. As leaders and companies, we must earn that trust.”

Bill Kelleher, CEO, UK & Ireland, IBM⁴



But how do you move ethics out of the lecture theatre and put it into practice?

The word 'ethics' can conjure images of philosophical debates and moral mazes, so it can be hard to identify the practical steps that lead towards a more ethical approach to AI. But in reality, meaningful ethics is an issue organisations already think about, through effective governance, employee empowerment, and customer engagement.

3. House of Lords Select Committee on Artificial Intelligence, AI in the UK: ready, willing and able? (2018)

4. Bill Kelleher, Winning trust in Artificial Intelligence (2019), <https://www.cbi.org.uk/articles/winning-trust-in-artificial-intelligence/>

Three E's for practical ethics

The CBI has developed three E's for practical ethics: embed, engage, and explain

Some governments are creating high level principles; some firms are setting up ethics boards. Based on extensive engagement with industry, the CBI has created this practical guide to support businesses like yours on their AI journeys, from first deciding to use AI to how you can communicate that approach to employees, customers, and the public.



Embed | Put ethics at the core of your approach to AI

Step 1: Update governance processes to minimise risk and address uncertainty



Engage | Empower your workforce when it comes to AI

Step 2: Take action so all employees can participate in an AI-powered workplace

Step 3: Improve data and diversity to eliminate unfair bias



Explain | Engage customers and the public on AI

Step 4: Increase data security and privacy to boost consumer trust

Step 5: Put trust in the hands of customers by explaining AI

Embed: Put ethics at the core of your approach to AI

Key action that can help put ethics at the core of your approach to AI

Step 1: Update governance processes to minimise risk and address uncertainty

How you use AI will shape the ethical questions that matter most to your business

AI is being used in different and increasingly advanced ways across businesses and sectors:



In **financial services**, AI is being used to help combat increasingly sophisticated money laundering and fraud threats by helping to detect suspicious activity.



Businesses in the **energy sector** are using AI to help optimise the efficiency of how energy is supplied.



The **legal sector** is using AI to improve how millions of contracts are analysed and empower lawyers to make more informed decisions.

Diverse AI use has led to diverse ethical challenges and priorities: autonomous vehicles will raise different challenges to chatbots. AI both makes existing challenges more urgent and leads to new ones – for example, how to keep humans involved in technology decisions.



“It’s important to remember the potential of AI to help solve some of our biggest challenges. In a range of cases, it would be unethical not to use AI.”

Dr Will Cavendish, Global Digital Services Leader, Arup



There is no one-size-fits-all approach when it comes to an ethical approach to AI. Crucially, ethics must have a meaningful impact on a business's day-to-day operations

The majority of businesses recognise the importance of engaging on ethical issues. Firms are taking different approaches – some are using high level principles, others are focusing on ethical impact assessments or setting up new committees. Often, firms are investing time and people in multiple strategies.

However advanced in their thinking and whatever their approach, businesses are united on one front: it's imperative to walk the talk. A common starting point to help AI impact on businesses' day-to-day is ensuring AI fits into governance and risk management frameworks.

“

“Ethics needs to flow through a whole business to be effective.”

Sue McLean, Partner, Baker McKenzie

Step 1: Update governance processes to minimise risk and address uncertainty

What is this ethical issue?

“Ethics is a process, not a set of answers. It's about how you take a decision, what the inputs are, and who gets a say.”

Judy Baker, Partner, Ward Hadaway

”

Governance and accountability go to the heart of how businesses engage with AI technology. By outlining principles and processes and assigning clear lines of accountability, businesses can make decisions with integrity even when faced with risk or uncertainty.

What's the challenge?

AI adoption doesn't happen in a bubble - it impacts on core business legal and governance responsibilities towards everything from accountability to privacy. Crucially, using AI technology can lead to new questions about classic business issues. Governance structures should be designed and refined to cope with the inevitable uncertainties that arise, and build accountability into how firms use AI.

What could my business do?



Evaluate how AI fits into existing governance frameworks or consider creating new principles or frameworks.

- **Map how AI fits into your existing frameworks and principles**

AI is an opportunity for organisations to reflect on and be clear about their values. Doteveryone's [Consequence Scanning manual](#) helps tech teams to ensure that the products and services they create are aligned with their organisation's values and culture.

- **Consider creating new governance frameworks specifically for AI**

There are a number of existing codes, guidelines, and frameworks businesses can look to and adapt, for example the [NHS code of conduct for data-driven health and care technology](#) or the [IEEE Standards Association's Global Initiative on Ethics of Autonomous and Intelligent Systems](#).

- **Use an algorithmic impact assessment (AIA)**

An AIA, for example [the AIA created by the Canadian government](#), assesses and mitigates risk and impact on society, and provides a framework for ethics by design.

- **Explore sectoral and technology voluntary codes of conduct**

Some businesses have signed up to voluntary codes of conduct, for example the Safe Face Pledge to prevent abuse of facial analysis technology.

Use the EU's guidance on AI governance

The EU Commission's pilot [Trustworthy AI Assessment List](#) can be incorporated into an organisation's existing governance mechanisms. The list aims to help businesses operationalise the EU's seven requirements for Trustworthy AI, which should be lawful, ethical, and robust.

Case Study: Barclays Bank

Barclays looked beyond banking to shape its approach to AI

When developing its AI principles, Barclays looked outside the financial services sector to other businesses' and countries' AI frameworks and principles, for example the European Commission's Ethics Guidelines for Trustworthy AI. The AI principles were refined with the input of external legal firms and internal data science teams to ensure that they both reflected best practice from outside and could be practically realised within Barclays.

AI projects must go through Barclays's existing governance structures (for example, product approval and technology governance), which require evidence of all principles being adhered to.

"AI technology is not new, and it has been used to different degrees of sophistication for a long time. Existing governance processes provide a good starting point and it is worth adapting best practice to embed it in existing processes, instead of starting from scratch."

Ana Perales, Strategic Transformation Director, Barclays Bank



Set up an AI ethics board or advisory council to test assumptions and make decisions using a range of perspectives.

Committee responsibilities might include setting ethical guidelines, evaluating whether new products and services align with ethical guidelines, or investigating complaints. To be effective they will have:

- **Diversity in a broad sense**

Diversity means a range of protected characteristics like gender, ethnicity, or age; representation of different disciplines and the different teams in an organisation; a wide mix of seniority levels. This will help contribute to robust, comprehensive conversations.

- **External representation**

Engaging with external stakeholders, including critics, might be desirable. Understanding a less positive view will lead to more thoughtful decision-making.

- **Transparency about role and remit**

Transparency on a board's role and when it should be consulted is essential. Not all AI technology will need to be put to an AI board or council. What's more, not all boards have decision-making powers – some act in a more advisory capacity and make recommendations to decision-makers. The purpose of the board or council should be made clear from the start.

- **Authority to make meaningful decisions or recommendations**

If a board has decision-making powers, make sure there are well-defined enforcement systems in place. If it acts in a more advisory capacity, ensure it has the power to make meaningful recommendations, and be transparent about when and why recommendations are rejected.



“At Brunel, researchers must submit all research projects involving human participants to the university’s Research Ethics Committee to ensure regulatory and ethical compliance. Everyone in the team is involved in getting ethical approval. An individual’s perspective might not be perfect, but a group is more likely to get things right. Industry could learn from this academic framework.”

Professor Panos Louvieris, Professor of Information Systems, Brunel University London



Introduce human oversight of AI systems.

All businesses we interviewed emphasised human responsibility as a fundamental pillar of governance and oversight processes, particularly given AI’s potential for autonomy. Build clear lines of human accountability that go to the top of an organisation. AI is simply a tool – people, not machines are ultimately responsible for its decisions.

Although having a ‘human-in-the-loop’ can impact the speed of decisions, businesses emphasised the importance of ensuring accountability and reviewing the quality of outcomes. Different levels of human intervention during AI design, decision-making, and monitoring cycles can help to reduce the chances of unexpected consequences.



“The less oversight a human can exercise over an AI system, the more extensive testing and stricter governance is required.”

European Commission Independent High-Level Expert Group on Artificial Intelligence⁵

“AI does not change the need for effective oversight. Humans have to evolve our approach to governance to stay ahead of the challenging questions posed by technology.”

Adam Green, Chief Risk Officer, Equiniti



Case Study: AnyGood?

AnyGood?'s hybrid recruitment model ensures humans are in the loop

After commissioning independent research into trust and recruitment, crowdsourcing talent platform AnyGood? had greater awareness of the technology trust and friction points in a hiring process – moments when people trusted algorithms and moments where they were more comfortable with the involvement of a person. This allowed AnyGood? to create a new model of hiring using a combination of AI and human intervention, rather than relying on a black box solution for sensitive decisions and processes, which although seemingly efficient actually risks undermining trust.



Continue to monitor how AI performs through technological means.

Humans won't necessarily be able to monitor every aspect of an AI system, so businesses should introduce other monitoring methods. Checking an algorithm continues to perform as expected will be better for businesses, too.



Check suppliers and customers align with your ethical principles.

Most businesses won't design their AI in-house but will rely heavily on third parties. Work closely with your suppliers to test whether their technology aligns with the ethical approach you expect – don't assume it will.

Businesses who sell AI can do the same. Some organisations, for example Microsoft, have chosen not to support specific scenarios where customers' use of the technology didn't match their principles.

- **Use your contract as an opportunity to align ethical approaches**

Some businesses are adding into contracts either their own AI principles or best practice guidance. For example, they might expect their suppliers to assist with providing explanations about AI decisions in line with the law and best practice. Engaging suppliers early on ethics can lead to a flow-down of ethical AI.

- **Ask your AI supplier questions on ethics**

Many firms have started to develop a standard list of questions they ask suppliers when buying AI technology, for example 'Will my business know if and when the supplier updates the algorithm?' or 'What data was used to train the algorithm?'. Businesses should make sure they are undertaking due diligence when choosing an AI supplier.

Engage: Empower your workforce when it comes to AI

Key actions to help empower your workforce

Step 2: Take action so all employees can participate in an AI-powered workplace

Step 3: Improve data and diversity to eliminate unfair bias

Great people practices: better for society, better for business

AI could boost businesses' productivity, with the potential to increase the UK's GDP by 10% by 2030.⁶ To allow everyone to participate in the economy of tomorrow, businesses can take action to manage AI's impact on the workforce and increase diversity.

Action will benefit businesses too



22%

Firms with the highest levels of employee engagement see profits 22% higher than those with lower engagement.⁷



72%

Companies that develop the strengths of their staff have reduced staff turnover by up to 72%.⁸



up to 35%

Businesses with the highest levels of gender and ethnic diversity are 15%-35% more likely to outperform their rivals.⁹

The evidence is clear: the same actions that make up an ethical approach on AI will also make businesses more competitive.



6. PwC, The economic impact of artificial intelligence on the UK economy (2017)
 7. CBI, Great job (2019)
 8. Ibid.
 9. McKinsey, Delivering through diversity (2018)



30%

of UK jobs could be at high risk of automation by the early 2030s

Step 2: Take action so all employees can participate in an AI-powered workplace

What is this ethical issue?

Technology is changing the way we live and work – and AI will impact upon all sectors and a wide variety of roles to a greater or lesser extent. Businesses must take action so that the benefits of AI are distributed throughout the workforce. They should ensure all employees are empowered to participate in a workplace that might already look very different to the one they entered and is fast continuing to evolve.

What's the challenge?

Businesses we interviewed emphasised AI's effectiveness at automating routine or repetitive tasks, leaving employees with more time to focus on high-level decision-making, creative assignments, and customer service.

But ensuring AI's benefits are felt by all will require business action. Although just 1% of businesses report using technology to eliminate headcount,¹⁰ the least optimistic estimates suggest that **up to 30% of UK jobs could be at high risk of automation by the early 2030s.**¹¹ AI could also create millions of new jobs, but certain demographics and sectors are at higher risk.¹² In England, women, young people,¹³ and those without degrees are more likely to be affected.¹⁴

There remains a challenge to prepare millions of currently low-skilled adults for the new future of work. Encouragingly, nearly nine in ten (87%) businesses anticipate upskilling employees in their current roles during the next 12 months.¹⁵ But with 23% of workers worried that their job might not be needed in the next decade because of technology, it's a business imperative to change people's perceptions by taking steps to improve trust.¹⁶ Industry must prevent fears about job losses becoming a reality and plan for the skills that will be needed in the future.

10. KPMG, Easing the pressure points: The state of intelligent automation (2019)
11. PwC, Will robots steal our jobs? The potential impact of automation on the UK and other major economies (2017)
12. PwC, AI will create as many jobs as it displaces by boosting economic growth (2018)
13. ONS, Which occupations are at the highest risk of being automated? (2019);
14. PwC, Will robots steal our jobs? The potential impact of automation on the UK and other major economies (2017)
15. CBI, Educating for the modern world (2018)
16. Research from Commission on Workers and Technology (2019)

What could my business do?



Prepare all employees for working in an environment that is rapidly changing.

Embrace a culture of lifelong learning, embed upskilling, and introduce personal training into training strategies. Harness your existing talent, by looking for hidden skills within your workforce and retraining at work.¹⁷

Case Study: Microsoft

Microsoft is supporting businesses to put people first when it comes to AI adoption

AI will create new roles, change existing roles, and impact how businesses work across all industries. Microsoft has created the AI Business School to share insights and practical guidance from top executives on how to strategically apply AI. Modules are available to help leaders and their teams work through defining an AI strategy, enabling an AI-ready culture, guidelines on how to develop and use AI in an ethical and responsible manner, and information on what the latest innovations are.

Microsoft also supports its partners to engage their employees. Johnson Controls, a provider of building technologies and solutions for over a century, leads the way today with its GLAS smart thermostat. The thermostat was built on the Microsoft cloud and AI solutions following the company's internal Tech Days challenge, which gives employees two days to present ideas to company leaders for potential future development.

"While it is important for organisations to embrace AI's potential, what matters is not just what AI can do, but what it should do. So, whilst repetitive tasks will be replaced by AI, creating time for people to think critically and make EQ-decisions, it is also important to ensure that everyone has the skills to embrace and thrive in an AI-enabled world."

David Frank, UK Government Affairs Manager, Microsoft



Engage employees who use or will use AI technology.



"Employees must be front and centre in any ethics oversight boards and governance processes."

Ivana Bartoletti, Head of Privacy and Data Protection, Gemserv



57%

of employees think that technology has had a positive impact on their role

Many firms are familiarising employees with AI technology through internal roadshows, tech demos, and speakers. But research reveals that nearly six in ten workers (58%) say that they're unable to influence how technology changes their workplaces.¹⁸

Businesses we interviewed reported increasing staff engagement by encouraging employees to explore where they would find it useful to introduce AI within their own roles. Positively, **over half of employees (57%) think that technology has had a positive impact on their role** – if they had more say over where it were introduced, that could be even higher.¹⁹



“Businesses should engage staff who are doing the work by encouraging them to propose AI use cases within their own roles. We’re open and honest with our employees: we tell them what we’re doing and where we want to be – and ask them to help us get there.”

Shamus Rae, Head of Innovation, Digital Labour and Cognitive Transformation, KPMG

Engage your HR team early, to assess potential impact on jobs or reskilling needs. Investing in the technology is not a foregone conclusion – for example, one firm we interviewed decided not to invest in AI after identifying it did not have the right reskilling training in place yet to support its workforce.

18. Research from Commission on Workers and Technology (2019)

19. Ibid.



16%

of people working in AI have left their job due to a decision they felt could have harmful consequences for people and society



Engage employees who are developing AI technology.

Taking an ethical approach is bottom-up as well as top-down. Embed ethics into products and systems by designating time and resource for developers to think about the impact of their products and services – for example, through consequence scanning. Have transparent escalation processes for raising ethical concerns.

AI workers want to work for firms who do the right thing – and are voting with their feet when they disagree with their employers. Research suggests that **16% of people working in AI have left their job due to a decision they felt could have harmful consequences for people and society.**²⁰ With each lost tech worker estimated to cost a company £30,000,²¹ and industry already reporting a shortage in AI skills as the main barrier to adoption, businesses can ill-afford to lose talent.²²

Step 3: Improve data and diversity to eliminate unfair bias

What is this ethical issue?

From identifying job candidates to recommending care pathways for patients, AI is helping to make decisions which could have lifechanging impacts. Businesses must continue to recognise the need to eliminate unfair bias and discrimination. AI alone won't solve the complex societal dynamics that that lead to unfairness, but fair algorithms could support greater social mobility and more equal opportunities and outcomes.

20. Doteveryone, People, power and technology: the tech workers' view (2019)

21. Oxford Economics, The cost of brain drain (2014)

22. CBI, Disrupting the future (2017)

What's the challenge?

- **Data:** If algorithms are trained on data that is incomplete or contains historic prejudice against particular groups, it will learn those unfair biases and incorporate them into its decisions, entrenching existing unfairness and barriers.
- **Teams:** Firms are at different stages when it comes to achieving diversity, but it is particularly important when it comes to algorithms. Diverse teams will be more innovative. They are more likely to spot problems in data, and challenge assumptions that could lead to unfair bias being programmed into AI.

Taking action to reflect diversity and inclusion in teams and data will lead to social benefits and makes good business sense too. Businesses who prioritise fairness and inclusion are more likely to create algorithms that make better decisions, giving them the competitive edge.

What could my business do?



Develop stretching targets and metrics to work towards fairness, and continue to collect data on, monitor, and adjust them.

'Fairness' has many definitions, both technical and social. Different metrics and standards might be required, depending on context and use case. For example, companies who have doubled their diversity and inclusion efforts to increase the number of female or BAME senior leaders in their business are creating adjustable targets, based on both where they are and where they want to be.

Make sure you listen to a diverse range of voices from the beginning when deciding what targets to set in your business.



Leaders should champion diversity and make it a shared priority across their business.²³

People in positions of power should champion greater inclusion, helping to create a culture that will encourage colleagues at every level to prioritise diversity and inclusion.



Question pre-existing practices.²⁴

Challenge the everyday. Questioning interview panels or candidate shortlists that have the same type of person will lead to more diverse teams, which have the range of perspectives that can design more inclusive AI.

23. CBI, *Mind the gap* (2019)

24. *Ibid.*



Use available technical tools to test and monitor algorithms and datasets.

A number of organisations have tools designed to detect AI bias:

- [IBM's Fairness 360 Kit](#) is an open-source toolkit containing metrics to check for unwanted bias in datasets and machine learning models, and algorithms to help mitigate such bias.
- Google's [What-If Tool](#) is a feature on the open-source TensorBoard. It lets users try five different mathematical types of types of fairness, as well as allowing them to analyse the performance of their AI models with new data or datasets.



When data is incomplete or contains historical unfair bias, take steps to make it more inclusive.

Data that is incomplete or contains unfair bias can be cleaned. Businesses are also taking steps like collecting more representative data to improve the inclusiveness of their datasets.

Case Study: Yoti Ltd

Yoti makes sure its age scanner works the same for people from all walks of life

Bias for Yoti's age estimation technology, Yoti Age Scan, arises when it is less effective for certain demographics than it is for others. Yoti undertakes assessments of its training data to understand what the mean error rating is, and when a demographic suffers a wider margin of error than average, it undertakes outreach projects to gather additional training data and ensure the algorithm becomes as accurate as possible across all demographics.

After realising its algorithm was underserving one demographic group, Yoti gathered further age-verified images from various locations, such as Nairobi, Kenya. It demonstrated the technology to individuals and asked if they would provide a verified date of birth and an image. The purpose of the exercise was explained to the individuals, and they were informed of the process for withdrawing their consent for Yoti to use their image, if they wished to do so at any point. The additional data helped improve the accuracy of the algorithm.



Collect new data with AI in mind.

Historic data has not predominantly been collected for AI training, and cleaning it can be expensive. Think about potential AI use cases you may have in a year or two to help define the fresh data you would like to collect.

Explain: Engage customers and the public on AI

Key actions that can help engage customers and the public

Step 4: Increase data security and privacy to boost consumer trust

Step 5: Put trust in the hands of consumers by explaining AI

Clear communication on the issues customers care about is imperative to building trust

People need to have trust in the way that all technology – and particularly AI – gets smarter and solves new challenges. Business is only one voice in the public conversation on how to best take an ethical approach to AI, but it's an important one. **Nine in ten people (92%) expect businesses to speak out on key social issues** and are likely to view CEOs more favourably for doing so.²³

Clear communication with customers will be one of the key enablers that builds trust – demonstrating that businesses care about the important issues. In our conversations with firms, they identified some key areas where industry could better engage customers and the public.

Step 4: Increase data security and privacy to boost consumer trust

What is this ethical issue?

AI is built on digital information, not just nuts and bolts. From digital personal assistants who can answer questions, to algorithms that recommend films based on what we've already enjoyed, the products and services we love are powered by our data.

With technology under greater scrutiny, the way a company treats personal data is the most important characteristic when a customer decides which businesses to buy from and work with.²⁴ Businesses must commit to upholding the highest data protection and cyber security standards.

23. CBI, *Everyone's business tracker* (2018)

24. *Ibid.*



92%

Nine in ten people expect businesses to speak out on key social issues

“

“With AI being deployed at scale, businesses must understand that they exist in a connected economy and are only cyber secure if the whole system is. We have to build resilience in. Building consumer confidence in data security will be critical as AI becomes normalised across society.”

Nathan Marsh, Director of Digital Transformation, Atkins

What's the challenge?

Privacy and security can be a barrier to trust: half of the public are wary about the motives of businesses who use their data – a view informed by the frequency of cyber security threats and a feeling of lack of control and transparency.²⁵ With 55% of firms facing a cyber attack in 2019, it's essential to improve cyber resilience and protect the data that underpins AI to turn the dial on trust.²⁶

Customers aren't the only ones who have concerns. AI pioneers – early adopters and champions of digital innovation – report concerns around security and privacy as one of the top three barriers to business adoption of AI.²⁷ For businesses, AI systems can bring with them new risks, as they introduce new kinds of complexity or make existing risks more challenging and harder to control.

25. CBI, Everyone's business tracker (2018)

26. Hiscox, Cyber readiness report (2019)

27. CBI, Everyone's business tracker (2018)

What could my business do?



Build robust cyber security practices for general IT systems and AI products.

With data underpinning AI, building privacy and security by design is essential. Introduce good data practices: make sure there are robust and audited data access protocols in place, outlining who can access data and under which circumstances; deidentify data where possible.

Firms must get the fundamentals right. Government resources to improve cyber security, such as the National Cyber Security Centre's (NCSC) [board toolkit](#) and [10 steps to cyber security](#), are important starting points. Government's [Code of Practice for Consumer IoT Security](#) contains guidance for businesses involved in the development, manufacture, and sale of consumer Internet of Things (IoT) products.



Increase consumer trust on data by making your privacy policy accessible to consumers

Data protection is about more than a legalistic privacy policy. As data use becomes even more global and complex, it's incumbent on a business to be proactive in explaining to customers their data rights, how their data is being used, whether and how it can be erased, and who has access to it. For example, some businesses are creating an easily-accessible web page which tells customers in everyday language about the 'Decisions we make using your data' or 'How to delete your data'.

Step 5: Put trust in the hands of customers by explaining AI

What is this ethical issue?

'Explainability' is the big buzzword in the AI ethics space. How do firms explain the increasingly complex decisions that AI technologies are arriving at? Customers want to know more about how their data is being used and the way algorithms make decisions.

People are more likely to trust things they understand, so there is huge value in businesses getting this right. Clearly explaining how collecting and using data directly benefits users will empower customers.



"Businesses should aim to put trust fully in the hands of consumers."

Paul Earnshaw, Technical Director for Innovation and Research, CGI

What's the challenge?

From black boxes to highly complex algorithms, businesses are grappling with how to identify, understand, and communicate the decisions made by AI in a meaningful way.

On the one hand, some businesses highlighted the potential trade-off between explainability and performance. Increasing the number of factors that influence an algorithm's recommendations or decisions might make them more accurate, but it can also make it difficult to track why an outcome takes place.

But other businesses emphasised that without explainability, it is harder to check for errors, problems, or unfair bias in AI decisions and recommendations. Explainability is worth investing in – accurately tracking an algorithm's performance could lead to better outcomes for business.



“We think that the ‘explainability’ of AI is an increasingly important issue. It already features in ethical best practice guidance and in some legislation relating to AI. For example, some legislation in the UK and EU already contains an ‘explainability’ requirement, such as in relation to data protection and antitrust law. We expect that this trend will continue and there will be more legislation relating to the ‘explainability’ of AI, particularly as authorities continue to place consumer welfare at the forefront of the regulatory agenda.

Businesses who use (or even develop) AI systems – particularly those that involve personal data – will need to understand how their AI systems work and be able to explain this in a comprehensible way.”

Simmons & Simmons

For many firms, deciding how far to balance the trade-off between explainability and performance will depend on the context. Explaining decisions and processes to customers also depends on context. Opening up a black box by explaining the technical workings of an algorithm often won't make sense for customers and can have IP implications for firms. In most cases, customers should be empowered to understand how an algorithm will return its results and the impact of an outcome.



What could my business do?



Conduct an internal audit of how personal data is being used

Make sure you know how any automated processes or AI systems are using personal data. Businesses will only be able to explain their use of data when they understand it themselves.



Map stakeholders and put yourself in their shoes to decide how to communicate on AI

Explanations will be context dependent. For example, AI that filters spam in a mailing system does not need a high degree of explainability. But AI that identifies and sorts potential job candidates does, so that both a business and candidates are reassured there is no unfair bias and so that a business can feed back on a candidate's performance.

Explanations should also be tailored for the audience. In the case of financial fraud, an explanation for an auditor should be very different to an explanation for an accountholder.

Case Study: Materials Processing Institute (MPI)

The Materials Processing Institute helps clients understand AI by comparing traditional and AI-generated outcomes

The Materials Processing Institute develops new technologies including AI for companies working in advanced materials, low carbon energy, and the circular economy. The Institute shows customers the outcomes generated by AI, comparing them to the outcomes that would have been seen without use of AI. The Institute has found that an outcome-driven process which demonstrates reliability and measures outcomes against a recognisable set of metrics drives trust and understanding of the technology among clients.



Inform people when decisions are being taken by algorithms

Make customers aware of AI systems as early as possible in the decision-making process. Take steps to outline and explain the factors that will impact on a decision and make it clear how those decisions can be queried.

Explainability can come into conflict with a business's proprietary information. Rather than opening up an algorithm, organisations should be able to demonstrate due process in these instances – reassuring customers that rigorous procedures are in place and that the right steps have been followed to produce outcomes.



“Where AI is used in decision making, it must be made clear that it is the mechanism by which a decision has been made. Those affected by the decision must also be aware of how to appeal – particularly because people tend to defer to a machine’s decisions.”

Sam Rowe, Research and Policy Executive, Yoti Ltd



Highlight to customers when they are interacting with AI

Make clear at point of use when AI is being used, even if it is not making significant decisions – for example when customers are talking with a chatbot rather than with a person.

What the law says: GDPR and explainability

It is lawful to inform customers when certain decisions are being taken by algorithms. The Information Commissioner’s Office (ICO) [Project ExplAIIn Interim Report](#) sets out the articles of GDPR that are relevant to AI. Many of the provisions have implications for AI’s transparency and explainability. For example, as the report states:

- Articles 13-15 give individuals the right to be informed of the existence of solely automated decision-making, meaningful information about the logic involved, and the significance and envisaged consequences for the individual.
- Recital 71 provides interpretative guidance of Article 22. It says individuals should have the right to obtain an explanation of a solely automated decision after it has been made.

The ICO is also developing an AI auditing framework, due to be released in 2020, which will assess the data protection of organisations using AI. It’s clear that explainability is already important for government and businesses – and is set to become even more so.

Produced by Khushali Dodhia and the Innovation team
To share your views on this topic or ask us a question, contact:



Khushali Dodhia
Policy Adviser
khushali.dodhia@cbi.org.uk