

C L I F F O R D

M I L L T O W N || P A R T N E R S

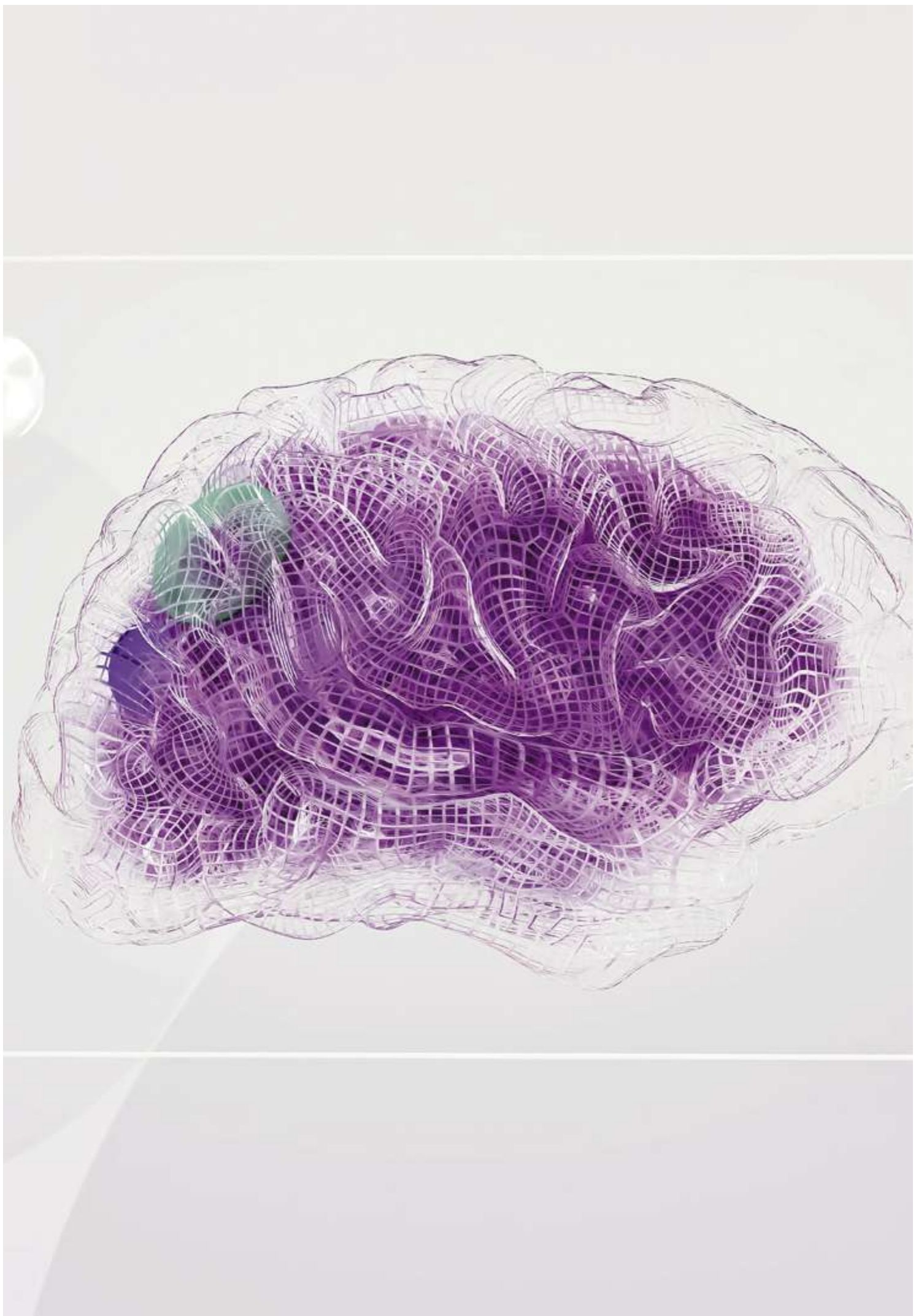
C H A N C E



RESPONSIBLE AI IN PRACTICE

PUBLIC EXPECTATIONS OF APPROACHES TO
DEVELOPING AND DEPLOYING AI

SEPTEMBER 2023



CONTENTS:

Foreword: The Changing Dynamics of AI	4
Introduction: Researching Public Attitudes as a Guide for Navigating Responsible AI	7
Executive summary: Six Insights on Responsible AI	8
Method at-a-glance	9
Research Findings and Analysis	
1. Overall Attitudes to AI	10
2. AI Policy and Regulation	16
3. Responsible AI Governance and Accountability	18
4. AI Transparency and Explainability	20
5. Consent	22
6. AI in Employment Settings	24
7. AI-Generated Content: Copyright, Intellectual Property and Plagiarism	28
8. Illegal or Harmful Content and Moderation	30
Methodology	34
Acknowledgements	36

FOREWORD: THE CHANGING DYNAMICS OF AI

A word from Clifford Chance and Milltown Partners

The release of generative AI models for use by the general public in late 2022 created a new societal and market dynamic. Consumer-facing AI products have proliferated, enabling people to interact with AI models in new and more immediate ways, and to better understand their potential. The impact and availability of generative AI has made it imperative for organisations to make AI a strategic priority, and they are developing and deploying models at pace. These pressures are being felt in all industries, as the disruptive potential of AI becomes clear.

While AI offers benefits and opportunities to businesses, people and society, its far-reaching potential means harmful impacts are also possible. To bring AI products to market, companies must therefore navigate complex legal and ethical issues. They must also understand and address the political and policy environment, which continues to develop with remarkable speed alongside the technology. The impetus to get AI strategy, governance, design and communications right for the stakeholders that organisations serve and work with – from consumers to employees, policymakers, business partners and the wider public – is stronger than ever before.

The task of guiding a company through these complex challenges falls in large part to legal, policy, product design and communications professionals.

As advisors, we are approached by clients daily with AI-related challenges – for example:

- How can I advocate for regulation that fairly reflects the responsibilities and capabilities of developers versus deployers of AI? (Policy)
- How do I generate excitement about my company's new AI product among consumers and investors, while meeting the expectations of policymakers and employees? (Communications)
- How can we design safeguards into our AI products that reassure consumers and earn their trust in relation to risks like bias and privacy? (Product design / IT / Legal / Compliance / Risk)
- Which laws and frameworks apply to AI use today, and what's on the horizon? How can I effectively manage legal risk and compliance in relation to AI development and/or use? How do these things affect my processes, policies, governance frameworks, contracts, notices, and wider strategy? (Legal).

We would be remiss to look solely to existing law or industry standards for the answers. These areas continue to develop as approaches to AI governance rapidly evolve across the globe and are shaped by political, societal and economic pressures. Important questions are being asked about how existing sectoral regulation applies to this novel technology, how new horizontal regulation and standards can govern frontier models, and what industry best practices should look like.

How, in this environment, can leaders and their advisors decide on the right approach for their organisations? We are continuing to focus on this question, and we hope this report drawing on our primary research can give you some valuable insight.

“The AI landscape is evolving so rapidly that identifying the right policy or communications strategy can feel like trying to hit a moving target. These findings shed light on what people expect from “responsible AI” but also the role they want businesses to play in the development of effective regulation.”

Rebecca Fitchett, Partner, Milltown Partners

“This report provides strikingly clear insights into how public perception of AI is constantly evolving. There are some regional differences in opinion, but those surveyed expect guardrails and oversight. The data shows that those surveyed want to see that vulnerable groups are protected from AI bias and prejudice and appropriate oversight frameworks to make AI a positive force are built.”

Jonathan Kewley, Partner & Co-head,
Global Tech Group, Clifford Chance





INTRODUCTION: RESEARCHING PUBLIC ATTITUDES AS A GUIDE FOR NAVIGATING RESPONSIBLE AI

What do the public really think about AI?

In 2021 we undertook research to understand what AI controls might look like and the public attitudes that underpin them. Our findings showed people did not regard AI as a top priority for regulation - instead regulating cybersecurity, data privacy, child sexual abuse online, misinformation, and tax contribution were viewed as more important.

But the world has changed since then, with both AI technology and the associated policy environment evolving. AI products are more readily available for use, and people's understanding of and expectations about the role AI should play in society are changing. Policy, law and regulation will likely be influenced by those societal expectations, in turn shaping the AI development and operating environment.

To help anticipate the pressures on policymakers and demands from stakeholders that will shape the legal and policy framework for AI, companies can build a deeper understanding of people's expectations around it. To get a clearer picture of what people think about AI, this year we spoke to focus groups of policy-informed individuals in Germany, the UK and the US to explore their perspectives on AI issues such as bias, consent, copyright, transparency, and content moderation. We've found that people's attitudes towards AI are developed, nuanced and remain optimistic about the potential of AI, but there is much more work to do to demonstrate that the right tools are available to realise the full potential of AI responsibly and safely.

No single piece of research can provide a comprehensive roadmap on how best to meet societal expectations - particularly on a topic as broad and complex as AI. Nonetheless, we hope these findings can serve as a guide, indicating the direction of travel, helping identify possible approaches to relevant issues, and pinpointing questions for deeper exploration.

"The report provides valuable insights into public perceptions about AI and offers practical analyses to help guide companies who are defining their AI strategies. Although regional and individual approaches may vary, the public appears optimistic about AI's potential and expects thoughtful collaboration among industry and the regulators on AI governance, accountability and oversight. We are excited to partner with Milltown Partners on this important research and look forward to continuing this foundational dialogue with our clients and our network."

Devika Kornbacher, Partner & Co-head, Global Tech Group, Clifford Chance

"AI will touch many aspects of our lives, facilitating our most important and trivial interactions. So research to understand public attitudes to AI and interrogate assumptions is urgent. Particularly as we collectively proceed towards a new era of AI innovation, law and policy globally. The findings contribute to sharpening the contours of the developing responsible AI dialogue, providing a platform for informed discussion on public expectations."

Herbert Swaniker, Clifford Chance

"Policy makers are driven by both public sentiment and the practical realities that impact their constituents. This study shows a vigilant public that is scrutinizing not only political decisions but corporate action in the AI space. This is already driving high-stakes law-making in Europe in the form of the EU AI Act and will ultimately influence which technologies society is willing to accept – or reject."

Dessislava Savova, Clifford Chance

"Throughout history public attitudes have offered a powerful signal for what companies and governments can do to develop and deploy technology responsibly - and to earn public trust. We've seen this with cars, the internet, nuclear power, genetic engineering, and more. The same is true for AI today, and this research contributes to tapping into and understanding that signal."

Aidan Peppin, Milltown Partners

EXECUTIVE SUMMARY:

SIX INSIGHTS ON RESPONSIBLE AI

1. There is public awareness of how AI could benefit society but also negatively impact individuals.

Many participants in our focus groups recognised that AI can help make advances in areas like health, science, environmental protection, productivity and more. But they were also wary that AI might negatively affect people's everyday lives, for example by exacerbating inequality or affecting jobs. People stated that they want AI risks mitigated so that its benefits can be realised.

This finding suggests that using the term "AI" when describing a service or product can signal innovation to users and investors, but could also raise concerns. To navigate this, companies can consider the right strategy for them by taking into account both the positive and negative connotations of the term "AI", and highlighting the steps they are taking to address AI risks.

3. Public attitudes about AI do not fall into opposing binary viewpoints.

On issues such as data use and encryption, participants views were often characterised into opposing camps of 'privacy vs national security' or 'free speech vs user protection'. However, participants' views on AI were not easily positioned into such opposing binary viewpoints.

This finding suggests that without distinct binary viewpoints defining - and potentially limiting - the debate about AI, organisations have the opportunity to educate stakeholders and the public and shape the debate in a favourable manner.

2. Public perceptions of AI are heavily influenced by perceptions of high-profile technology companies.

Participants tended to see the impact of AI through the lens of practices they associated with large, well known tech companies. For example, participants were influenced by recent debates about issues such as content moderation and online safety, and whether the safeguards put in place by these companies are sufficient.

This finding suggests that the public starts from a position of scepticism when it comes to AI. Companies that recognise this have the opportunity to highlight aspects of their products which the public may find reassuring.

4. Participants wanted information and choices about how AI is used in their lives.

Participants wanted companies to provide information about AI in ways that help them understand and make informed choices about how to interact with AI products and how AI impacts them. This was particularly the case for AI that might have a significant and immediate impact on individuals, such as AI models that assess people's eligibility for a loan.

This finding suggests that companies can consider the different levels of information their users and stakeholders expect for different AI use cases, and provide the appropriate information clearly as part of the experience of interacting with an AI product or output.

5. Growing public awareness of the near-term impacts of AI means companies face increased scrutiny.

Participants demonstrated a developed understanding of the near-term risks of AI, such as bias, transparency and accuracy, and the impact they may have on people and society. Though some participants were aware of conversations about longer-term existential risks associated with AI, most expressed less concern about them.

This finding suggests that the public are developing the capacity to critique companies' current approaches to AI with a significant amount of nuance. Therefore, efforts to earn public trust may involve organisations developing AI strategies that focus on AI's potential impact on people's lives today rather than these less-concrete existential impacts.

6. Participants expected regulators and companies to collaborate to ensure AI technologies are developed and used responsibly.

Participants felt that companies have an important role to play in realising AI's benefits, and in doing so safely. But they expressed little trust that companies by themselves will ensure AI is developed and used responsibly. Rather than slow down or pause development of AI, participants wanted companies and governments to work together to develop effective guardrails, regulation and standards effectively and swiftly.

This finding suggests that, as legal and regulatory frameworks are developed around the world, there is a window of opportunity to inform and collaborate with policymakers at a regional, national and international level.

Method at-a-glance



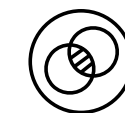
1. Identify AI topics

We reviewed academic and industry literature, spoke to experts, and drew from our own experience advising on AI strategy, to identify prominent responsible AI topics and current or potential approaches companies could take to address them.



2. Focus groups

We explored public perspectives towards the identified responsible AI topics and approaches through a series of focus groups in the UK, US and Germany.



3. Analysis

We conducted thematic analysis of the focus group transcripts to understand key themes in the views expressed by the participants.

We also gathered expert input to refine our findings and conclusions.

RESEARCH FINDINGS AND ANALYSIS

1. Overall Attitudes on AI

Across the focus groups, no dominant group of AI optimists or pessimists emerged, participants instead saw both benefits and risks of AI.

- Participants saw potential benefits from AI development and use, for example through awareness of where AI could be used to make things safer, advance science and medicine, improve services, and boost productivity.
- They expressed concerns about the possibility of AI making errors that lead to real-world harm, AI use amplifying existing bias and unfairness in society, or where AI might not have capacity for human qualities such as empathy or the ability to take unique circumstances into account.
- Participants saw benefits and concerns in parallel: even if they had a concern about a particular AI use case, they could recognise the upsides, and vice versa.
- Across the discussions, participants consistently viewed companies as the drivers of AI developments and adoption. However, there was an underlying assumption that these companies are driven by profit, so if there was a tension between responsibility and financial gain, companies would always ultimately favour the latter. This was accepted as 'what companies do' and not necessarily seen negatively, but it does provide a filter through which participants viewed 'responsible' behaviours.
- Participants typically saw AI use as being motivated by increasing efficiency or reducing costs through automation, rather than improving quality. They expressed doubt about AI products' ability to outperform human beings on certain tasks that they considered required 'human' judgements, such as evaluating a person's performance or making unbiased decisions about a job applicant.

"I think the standout one that's beneficial is [using AI] to develop new medicines. I can't really think of even a potential downside to that one."

UK focus group participant

"The overriding issue with AI is there's no empathy. You need a human overseeing it, because not everything is just mathematical."

UK focus group participant

"There are so many reasons why [an employee's] performance could not be up to scratch. Without human input to understand that or have the conversation, you could quite easily write off someone."

UK focus group participant



“ I find using AI to discover new treatments or medications very interesting. I could imagine a possible treatment might be more effective through that. ”

DE (Germany) focus group participant



“ One of my concerns is that the data AI is trained on is often biased. So, inherently, the AI is trained to be biased, right? [That] could amplify existing bias and make it even worse. ”

DE focus group participant

Participants demonstrated a good understanding of potential bias and inaccuracy in AI outputs, and were less focused on concerns related to existential risk.

- Participants understood how bias can be imported into AI through datasets. They expressed concern that it could create unfair consequences, and that unfairness might also occur if certain groups do not gain from the benefits of AI, both on an individual level (for example, in the workplace) or societal levels (for example, science and healthcare). Notably, they also discussed the role AI can play in reducing unconscious bias in human decisions.
- Many participants had first-hand experience of using AI products and finding the outputs to be inaccurate or poor-quality, such as large language models creating poorly written children’s stories or unconvincing song lyrics, or content recommendation algorithms suggesting irrelevant songs or videos to watch.
- Participants across all the focus groups expressed a generally high level of positivity towards AI, which some may consider surprising given industry and media discourse around existential risk. However, some participants did express concerns about existential risk, for example thinking about the implications of complex, misaligned AI models or potentially dangerous AI uses.

Context is crucial: participants’ attitudes towards AI depended on the specific use case, although there were some recurring areas of focus in relation to benefits and risks.

Participants identified several benefits and concerns across the six different use cases that we presented to them, as set out in the table below. While participants focused on benefits and risks specific to each use case, some common themes included:

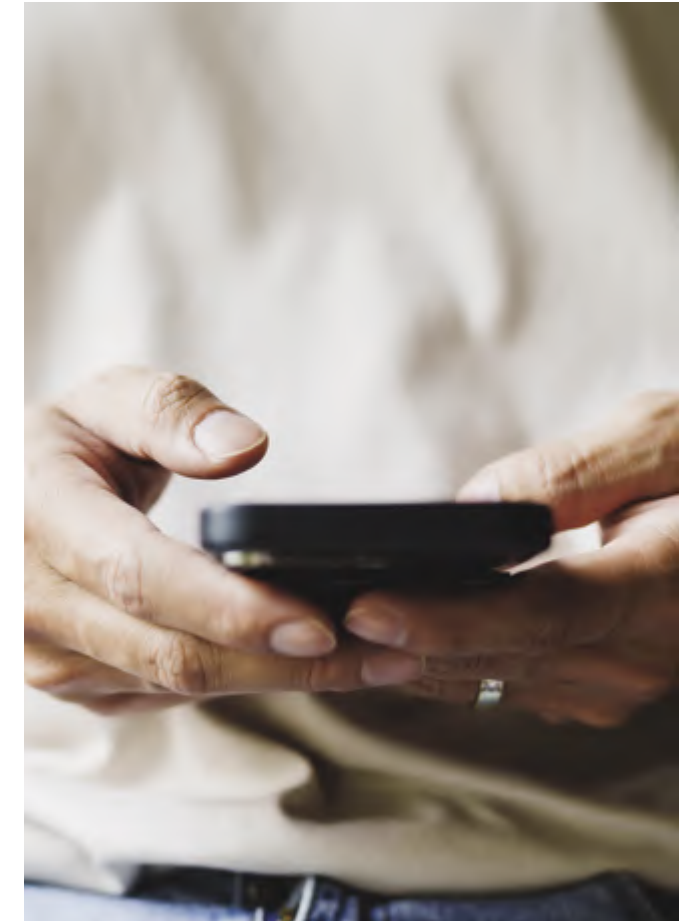
- Recognising that using large data sets and removing human subjectivity can reduce human bias or errors, while also being concerned that these could also amplify existing bias and unfairness, or lead to dangerous errors.
- Recognising the benefit of increased human productivity through AI use, while also being concerned about the potential impact of AI on the role and rights of humans in the workplace and across society.

Participants identified a range of benefits and concerns across the use cases we presented to them:

AI USE CASE	PERCEIVED BENEFITS: POTENTIAL TO...	CONCERNS: RISK OF...
Analysing applications for loans or insurance claims	<ul style="list-style-type: none"> Speed up the process of reviewing applications. Reduce human bias in loan and insurance assessments. 	<ul style="list-style-type: none"> Amplifying bias and unfairness in loan and insurance assessments. Reducing transparency of how loans and insurance decisions are made. Removing human empathy and subjectivity from the assessments and decision making.
Reviewing job applications or monitoring employee performance	<ul style="list-style-type: none"> Enable more applications to be reviewed, increasing fairness. Reduce human bias in recruitment. Increase fairness and accuracy of employee monitoring. 	<ul style="list-style-type: none"> Amplifying bias and unfairness in recruitment processes and employee monitoring. Increasing unfair surveillance of employees.
Driving assistance and autonomous driving	<ul style="list-style-type: none"> Increase safety on the roads. Enable greater inclusion and access for those unable to drive (for example, due to disability). Free up people's time to do other things instead of driving. 	<ul style="list-style-type: none"> Errors or inaccuracies in the AI leading to less safe roads. Negative impacts on jobs for taxi or lorry drivers.
Discovering new medicines and treatments	<ul style="list-style-type: none"> Improve health for everyone. Discover treatments for rare diseases. 	<ul style="list-style-type: none"> Disproportionate profit making at the expense of patient privacy or wellbeing. Sensitive personal health data being exposed.
Generating images, music, text, sounds or video	<ul style="list-style-type: none"> Increases who can access artistic and creative tools. Create opportunities for new types of art. 	<ul style="list-style-type: none"> Increased misinformation. Negative impacts on jobs in creative and other industries. Compromising copyright and ownership.
Automating workplace tasks	<ul style="list-style-type: none"> Increase workplace efficiency and productivity. Enable employees to focus on more interesting and creative work. 	<ul style="list-style-type: none"> Negative impact on jobs.

When thinking about AI, participants drew comparisons to existing societal and technology issues, like inequality or social media.

- When discussing AI, participants tended to make references to existing technology debates or social, political and economic issues. For example, they talked about the benefits and risks of social media platforms or existing data protection practices.
- They also tended to reflect on their personal experiences, or the experiences of people they know. For example, they referenced what happened when they last applied for a loan or how people may already face discrimination in the workplace, and whether the use of AI could ameliorate or exacerbate those experiences.
- Well-known leaders of tech and AI companies were often discussed in the focus groups, suggesting that certain influential voices will play a role in shaping public perceptions around AI.



ANALYSIS

The findings suggest that current understandings of AI are shaped by people's perceptions of risks that could directly impact people's daily lives, as well as the tangible benefits for society. Future existential risks were less pressing, more abstract and uncertain for our participants. This suggests that people may expect companies developing AI strategies to focus on topics such as accuracy, reliability, fairness, transparency, human oversight, respect for rights (such as privacy and intellectual property rights) and consideration of their workforce, rather than on AI existential risks.

Participants saw innovative technologies like AI in the context of more familiar digital technologies like social media or modern computer programmes in general. This opens up a field of comparisons and metaphors that companies can use to help people understand AI, and suggests that AI products are more likely to be understood and accepted if the surrounding communications draw on past precedent and principles - for example, by describing generative AI as a powerful tool that creatives can use, like photo editing or digital sound mixing, or positioning AI loan decision systems as an improvement on existing algorithmic loan assessments.

2. AI Policy and Regulation

Participants want appropriate guardrails and standards for AI, and they expect government and industry to work together to develop them.

- Across all the focus groups, participants thought that more effective governance and guardrails for AI were ‘essential’, and that appropriate guidance, standards and rules around AI will be crucial in ensuring that AI is safe and trusted, allowing its benefits to be realised.
- Participants felt that, as AI is technically complex and fast-moving, governments might not have the knowledge or expertise necessary to develop effective regulation, standards and/or guidance. There was broad sympathy for the challenges governments and regulators face around regulating AI, compared with more well-understood sectors like energy or finance.
- Most participants expected technology companies to provide input to governments to ensure that appropriate regulatory frameworks were developed. Participants felt companies should focus on helping to bring policymakers up-to-speed.
- However, many participants were wary of companies interfering with policymaking and shaping it to suit their own interests. Instead, they expected there to be a ‘balance’ in the way companies engage with governments on AI policy. Some described this as companies giving evidence, offering advice or presenting case studies to government officials during the policymaking process.

“It needs guardrails. But I don’t think that’s necessarily at odds with innovation. Guardrails prevent not just bad language but deepfakes and things that can really harm society. I think there can be a lot of innovation with guardrails.”

US focus group participant

Rather than pause or slow down AI development and use, overall participants wanted guardrails for AI to be developed more quickly and align globally.

- Participants wanted companies to make the most of AI. While some thought AI companies should wait for appropriate guardrails to be in place before pressing ahead with the development and deployment of new AI technologies, particularly for higher-risk use cases, most felt that the process to develop appropriate guardrails should speed up.
- Some participants felt that policymaking processes are much too slow. They expressed a desire for cooperation between industry, researchers and governments to accelerate the development of suitable standards and regulation, to better keep pace with the evolving risks posed by AI so that benefits could be realised sooner.

“I don’t think AI companies should slow down. I think the regulations should speed up.”

UK focus group participant

Participants’ views suggest that the values underpinning AI legislation would be an important factor in how much that legislative approach earns public trust.

- Most participants felt that the jurisdictions leading on developing AI regulations - and the national values that would underpin that legislation - were important. Some felt that ‘democratic values’ should drive AI regulation for it to be trusted. This view was consistent across UK, US and German focus group participants.
- Some participants in the US groups feared that their government was lagging behind others, such as the EU and China, on AI policy.

“What’s decisive for me is who stands behind regulation? If, let’s say, a democratic country uses AI and rigorously tested it, I’m in favour of regulation.”

DE focus group participant

“I’m in favour of the benefits of AI and using them so that humanity can live better, but if figures like Altman or Musk themselves state that we should wait and regulate this, we should probably listen to them.”

DE focus group participant

“ I would say that experts have a duty to share their knowledge and comment on topics where they know more than the lay person or the people making policies. ”

UK focus group participant



ANALYSIS

Regulation of AI is multi-layered. Existing laws such as privacy, competition, product safety, intellectual property, civil rights, employment and cyber rules already apply. AI-specific legislation globally is emerging with forthcoming rules such as the EU’s AI Act. Regulatory bodies are increasing their activity (the CMA and ICO in the UK, and the US FTC and Department of Commerce being examples). The compliance requirements will vary between sectors and will depend on an organisation’s position in the AI value chain.

However, the research findings also suggest that, although users have certain expectations about what companies should do to demonstrate responsible AI development or deployment, there isn’t yet a firm consensus on what this looks like in practice, and the degree to which that is reflected in regulation. While AI-specific regulatory frameworks are developing, organisations have an opportunity to inform and collaborate with policymakers, and can establish a position of leadership in constructively working to earn public trust in the wider AI ecosystem.

3. Responsible AI Governance and Accountability

Participants generally liked the idea of AI ethics or oversight boards, but they hold strong concerns about their independence and efficacy.

- Participants thought that ethics boards would provide useful broader perspectives on responsible AI issues by drawing on a range of experts, and felt that would offer more diverse and effective oversight of AI than relying solely on company staff.
- However, participants questioned whether an ethics board would simply be another layer of bureaucracy, and if it would actually have any power to effect change.
- Across all the focus groups, participants expressed views that an ethics board would have to be independent in order to be effective. When asked what ‘independence’ would look like in practice, they usually meant that the board members would not be compromised by benefiting from the company’s profits.

“The phrase ‘independent regulator or board’ gave me a little more confidence that this would be the right thing to do.”

US focus group participant

“It has to be an independent board because the board of a company is going to be bottom-line and shareholder-oriented.”

US focus group participant

Participants saw responsible AI principles frameworks as a baseline that companies should have by default, though they would need to be complemented by other governance mechanisms to ensure accountability.

- Across all the focus groups participants expressed support for, and positive views on, responsible AI principles. They said that these would offer employees and executives a clear framework to guide decisions related to AI development and use, and that if made publicly available it would provide a way for people to hold the company accountable to its pledges.
- Some participants questioned who would determine what the right principles were, and said that cross-industry or even international consensus or standards would strengthen the legitimacy of any principles a company adhered to.
- Many participants noted that principles frameworks were a basic step that all companies developing or using AI should have, but that they are only the first step, and should be supplemented by additional governance mechanisms, such as aboard or appointed individual.

“As a user, I’d very much prefer the option to see the values and the principles promised by the company. If I’m able to see whether they act against them or not, I can make the decision myself.”

DE focus group participant

Participants had mixed views on companies appointing an individual to be ultimately responsible for AI governance. Where some liked the clarity and certainty of a single accountable individual, others worried it would be too much responsibility for one person.

- Some participants likened the appointment of an individual to oversee responsible AI governance to existing roles such as a Data Protection Officer or a Money Laundering Reporting Officer, and felt this would offer a suitable parallel.
- Most participants were concerned that a single individual would not offer a diverse perspective on the issues AI could raise, and that they may act in their own interests or not appropriately reflect the interests of the wider company.
- Some participants raised a concern that some irresponsible companies could use an appointed individual as a ‘scapegoat’ that they could blame any failures on, rather than as a genuine official charged with holding the company to high standards.

“A board of directors or some team, not an individual. It’s got to be a group who actually has these policies together.”

US focus group participant

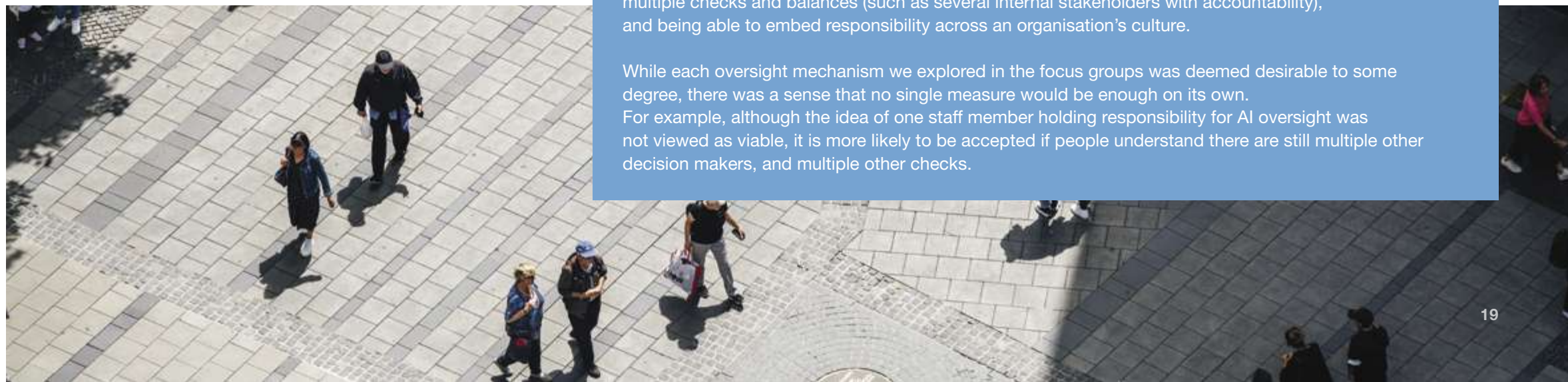
Participants saw rigorous testing of AI products ahead of public release as important mitigating potential risks and demonstrating diligence.

- Some participants recognised that real world environments provide an ideal way to test and develop AI models. But most participants expressed concerns that publicly releasing an AI product still in development or before sufficient testing for potential harms could have far reaching consequences.
- This was compounded by a concern that AI products might be misused, and it would then be too late to address any harms arising as a consequence.
- Participants felt that the responsibility to test or assess the risks of an AI product should sit with the company developing or deploying it.

ANALYSIS

The research reveals a set of strong and consistent views about the characteristics the participants felt AI governance should exhibit if it is to be effective: independence (particularly financial independence), multiple checks and balances (such as several internal stakeholders with accountability), and being able to embed responsibility across an organisation’s culture.

While each oversight mechanism we explored in the focus groups was deemed desirable to some degree, there was a sense that no single measure would be enough on its own. For example, although the idea of one staff member holding responsibility for AI oversight was not viewed as viable, it is more likely to be accepted if people understand there are still multiple other decision makers, and multiple other checks.



4. AI Transparency and Explainability

Participants expected companies to be transparent and provide information that helps people understand and make informed choices about how AI impacts them.

- When shown different types of information companies could provide about their AI products and tools, participants favoured information about the criteria that guides an AI decision or output - such as what data it uses or what it prioritises in the decision making process - and information about how they can appeal a decision or output to a human.
- Some participants also expressed a desire that companies would provide information about the data used by the AI system, and that a company would indicate where AI is used, such as in chatbots or generated images. Participants did not distinguish between which types of AI they would want labelled (e.g. all AI or high-risk AI uses only).
- Participants' preferences for the method and level of transparency depended on the perceived potential for negative impacts. For AI products used in healthcare or finance, they wanted information about data use, decision-making criteria and how to make an appeal. For AI-generated content, visual labels were more important.

“I want to know if I'm interacting with an AI or an individual. With genetically modified food, you don't know if it's real or if it's not: AI's like that. You're giving me all the information that I need, and yes it's accurate, but am I talking to a human being or am I talking to AI?”

US focus group participant

“Transparency is also about what data the AI was fed so that it operates the way it does. When I read a text, I want to see the sources so I can decide whether it's a source I trust. From my perspective, this is part of media literacy and it is important for us as society to protect ourselves against just believing anything.”

DE focus group participant

Participants preferred clear, accessible information about AI, but felt that companies should also make more technical information available to experts for scrutiny.

- Participants expected any information a company provides about AI to be clear and simple enough that a non-expert could understand it without any difficulty. They were less interested in technical information being available for users or the general public.
- However, when considering the issue further, several participants said they also wanted to know that technical information would be available to experts who would review or inspect AI systems to ensure they work safely and lawfully.

“I wouldn't get the technical details at all and honestly, I neither have the time nor do I feel like “drilling myself into this”. However, if it can be conveyed to me in a relatively simple way, for example how I communicate with AI, then I'd find it very interesting.” DE focus group participant

“I like the simple language and also technical detail. Sure, there will be a lot of people that don't understand or honestly even care to read through all of it. But for those that are - and are willing to be part of scrutiny - the most amount of information you can give will be the most helpful for making informed decisions.” US focus group participant

“Peer reviews - the experts being able to review and access information - would reassure me when using AI or seeing things that have been developed by it. Knowing that it's not just one individual company but it's been peer reviewed by the experts.” UK focus group participant

“ I'd like a label for when an output has been created by AI. I think it's [a form of] transparency and, at this stage when AI is relatively new to the majority, it's good to have that straight up acknowledgement. ”

UK focus group participant

ANALYSIS

When it comes to transparency in AI, participants didn't always expect or want highly-technical information. Instead they wanted to know there's a wider system in place to ensure safety, and that they only need to engage with the key facts necessary to make the right choice for themselves. A challenge for organisations using AI is determining their role in contributing to the creation of a wider AI safety ecosystem that people feel they can trust. Companies may wish to consider the level of information users and their stakeholders expect for different AI use cases, and consider policymaker guidance and wider industry standards in defining the technical information that might be provided by experts who design, deploy and manage AI.

That said, AI offers novel challenges for transparency and explainability that don't have parallels with other technologies. The research suggests the following prioritisation of information companies may wish to consider when designing explainable AI approaches (with importance varying by market, use case, and strategy):

1. Transparency about where and when an AI model is being used, such as labelling AI products and outputs.
2. Transparency about the criteria an AI model uses to make a decision or produce an output.
3. Information about a person's right to appeal to the use or outcome of an AI model, and how to do so.
4. Provision of human oversight over AI products and services.
5. Non-technical, plain language explanations of how an AI model works and where it is used.

5. Consent

Participants expected consent processes to be more stringent when AI might have a serious impact, compared with more mundane, every-day use cases.

- When considering different ways AI might be used, such as in self-driving cars, making loan decisions, or generating images, participants recognised that the approach to consent should be proportionate to the use case.
- Where the consequences of an AI application are minimal for users and the use case is more everyday, like using an AI-powered smartphone assistant, participants thought that consent would be less necessary, and might be only required at initial point-of-use. However, where the use of AI has more material consequences for someone affected, like a decision about a loan, participants thought that people deserved the right to consent every time.
- Many participants noted that consent is not just about giving people the option to say yes or no, but providing the information they need to make an informed choice. At the same time, participants expressed strong support for the principle of consent and expressed suspicion that companies might appear to offer consent, but in reality hide opt outs behind a confusing or complicated user experience.
- Participants highlighted a tension between transparency, consent and a seamless experience that allows them to more easily use a service.

“I feel like it's too extreme having to consent every single time. It depends on what you are using it for. If it's a virtual assistant or directions, [you should consent] at least the first time.”

US focus group participant

“ I wouldn't want it having access to my photos. I wouldn't want it having access to client files. I wouldn't want it having access to my microphone or my camera. There are certain things that you need to ring-fence. ”

UK focus group participant

Participants expected options to consent to, or opt into, certain AI uses, especially where AI might access or use personal data.

- When we presented the participants with different options for how companies could offer options to consent to the use of AI, most initially selected offering users the option to consent either every time of the first time they (or their data) interacted with AI.
- Participants felt that providing more granular options when seeking consent for AI use was more transparent, and granted people more control over how AI affects them. Several UK and German participants stated that, where AI relies on processing personal data, they expect similar approaches to consent to those required under European data protection law.
- Some participants, however, considered that AI products could have fewer barriers to use compared with other consent-related processes they were familiar with, such as cookie notices (which some participants in Germany found frustrating). Several participants suggested that exercising the choice to use or not use an AI product was sufficient in most cases to meet their expectations regarding the degree of control they wish to have when interacting with AI products.

“I would rather opt-in every single time, no matter how mundane the form or the task, just to stop and think about where the information is going.”

US focus group participant

ANALYSIS

The issue of consent in the context of AI is highly complex, and the focus group participants' attitudes towards it were not yet fully formed. This is likely to be the case among the wider public and policymakers too.

This is compounded by the nascent regulatory and market-practice landscape around consent in relation to AI products. Our findings suggest that ultimately people may think that - in addition to existing circumstances where a lawful basis like consent is needed to process personal data - the right approach is for consent to be required proportionately to how significant an impact an AI product will have on a person using or affected by it.

There is space to explore what 'consent' means in practice, balancing legal and technical definitions with people's desire for control and information alongside a streamlined user experience. Companies may wish to consider the balance (according to the use case) between giving people choices about AI and providing transparent information they need to make a choice, and not burdening users or putting unnecessary barriers in place to them receiving the benefits of AI.

6. AI in Employment Settings

Participants were fairly comfortable with the idea of using AI to improve productivity or keep employees safe, for example by using AI to monitor for hazards. However, they were wary of where AI might be used in monitoring or assessing employees.

- Participants felt that the introduction of new technologies was inevitable and accepted that it was part of a natural effort to improve productivity and efficiency. However, they had strong concerns around the introduction of AI in the workplace.
- In particular, participants expressed views that existing attempts to monitor performance can be unfair. They worried that AI might amplify rather than correct that unfairness, especially if it made errors or replaced human subjectivity and empathy from assessment processes.

“If the AI worked like spellcheck, and it said to the user ‘Have you thought of doing this?’ rather than reporting upwards, then that is very useful.”

UK focus group participant

I don’t like AI performance monitoring: don’t hire people you can’t trust to work honestly.”

US focus group participant

“On the one hand AI is probably much better than humans in evaluating humans. On the other hand, humanity somehow gets lost. I’m not sure whether AI can properly assess the human aspect.”

DE focus group participant.

“Monitoring employees’ performance using AI is an absolute dystopia.”

DE focus group participant.

Participants expected humans to be involved in AI decision-making around recruitment and employee management, and that workers should have the ability to exercise some control over its use.

- Participants recognised the potential for AI in reviewing job applications or assessing performance - to enable more candidates to be considered, or to reduce human unconscious bias - but they expected humans to be involved in any final decision-making.
- Participants frequently used language focused on “fairness”, both in the sense that AI might mean certain employees are advantaged or disadvantaged compared with others, and in expressing a view that AI should be used in ways that reflect reasonable and good treatment of employees by employers. There were concerns across all the focus groups about bias and whether workplace applications of AI may negatively impact already disadvantaged groups.
- Where AI tools might be used to monitor employee’s performance or to improve productivity, many participants liked the idea that employees could have some say over where and how it was used, for example by ‘switching off’ AI tools when they did not want to be monitored, or where they wanted to complete tasks manually. However, some said that this might mean employees who do not use AI as much might be unfairly disadvantaged. Some expressed concern that the option to ‘switch off’ AI at work might cause mistrust in employees who took the option.



“ Monitoring of employees in any sense - human or otherwise - would be detrimental to wellbeing, mentally and physically. ”

UK focus group participant

“I believe the more AI is used, the more important the human aspect gets.”

DE focus group participant

“Having the humans make the final decision - just to oversee that all the tasks done by AI are as efficient as possible - is definitely the key.”

US focus group participant

Although there is general concern about AI replacing jobs, participants also expressed more nuanced concerns about how AI might affect the quality and nature of people’s jobs.

- Participants felt that AI-based systems would be unable to consider context, and that a lack of empathy risks diminishing the experience of work environments for employees.
- Some participants expressed concerns around learning and development if overreliance on AI products steers people away from mistakes. While some participants recognised that AI could play a role in flagging mistakes, others said that AI could prevent people from failing and therefore learning from their errors.
- Many participants said that soft skills - such as helping a team gel together and work together effectively - might not be skills an AI could watch for and reward and could be driven from the workplace.

“A lot of the good stuff comes from the bad. So what you wouldn't want to do is prevent [the possibility of] a meeting where people brain-dump ideas, but something rubbish then evolves into something brilliant.”

UK focus group participant

ANALYSIS

Participants understood why employers might want to use AI tools, and can see its benefits for productivity and efficiency. But they are wary of anything that might diminish their rights as employees or the quality and experience of their job. Employers that focus on ensuring AI increases (or at least maintains) fairness and can demonstrate advantages of AI use both for the organisation and its staff are more likely to secure employee support for AI projects that affect working practices or HR-related processes.

The idea that some AI applications could be 'switched off' will be a challenge where AI is already used for legal and commercial reasons in areas such as data loss prevention and user activity monitoring. There is an opportunity for employers to engage with employees to understand which AI uses they would like to have more control over and balance this against the objectives of particular AI use cases and capabilities of the relevant AI system.

The participants' views suggest that concerns about job losses caused by AI are not currently an urgent issue for some people, but that this could change quickly if widespread and rapid job disruption does occur.



“ There is relatively little danger that AI will replace the person or human work itself. AI inherently has more of a supporting and balancing factor for bothersome, repetitive work or processes. ”

DE focus group participant



7. AI-Generated Content: Copyright, Intellectual Property and Plagiarism

When considering AI-generated content and copyright, participants expressed strong views about the importance of protecting creators' works and their intellectual property, but had different opinions on how that should be achieved.

- There were mixed views on how companies should manage issues related to AI-generated content and copyright. Some thought that text or images should only be used in AI models if the artist had given permission, others thought that was impractical.
- Some participants thought that an artist should have the right to claim ownership over AI-generated content that mimics their own work. Others thought that put too much pressure on artists to protect their own intellectual property.
- However, across all the focus groups we consistently heard that, whatever approach companies take, it should respect creators' rights and property of creators - whether that is a creator using AI in their own work, or a creator whose work is used in AI.
- Participants frequently drew parallels to the way they understood existing copyright laws to work in protecting creators' intellectual property, and felt that the copyright rules should broadly be applied to AI-generated content too.
- However, they recognised that the capabilities of AI mean some changes will be needed, for example protections that prevent a singer's voice being mimicked in an AI-generated song or an actor's face to be used in an AI-generated movie without their permission. Participants felt that creators should have the ability to protect their personal image and brand in a world of AI-generated content.

"The artist needs to get credit, they probably need to license, they probably need to approve, and then the existing standard should go for plagiarising: if AI sounds too similar you could be sued by that artist."

US focus group participant

"How can we evaluate and classify the percentage of AI content that is original or plagiarised, and how can we protect living artists?"

DE focus group participant

“ Credit where credit's due. When someone has spent their life's work creating something, AI should definitely require permission. AI can't just come in and in 10 minutes confuse it all. I think that undermines the original piece. ”

UK focus group participant

Many participants felt that art and creativity is uniquely human. They did not feel the same level of connection to AI-generated content compared with content created by a human.

- Participants understood that AI-generated content is based on an AI model being trained on existing images, music, text or video. Some felt this was not very different to human artists, who learn their craft by studying others and take inspiration from existing works to create their own. Others commented that if AI-generated content appeared to be an exact copy or a close imitation of an artist's work or style, they would consider this akin to plagiarism.
- For AI-generated content to be original, some participants thought it would have to offer or create something that was noticeably distinct from what had been created before.
- Overall, however, participants thought that AI-generated content would not have the qualities of human empathy, experience or fallibility, nor the creative intentions to produce genuinely original and creative work.
- Participants' reactions to AI-generated art suggest that they think recent examples of AI used to create songs or podcast episodes are a gimmick that will not replace the connection people feel with human artists.

"We need to think about whether we want to maintain art as what it is and as what it is intended to be - a political medium - or whether we want to step on the path where it is purely about aesthetics."

DE focus group participant

"We think about art as being creative and being a very human process and being imbued with human emotions. And I don't know how I'd feel about listening to a piece of music that was completely AI-generated."

UK focus group participant

"It's basically just some kind of photomontage or something like that. AI helps with existing things, but it's nothing really new. I believe that creativity and artificial intelligence exclude each other anyway. It can't really make something up."

DE focus group participant

ANALYSIS

Should artists and other content providers maintain control of their own creations, or at least get compensated where their work has been used to train AI? The general answer from respondents is "yes". This creates complexities for policymakers, who need to balance the interests of the creative industries against the risk of limiting new applications of AI that could be economically and societally beneficial.

Participants saw AI as a tool creatives can use to create new works, and though it may offer new opportunities, people still want to connect with the human behind the art. Without that human element, it is not clear that participants wanted copyright to extend to works solely created by AI.

Our research also suggests that there is strong public support for creators who want to protect their output and public figures who want to protect their image rights. This could have wide-ranging implications both for access to training data, the development of AI models themselves, control over generative AI outputs, and the speed of deployment of new systems.

8. Illegal or Harmful Content and Content Moderation

Participants expected companies that make or apply AI (particularly generative AI) to take appropriate steps to limit the potential for illegal or harmful content.

- Participants consistently expressed concern about the fact that AI could be used to create harmful, illegal or otherwise inappropriate content, such as hate speech.
- Participants accepted that, where the definition of 'harmful content' was concrete, companies had a responsibility to prevent that content from being generated or take action to remove it if created. When asked what kinds of content fell into this category, participants often referenced types of content that already have relevant laws preventing their creation and distribution in many jurisdictions, such as child sexual abuse material, banned symbolism, and incitement of criminal activity.
- Participants were split on how companies should address 'harmful content', the definition of which they considered to be more subjective, citing misinformation as one area where definitions of 'harm' can differ. Some felt that companies should take proactive steps to prevent an AI product from generating harmful content, for example by building in guardrails that would prevent certain outputs. However, other participants felt this would not be effective – noting that there would always be workarounds.


“If it's possible to post whatever kinds of beheading videos, etc, then of course the person who uploaded it is responsible, but [the social media platform] is liable at the same time.”

UK focus group participant

“I think the company has to take the ultimate responsibility. It's too easy to say it's the individual user who's created this horrible cartoon or something. The platform has to take some responsibility.”

UK focus group participant

- Many participants were wary of the idea that, if companies placed guardrails on AI-generated content, this might limit people's freedom of expression or censor certain content based on the company's own political or ideological views. Some of these participants preferred the idea that users should be free to post whatever content they wanted (unless it was illegal) and other users should be able to flag or report any content they found inappropriate. Participants in the German groups described this in terms of censorship or propaganda, whereas US participants expressed this through concerns about companies' ability to correctly balance content moderation with freedom of expression.
- Where it was suggested that companies should rely on users to report harmful content, there was some concern that the most vulnerable in society do not have the voice to "sound the alarm" or draw public scrutiny, rendering reporting an ineffective solution.



“ By putting the block into the tool itself, hopefully you can stop people misusing it at the source. ”

UK focus group participant

When thinking about how companies might mitigate harm from AI-generated content, participants drew comparisons to content moderation on social media platforms, referencing what has worked in the past and where they perceive companies to have got it wrong.

- When thinking about what companies should do about AI-generated content that might be harmful, participants frequently drew comparisons to the spread of misinformation, hate speech and other harmful content on social media and online more generally. For some who made this comparison, AI-generated content did not offer any new concerns – companies already deal with harmful content and should apply those approaches to AI.
- However, many participants recognised that AI might potentially increase both the scale at which harmful content is generated, and the severity of the harm it might create, for example because it can create incredibly convincing or authentic-looking fake images.
- Participants recognised that addressing harmful AI-generated content is a serious challenge. Many acknowledged there was no easy answer. Based on their perceptions of how well social media companies have already dealt with content moderation, few thought that tech companies should tackle this issue by themselves, nor did they trust them to get it right if they did.

“[Social media platform] now has user reporting and you can report harmful content, and they take their good time and investigate it, and it may work, and it may not.”

US focus group participant

ANALYSIS

Issues relating to illegal or harmful online content and content moderation are currently in the regulatory spotlight – for example in the EU's Digital Services Act, the UK's Online Safety Bill and the USA's Kids Online Safety Act. Requirements for controls around certain content generation also feature in a number of AI-focused bills and laws.

Participants in our research drew comparisons between mitigating harmful content generated by AI and existing content moderation practices on social media platforms and other digital technologies. They expressed concerns that AI-generated content amplifies existing challenges for content moderation, particularly the scale and speed at which content can be generated and the increased ability for AI to blur the lines between truth and reality, such as through the creation of 'deepfake' videos.

The views expressed by the participants suggest that people believe companies have a responsibility to take active steps to remove clearly harmful content, rather than relying solely on user reports or guardrails in the AI itself.



“ Existing user reporting functionality has done little to prevent harmful content being widely shared. ”

DE focus group participant

METHODOLOGY

1. Identifying approaches to responsible AI

To start our research, we identified a range of current and potential approaches to responsible AI through the following method:

1. We reviewed academic, policy and industry literature on responsible approaches to AI, developing a longlist of responsible AI topics, such as governance, transparency, bias or content moderation.
2. For each topic, we identified a challenge companies (across multiple sectors) were facing, and outlined two or three approaches companies might take, based on existing best practices or theoretical frameworks.
3. We then engaged experts from within industry and academia to identify which topics and approaches are the most pressing for the current state of responsible AI, and which would benefit most from better understanding public perceptions on them.
4. From this we identified the following shortlist of responsible AI topics to include in our focus groups: AI and Employment | Copyright and Intellectual Property | Content Moderation | Transparency and Explainability | Consent | Corporate Governance | Regulation

2. Focus groups

To explore public perspectives towards these responsible AI topics, we conducted six 2-hour focus groups across the UK, US and Germany, with two groups of six people in each country, totalling 36 participants.

1. We recruited the participants to include people who engage in political issues and have some knowledge of AI, but who were not experts or professionals in either policy or AI.
2. In each focus group, we presented participants with stimulus materials we developed based on the responsible approaches. This included short explanations of certain AI use cases (for example AI to generate images, or used in self-driving cars), as well as short descriptions of different approaches companies could take to developing and using AI responsibly.
3. Following a set discussion guide, standardised for each group, we asked the participants to reflect on the different approaches to responsible AI and probed on what they thought would be the most responsible approach, and why.

Although the sample size was relatively small, the combination of the complex nature of AI as a topic, the extended 2-hour length of the focus groups, and the fact our participants had been selected to already have a level of engagement with the topic meant the qualitative data we gathered was remarkably rich.

3. Analysis

Using the transcripts of the focus groups' discussion and participants' written comments we applied thematic analysis to code the discussions and identify key themes, which formed the structure of the findings we have reported.

We discussed these findings with experts from industry, via a series of meetings and roundtable discussions in July and August 2023. Using this input, we refined our findings and developed the conclusions outlined at the start of this report.

Note on this research

This paper reports the focus group participants' views based on analysis of their discussions and contributions during the focus groups. The qualitative research was based on 6 focus groups across 3 markets, with a total of 36 participants. It also dealt with a wide range of topics rather than exploring each topic in depth.

The research therefore provides a valuable snapshot of public expectations on responsible AI, and directional insight on some of the challenges companies are facing.

However, the small sample and breadth of the research means readers interested in investigating public attitudes or data to inform decisions should seek context-specific data to inform decisions. The analysis is intended to provide insights into the views and opinions of focus group members and should be read in light of the limitations of an exercise of this nature.

Participants' quotes have been edited for length and clarity, but their meaning has not been altered.

Acknowledgements

This report is based on research conducted in partnership by Milltown Partners and Clifford Chance.

Authors & contributors

Milltown Partners

Aidan Peppin
Rebecca Fitchett
Matthew Burchell
Alex Harrison

Clifford Chance

Devika Kornbacher
Jonathan Kewley
Holger Lutz
Dessi Savova
Herbert Swaniker
Phillip Souta
Rita Flakoll
Don McCombie

About Clifford Chance

Clifford Chance is one of the world's pre-eminent law firms with significant depth and resources across five continents. Clifford Chance's Tech Group is a global, cross-practice team of over 600 tech-savvy lawyers with extensive industry expertise. We help clients from all sectors explore the opportunities and risks that technology brings.

<https://www.cliffordchance.com/>

About Milltown Partners

Milltown Partners is a global advisory firm working with influential organisations and individuals on the communications and public policy challenges that define their reputations. We partner with leaders in their fields: ground-breaking technology companies and their founders; global businesses and their CEOs; and prominent institutions and their leaders.

www.milltownpartners.com