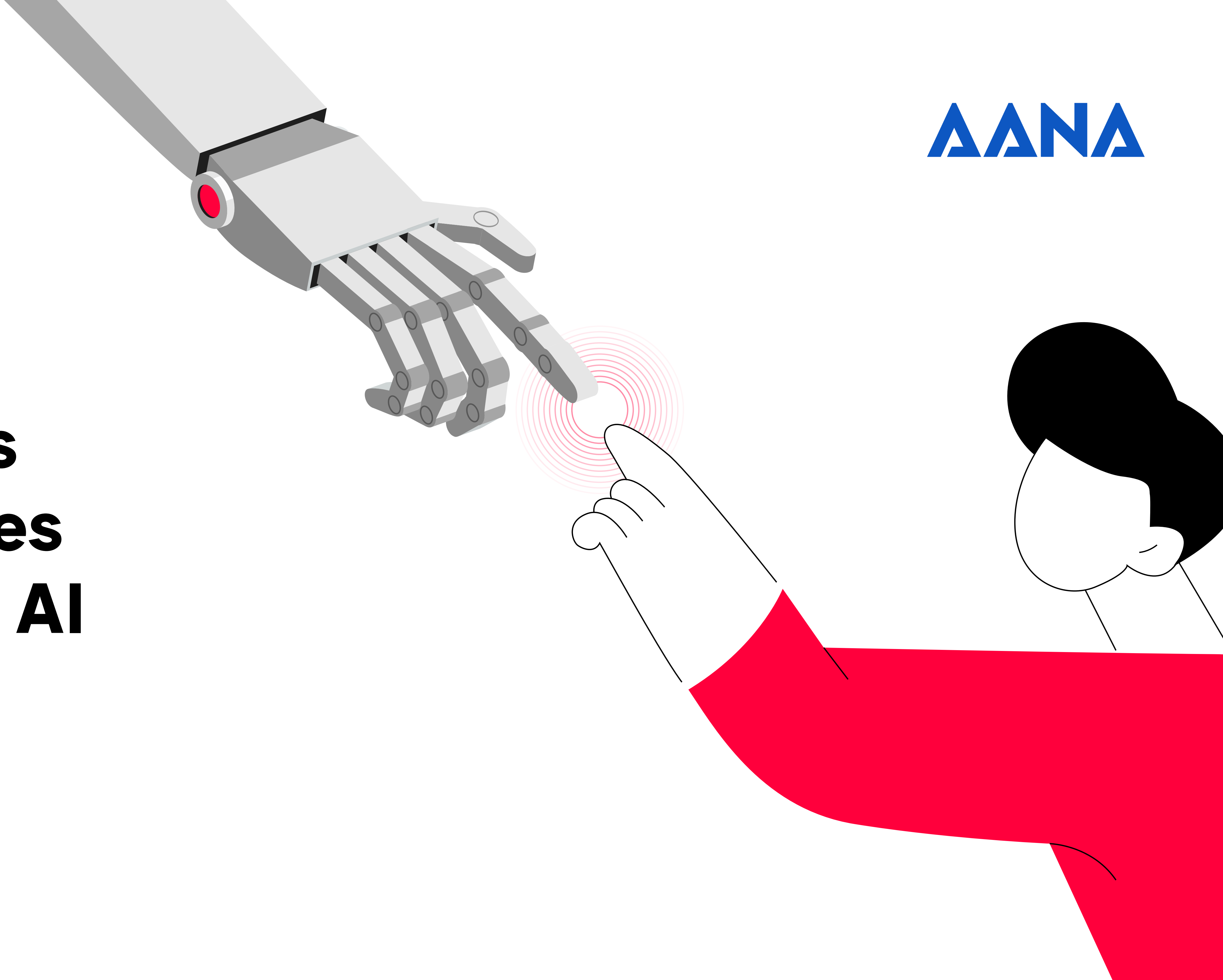
WFA World Federation of Advertisers

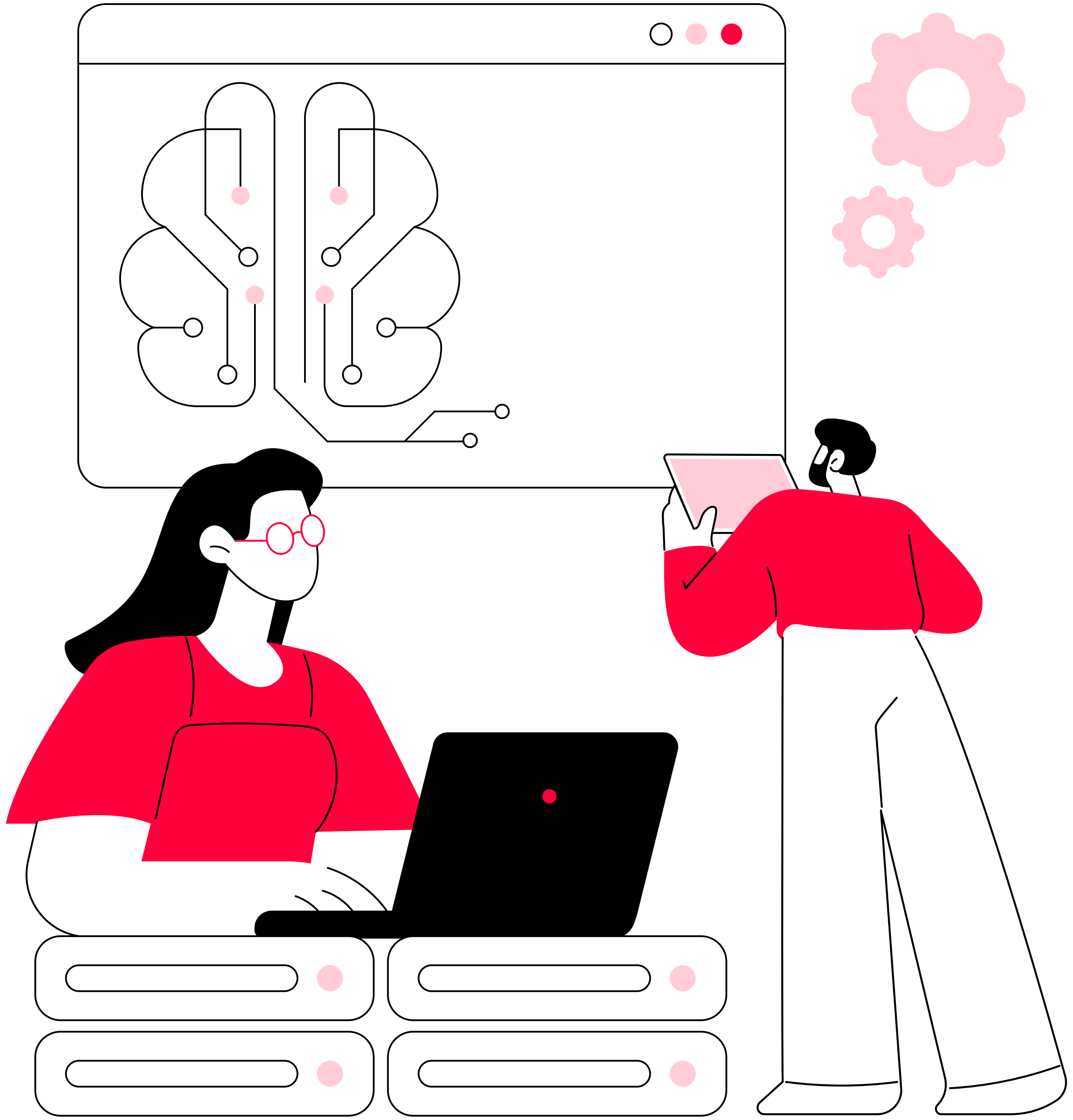
AANA

WFA Primer: opportunities and challenges in generative AI



February 2024

Contents



Executive summary

In November 2022, the public release of OpenAI's generative artificial intelligence (AI) chatbot ChatGPT resulted in the fastest adoption in human history of a consumer technology, amassing over 100 million users within the first two months.

Since then, the general availability of generative AI tools capable of writing text, composing music, creating art (and more!) has demonstrated its potential to revolutionise any industry where creativity is key.

It is no surprise therefore that three in four of the world's largest brands are already using generative AI in their marketing or are planning on doing so soon. From content ideation and creation to more personalised customer segmentation, task automation, customer service and product innovation, generative AI is poised to play a significant role in driving marketing creativity, effectiveness and efficiency.

However, generative AI's impact on marketing is yet to be fully understood and its use has already raised several legal, reputational, and ethical challenges for brands.

Concerns around data protection and intellectual property, company confidentiality, diversity, equity and inclusion and brand safety are most common, but there still remains a lack of understanding among marketers about what can be done to address these.

This primer aims to provide an overview of what generative AI is and how it impacts marketing, what the considerations of its use are for brands, and what initial steps could be taken to mitigate risks.

We believe this primer is crucial in equipping brands with the knowledge they need to make more informed decisions and harness the potential of generative AI in a responsible, ethical and sustainable way and without compromising on trust, safety and integrity.

In the coming months, WFA will continue to work with brands and industry partners to further develop these recommendations and practical solutions that can be adopted to propel safe and suitable AI use.

"The rise of generative AI has sent waves of enthusiasm across the marketing industry, heralding promises of heightened productivity, creativity and efficiency. However, amid this fervor, there is a nervousness surfacing amongst brands about the legal and ethical implications inherent in the use of generative AI, and how these could undermine brand reputation and trust.

This paper aims to drive a harmonised understanding of what generative AI in marketing means and bring to the forefront the key considerations for brands. This is a crucial step towards leveraging generative AI in a responsible way, and WFA looks forward to helping brands make the most of the exciting possibility that generative AI can offer to deliver on creative and marketing effectiveness."

Stephan Loerke
CEO, WFA

Introduction

Over the last two decades, artificial intelligence (AI) has transformed the marketing function, helping brands automate the buying and selling of ad inventory, generate marketing copy at scale, personalise customer service via chatbots and virtual personal assistants, recommend personalised products and content, analyse and predict consumer behaviour and segment audiences to target ads to the right people at the right place at the right time.

However, the advent of generative AI, often thought of as the 'next generation' or 'next frontier' of AI because of its broad utility and its perceived ability to emulate natural language conversation and content has taken the marketing industry by storm.

Its capacity to write text, compose music, create digital art, personalise content, predict future trends and elevate customer experiences is already disrupting marketing, promising opportunities for more personalised, creative, and effective marketing communications.

McKinsey estimates that generative AI could increase the productivity of the marketing function with a value up to 30% of total marketing spend.

According to WFA research, three in four of the world's largest brands are already using generative AI in their marketing or are planning on doing so soon.

And brands today are generally optimistic about the technology's potential to drive business growth, improve productivity and efficiency, reduce costs and drive creativity.

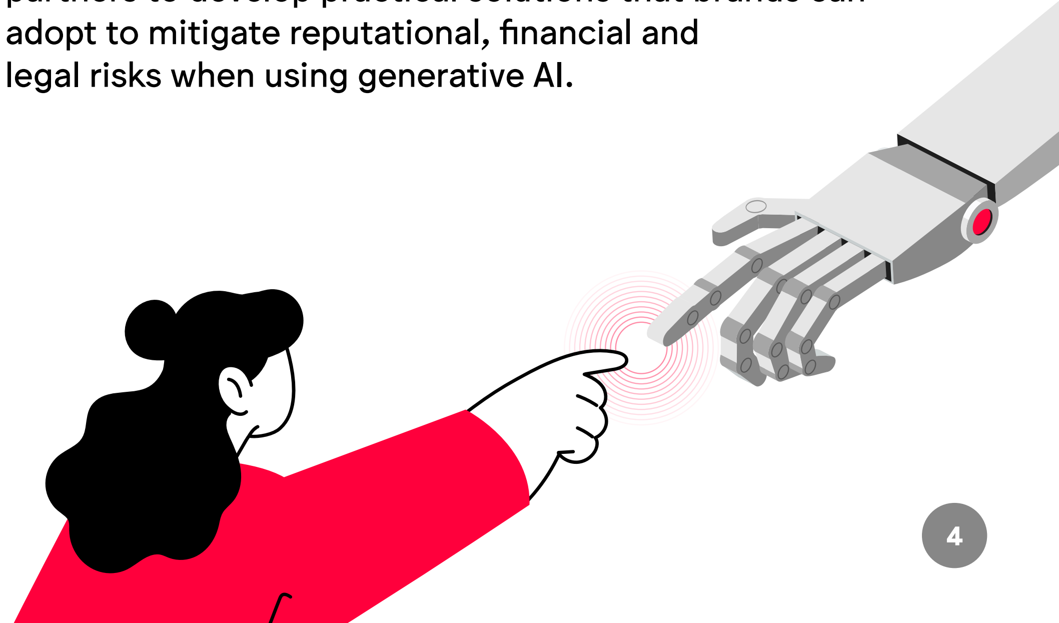
As the technology continues to evolve, generative AI is poised to play an increasingly pivotal role in marketing. However, while it has significant potential, its impact on advertising is yet to be fully understood and its use has already raised several legal, reputational, and ethical considerations for brands.

WFA research has found that concerns around company confidentiality, data protection and privacy, IP and copyright, diversity, equity and inclusion, and brand safety are most common, although broader societal considerations such as the impact of generative AI on the creative industry and the environment continue to be points of apprehension.

Despite such concerns, there is a lack of common understanding among marketers about the risks of generative AI and what can be done to mitigate them. Understanding these is essential for brands to safeguard their reputation and intellectual property, maintain legal compliance, and protect their consumers.
This primer therefore aims to provide an overview of what generative AI is and current main use cases in marketing, develop a shared framework for understanding the risks of its use and put forward some of the steps brands could take to mitigate these.

We believe this is necessary to equip brands with the knowledge they need to make more informed decisions and harness the potential of generative AI in a responsible, ethical and sustainable way.

WFA will continue to work with members and industry partners to develop practical solutions that brands can adopt to mitigate reputational, financial and legal risks when using generative AI.



A short history of AI

Traditional artificial intelligence (AI) models are generally defined as ‘pattern recognisers’, trained in such a way that they can identify trends in data sets and make decisions or predictions based on predefined rules. Such AI systems have long been used for marketing purposes, including for ad personalisation, predictive analytics and customer segmentation, product recommendations and for the automated buying, selling and placement of ad inventory.

Over the past 40 years, however, technological breakthroughs in AI development have enabled the rise of generative AI:

1980s

In the 1980s the advent of big data and machine learning marked a shift from models based on explicit rule programming to models which make predictions or decisions based on patterns they detect in data. An example includes online content recommender systems, which identify patterns in user choices and behaviours to suggest content users may be interested in.

2010s

Around 2010, the development of deep learning models revolutionised the field by enabling machine learning systems to autonomously discover complex patterns and carry out much more specialised tasks such as image recognition and classification, sentiment analysis, and object detection.

2017

In 2017, the rise of foundation models represented a transformational shift in the development of AI. Foundation models are based on emulating neural structures found in the human brain to develop synthetic content in the form of text, image, video, code or other.

November 2022

These developments culminated, in November 2022, in the release of a much more capable generative AI tool; the OpenAI chatbot ChatGPT 3.5. Its launch resulted in the fastest adoption in history of a consumer technology, amassing over 100 million users within two months.

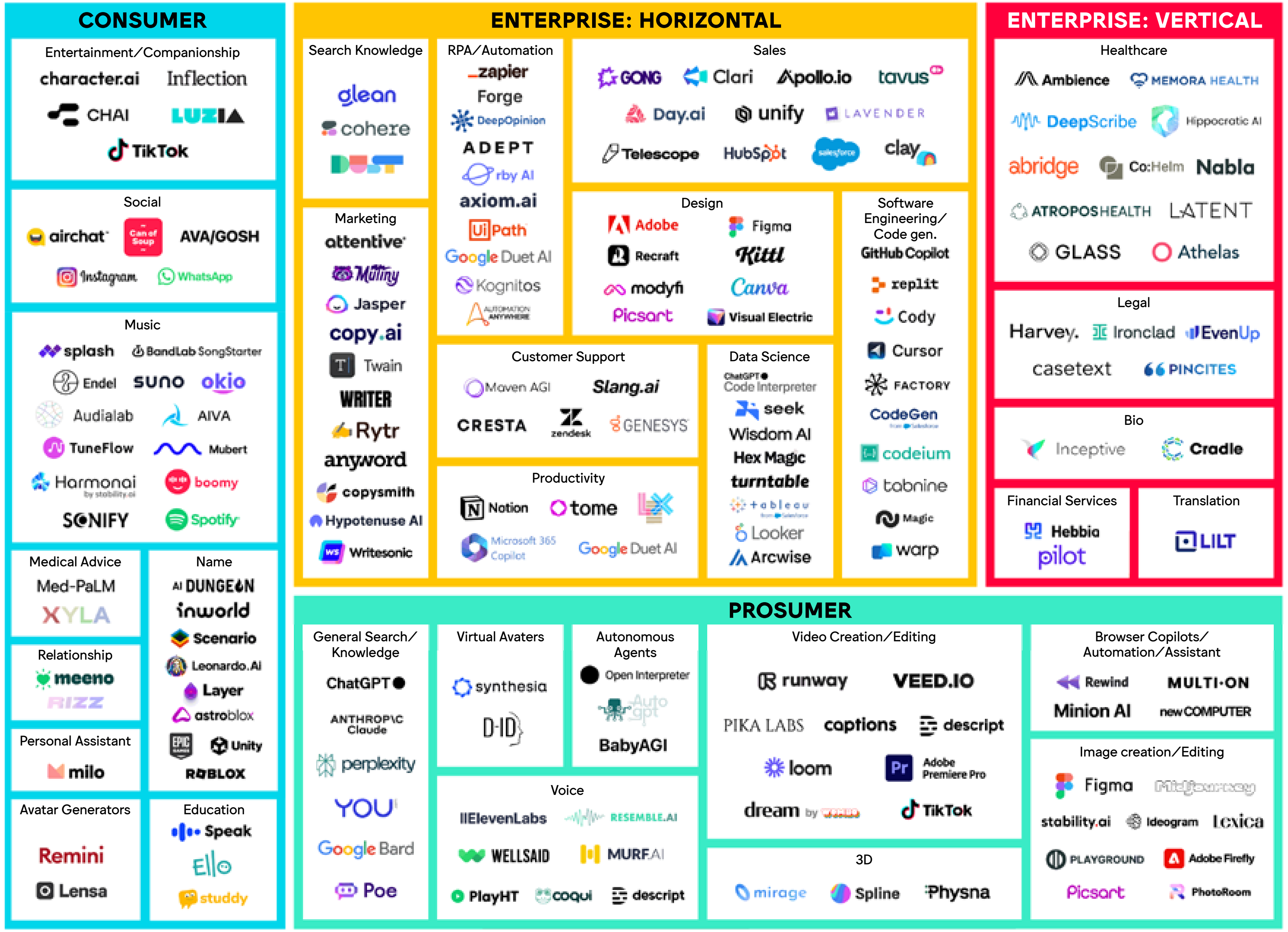
What is generative AI?

At its most basic, generative AI can be defined as a form of AI which produces unique content ("outputs") on the basis of both the patterns it has learned through training data and user prompts such as information, sentences and/or questions ("inputs"). According to predictions reported by Europol, given its wide scale adoption and public availability, within just a few years, up to 90% of content on the internet could be synthetically generated or manipulated.

Generative AI tools have captured the attention of people all around the world in a way that no other technology has not only because of their broad utility, but also because of their ability to emulate content and conversation in a way which closely resembles that produced by humans.

It is its ability to write text, compose music, create digital art, develop computing codes and more that has captured headlines. And as a result, generative AI has significant potential to revolutionise any field where creation and innovation are key, including marketing.

As demonstrated by the market map (right), there are already hundreds of generative AI tools available for use, including for marketing purposes; from text, to image, voice, video, music generators, copyediting, coding assistants, data analytics tools and more. The generative AI market is expected to grow exponentially over the coming years, and Bloomberg estimates that it will be worth $1.3 trillion by 2032, up from just $40 billion in 2022.



Generative AI in marketing

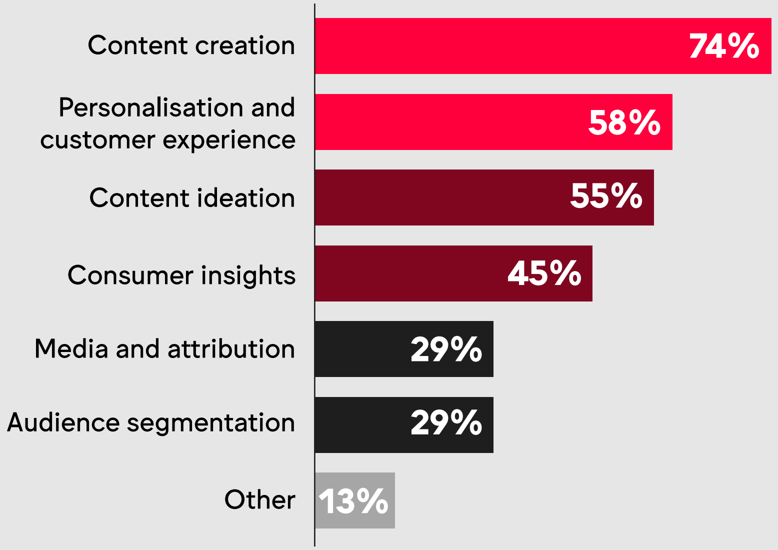
According to McKinsey, generative AI in marketing and sales is likely to represent most of the technology's impact across potential corporate use cases and could increase productivity in marketing by up to 30%.

In particular, generative AI is considered to be helpful when it comes to:

- **Content ideation & creation (marketing creative):** helping produce brand advertising, headlines, slogans, social media posts, voiceovers, sonic branding and product descriptions.
- **Customer service:** generating immediate, personalised responses to complex customer inquiries regardless of the language or location, in a way which can emulate human-like qualities such as empathy.
- **Personalised customer experience:** analysing customer data, past purchasing behaviour, consumer preferences to tailor services, products, interactions and messages to individual customer needs. This also includes providing personalised experiences like shopping assistants, generative AI-powered search engines and more.
- **Data-driven decisions and predictions:** analysing fragmented and unstructured data across different media and content to uncover correlations and break down information silos.
- **Search Engine Optimisation:** identifying relevant keywords and phrases that have high search volumes and low competition and dynamically optimise and personalise content to target keywords and content that are most likely to drive organic traffic to websites.
- **Product innovation:** using AI-driven consumer insights throughout product development process to optimise effectiveness.
- **General task automation:** streamlining the creation of marketing assets such as media plans, quarterly reviews and meeting agendas.

According to WFA research of 65 global brands, content ideation, content creation and customer experience are the most cited use cases of generative AI in marketing right now.

For what marketing purposes do you use Generative AI today?

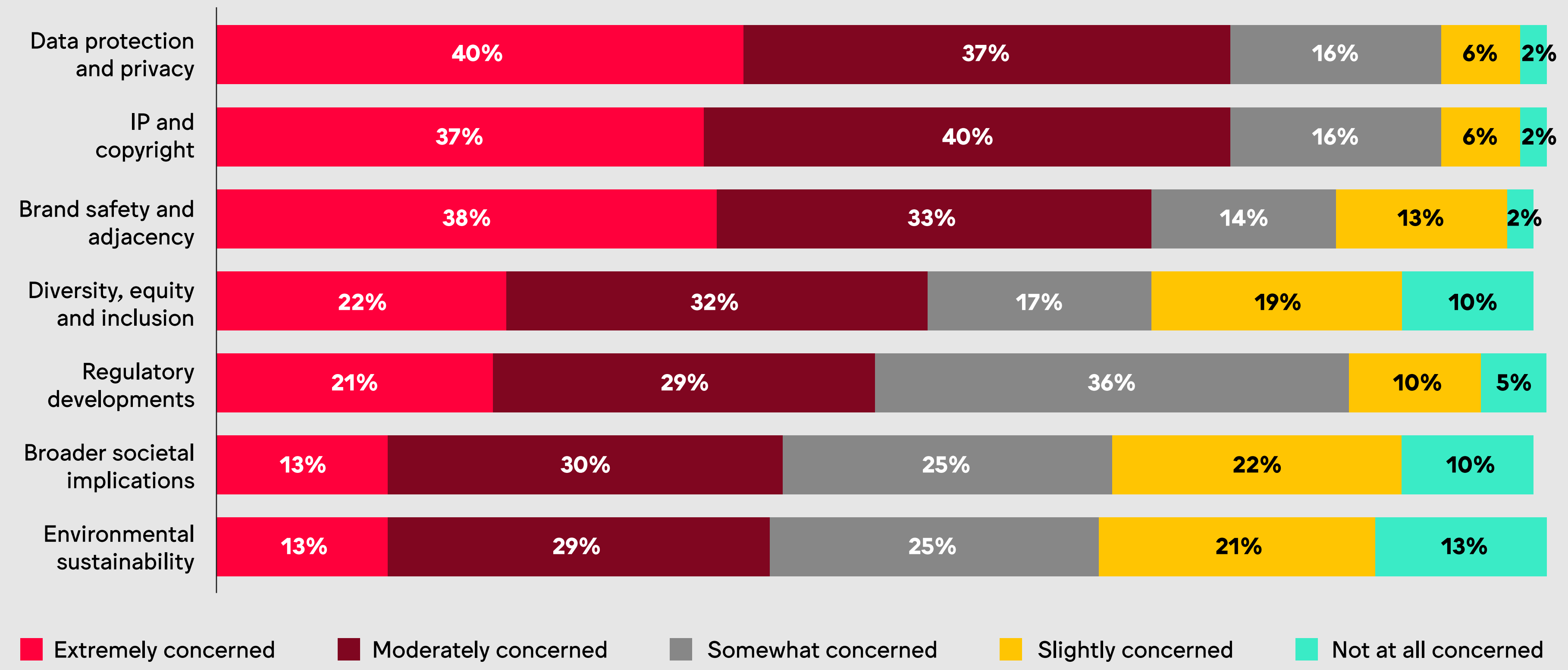


Challenges and risks of Generative AI

Although WFA research reveals that most marketers are enthusiastic about the potential of generative AI in driving business growth, they are equally concerned about the legal, ethical and reputational risks of generative AI use.

Data protection, IP and copyright and brand safety top the list of concerns for brands, with over 75% extremely concerned or moderately concerned about these. In addition, over 50% are concerned about the impact of generative AI on diversity, equity and inclusion efforts, the creative industry and environmental sustainability.

How concerned are you about the following when it comes to the use of generative AI?



On the basis of this research, WFA has identified six key priority challenges for brands:

1. **IP and copyright**
2. **Data protection and privacy**
3. **Confidentiality and security**
4. **Reliability, safety and integrity**
5. **Diversity, equity and inclusion**
6. **Broader societal considerations, including environmental impact**

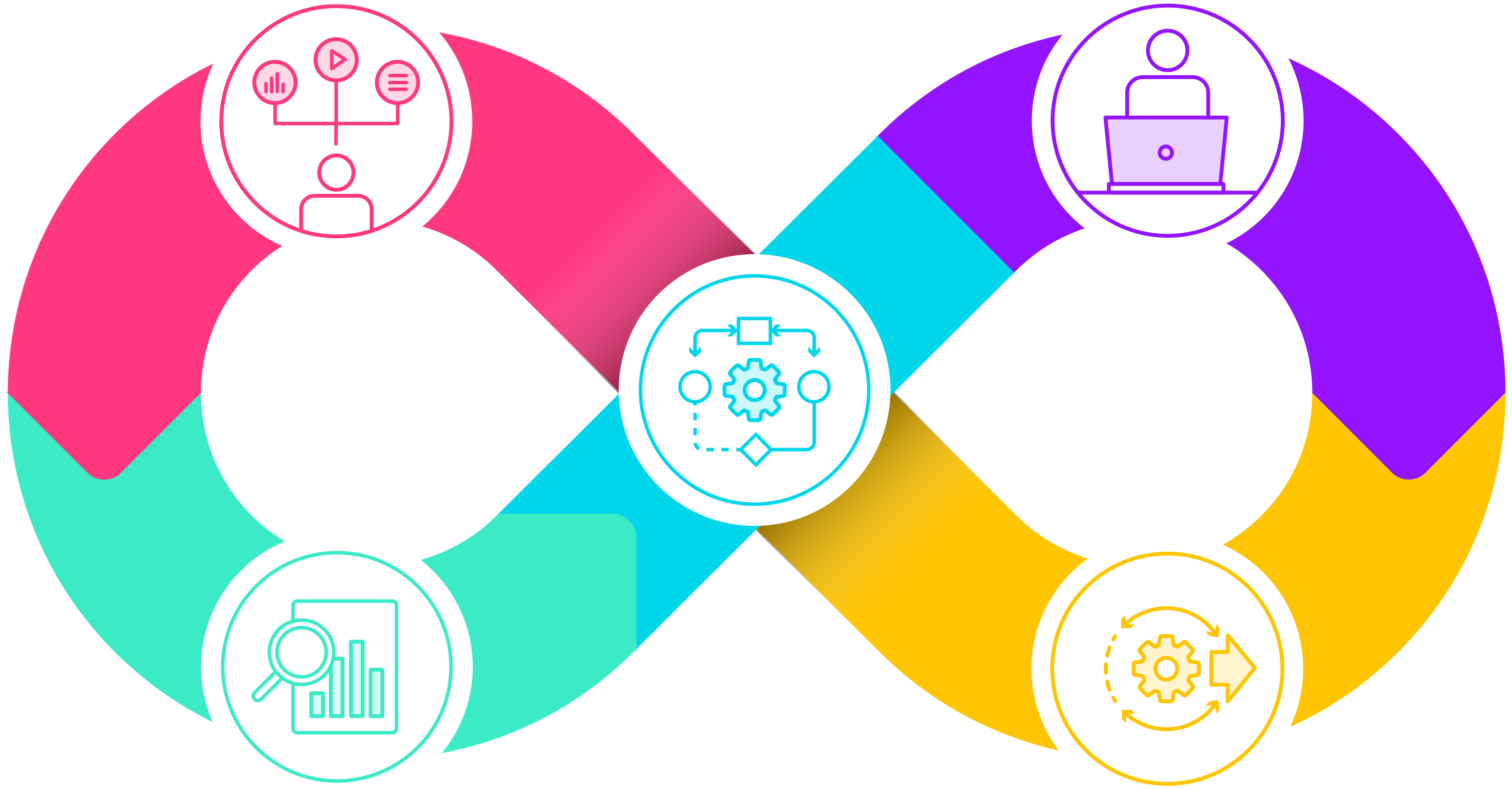
Although these risks won't be unique to generative AI and could apply to any AI and machine learning technologies, generative AI has the potential not only to exacerbate these but also bring with it novel challenges. This rests primarily on its ability to produce content in such a way that it may not be easily distinguishable from human-generated content.

The following sections of this paper will introduce a framework through which to assess these six generative AI risks and set out actions that both brands and generative AI providers could take to drive responsible AI use.

A framework for categorising generative AI risks

To categorise and understand the challenges of generative AI and respective responsibilities for addressing these, WFA has developed a framework for considering risk across the entire use cycle of a generative AI tool. This framework follows that introduced in WFA's GARM Safe & Suitable innovation guide.

The framework categorises risk into five pillars:



1 **Use case:** The risks arising from the purposes for which generative AI is being used and potential implications of such use.

2 **Information sources:** The risks emerging from the data upon which the model has been trained.

3 **Algorithmic processing:** The risks resulting from the algorithmic processing arrangements and the ways in which the model takes decisions.

4 **Protection of the end user:** The risks for the end user of prompting their information into the tool.

5 **Safety of outputs:** The risks arising from the outputs generated through the use of the tool.

Use case

Generative AI's broad utility and public accessibility has raised concern among regulators and consumers alike about the ways in which it may be used to undermine individuals' health, safety, fundamental rights and democracy and the rule of law. As a result, policymakers across the world including in the US, the EU, China, Singapore and the UK are introducing new rules aimed at regulating AI.

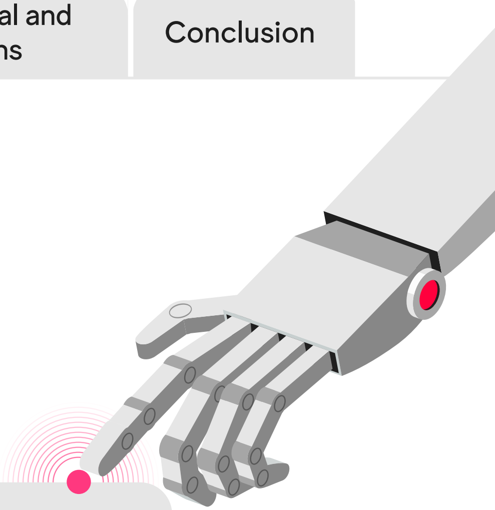
Regulators are focused in particular on the use of generative AI to infer political, religious or philosophical beliefs, sexual orientation and race, to develop facial recognition databases, to create 'deep fakes' (digital content purporting to depict a real person or scenario) for pornographic purposes and for social scoring based on social behaviour or personal characteristics. Many of these use cases are now being restricted in markets such as the EU.

While these rules don't pertain directly to marketing use cases, a complex patchwork of approaches is emerging and brands will need to assess any potential health, safety and fundamental rights implications of their use of generative AI.

For example, using generative AI to develop a chatbot to provide individuals with AI-generated medical advice, or to create content purporting to depict a real person, or to profile individuals for targeted advertising purposes will warrant closer examination and assessment. Such use cases are likely to bring higher risks compared to those aimed at dynamically optimising ads in real time or for basic customer service chatbots, for example.

More generally, brands may want to consider the safety and suitability of the selected generative AI tool to avoid reputational risks. Many of the largest generative AI providers now have community guidelines in place which prohibit the use of their systems for malicious purposes such as for reasons pertaining to hate, self-harm, sex, violence, harassment and deception. However, while some have promised to take proactive steps to prevent their tools from being abused, under most jurisdictions they have no obligation to do so, making it difficult to hold them accountable how their tools are being used to spread harm.

Recommendations: Use case



Brands could consider:

- **Developing a set of AI principles** to guide their organisation’s use of generative AI. These could be based on principles such as accountability and transparency, explainability, fairness and inclusion.
- **Defining a clear governance structure for AI** internally which allows for diverse collaboration across functions (including IT, marketing, legal and policy).
- **Creating an assessment tool** which identifies a set of acceptable and unacceptable use cases so that legal, procurement and marketing teams can assess generative AI risk depending on use case and tool.
- **Training teams** on the use of generative AI and the ways in which the technology can be misused.
- **Keeping track of global regulatory developments** to ensure legal obligations are fully understood.
- **Seeking clarity** on the measures AI providers have in place to avoid misuse and abuse of their tool, such as community guidelines, and the steps they take to ensure compliance with legal obligations.

Generative AI providers could consider:

- **Detailing how they are complying** with applicable laws, regulations and guidelines as it pertains to development and deployment of their AI systems.
- **Putting in place community guidelines** with clear rules on how their tool can and can’t be used and taking proactive measures to identify and prevent harmful practices.
- **Providing transparency** to business users about the ways in which their tools are being used and what mechanisms are in place to prevent misuse so brands can assess their safety and suitability.

Information sources

Generative AI tools are trained on vast amounts of data, often collected using web 'scrapers' or 'crawlers' - software programs that systematically navigate the internet, visiting web pages and collecting data from them. This can include data such as intellectual property and copyright protected works, personal data, and harmful content, including disinformation.

Given that outputs are generated on the basis of both the data upon which the model is trained and the data prompted by the end user, the quality of training data can have significant implications on the quality of user (brand) outputs. While this data and the way it was collected is often not of direct responsibility of the users of the tool, it can bring with it a number of risks.

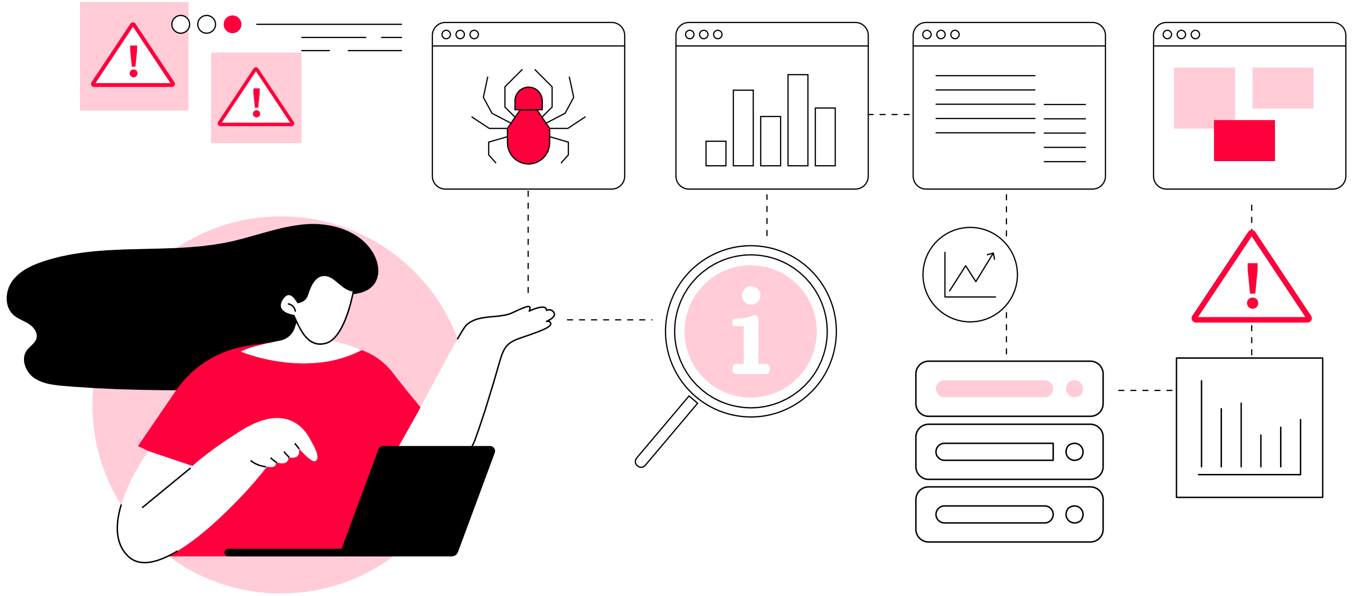
Intellectual Property and Copyright

Many generative AI models are trained on data which includes intellectual property or copyright-protected works. However, oftentimes, generative AI providers do not hold have licensing agreements in place to ensure that the use of such works complies with applicable laws and that there is sufficient protection, attribution, recognition and remuneration given to rightsholders.

From a brand perspective, this can mean that their own intellectual property (IP) is being used to train generative AI models, therefore running the risk of appearing in the outputs generated by others.

A number of class action lawsuits have already been launched against generative AI providers alleging that their works have unlawfully been used to train AI systems and mimic their style. Some tools are now also emerging, such as 'Nightshade', which allow artists to 'mask' and 'poison' their own works to prevent them from being scraped and used by generative AI providers for training purposes.

As a result of these lawsuits, some providers are introducing ways for creators to opt out of model training and developing licensing agreements with some publishers. For example, OpenAI created a channel for artists to inform the company it can't use their artwork for the training of DALL-E, its text-to-image tool. Other providers have started communal funds to share the revenues with the artists whose works were used to train their systems. However, the exact conditions for these remain unclear.



Data protection and privacy

Generative AI models can be trained on personal data, including sensitive information such as biometric data (images, videos of individuals), financial records, personal communications, as well as the data of minors.

In markets with robust data protection and privacy laws, like the EU or California, the collection and use of such data by generative AI providers can trigger specific legal obligations and responsibilities. For example, in the EU a clear legal basis is required for the collection and processing of personal data, even if it's publicly available.

More importantly, in certain jurisdictions there are also obligations on business users of generative AI tools to understand and take accountability for the data upon which the model they are using is trained, the conditions under which such data was collected, and who it is shared with. However, there is often a lack of transparency regarding how the model was trained and whether this was carried out lawfully, making it challenging for brands to assess their legal risks.

Despite data protection and privacy rules already in force in over 126 markets globally, privacy regulators are increasingly concerned about 'data scraping'. 12 of the most important data protection regulators recently issued a joint statement claiming that "individuals lose control of their personal information when it is scraped without their knowledge" and that social media companies must take steps to protect their users from "unlawful data scraping". Most recently, the Italian privacy regulator announced that it is opening an investigation into such data collection to understand whether website owners have appropriate security measures to prevent personal data from being 'scraped'.

As consumers become more aware of how their personal data is being used to train AI models, we may see a rise in lawsuits against AI providers claiming that they have unwittingly used and commoditised their data. In the US, a class action lawsuit has already been filed alleging that OpenAI stole personal data for AI training purposes and that this was an 'invasion of privacy'.

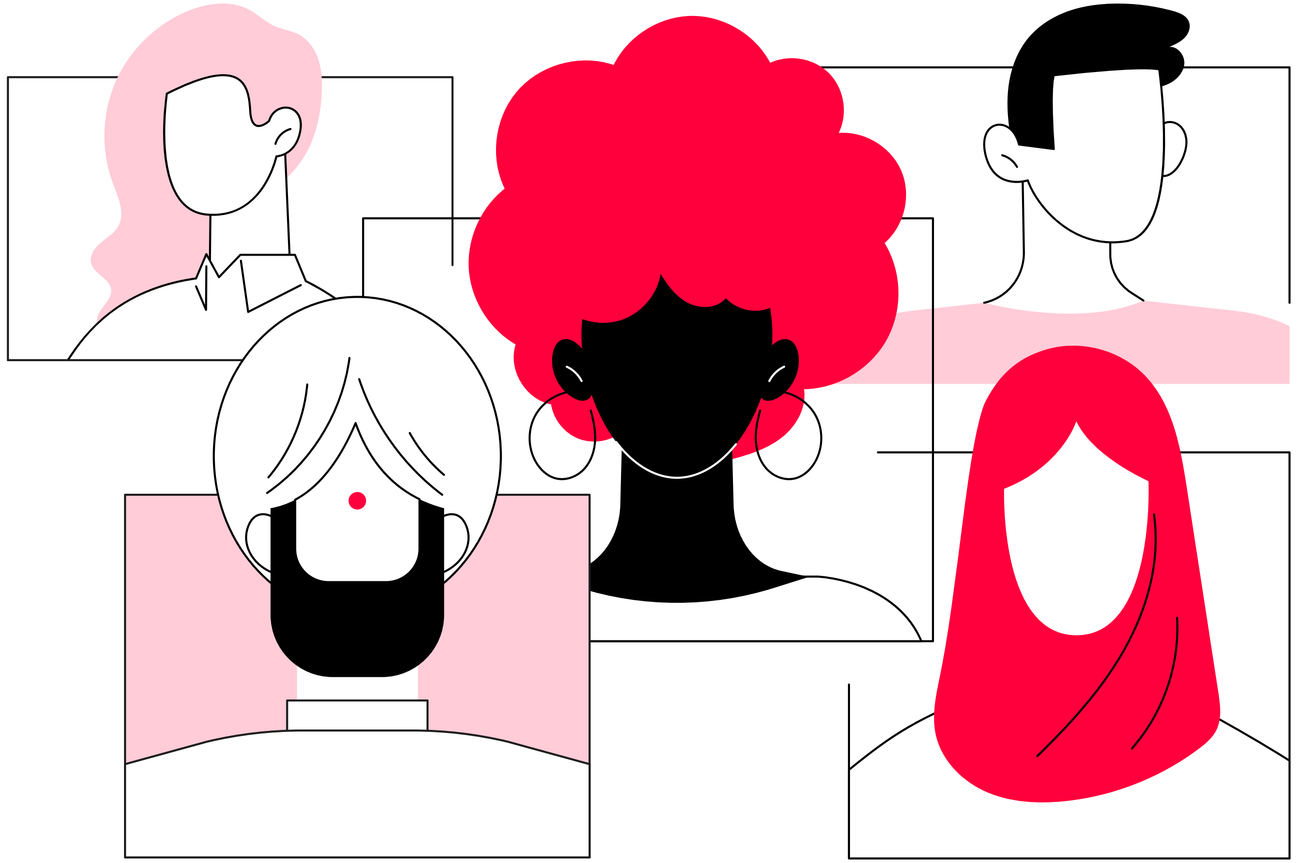
Brands may therefore face reputational risks if they are found to be using models which are considered to have a general disregard for people's right to privacy.

Reliability, safety and integrity

Any harmful or illegal content inadvertently featuring in generative AI training data could make its way in AI outputs. For example, an AI researcher trained a generative AI model using 3.3 million threads from 4chan's infamously toxic 'Politically Incorrect' messaging and image board. The tool then generated racial slurs and antisemitic hate speech which resembled those found on the website.

Most generative AI providers do not moderate the content feeding into the training data, and Stanford University research found that some of the world's most advanced AI models have been trained on graphic child sexual abuse images.

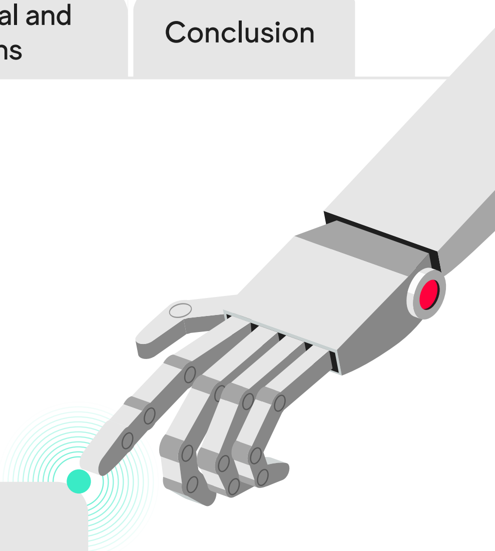
China has become the first country to introduce new rules which would require generative AI providers to conduct a security assessment of all content used to train their public-facing models. Should that assessment reveal that over 5% of the training data contains content considered in China as 'harmful' or 'illegal' (e.g., advocating terrorism or 'undermining national unity') then the model would be prohibited. Although the definition of what constitutes illegal content in China is different to that which can be found in other markets, regulators elsewhere could take similar approaches, requiring providers to take more proactive measures to prevent such content from being used to train their models.



Diversity, equity and inclusion

It is broadly acknowledged that generative AI models can be inherently biased, replicating or even exacerbating societal and historical inequalities. Experts argue that there are two main ways that bias show up in training data: either the data being collected is unrepresentative of reality, or the data reflects existing prejudices and bias.

In the former case, a generative AI model could be fed more photos of light-skinned faces than dark skinned faces. This could mean that the resulting facial recognition system would inevitably be worse at recognising darker-skinned faces. In the latter case, for example, it could result in female candidates being dismissed from certain recruitment processes, because the algorithm has learnt that similar jobs tend to be occupied by men.



Recommendations: Information sources

Brands could consider:

- ✓ **Seeking clarity on the types of data the generative AI model has been trained on** (personal data, biometric data, children’s data, IP and copyright protected works, etc), how that data was collected (i.e. using web scraping) and the steps that were taken by the provider to ensure compliance with existing laws.
- ✓ **Seeking clarity** on what measures the generative AI provider has taken to **ensure data being used to train the model is safe and suitable** (for example, in line with the GARM suitability floor).
- ✓ **Seeking clarity** on the steps the generative AI provider has taken to **ensure the training data is representative**, and not unintentionally biased.

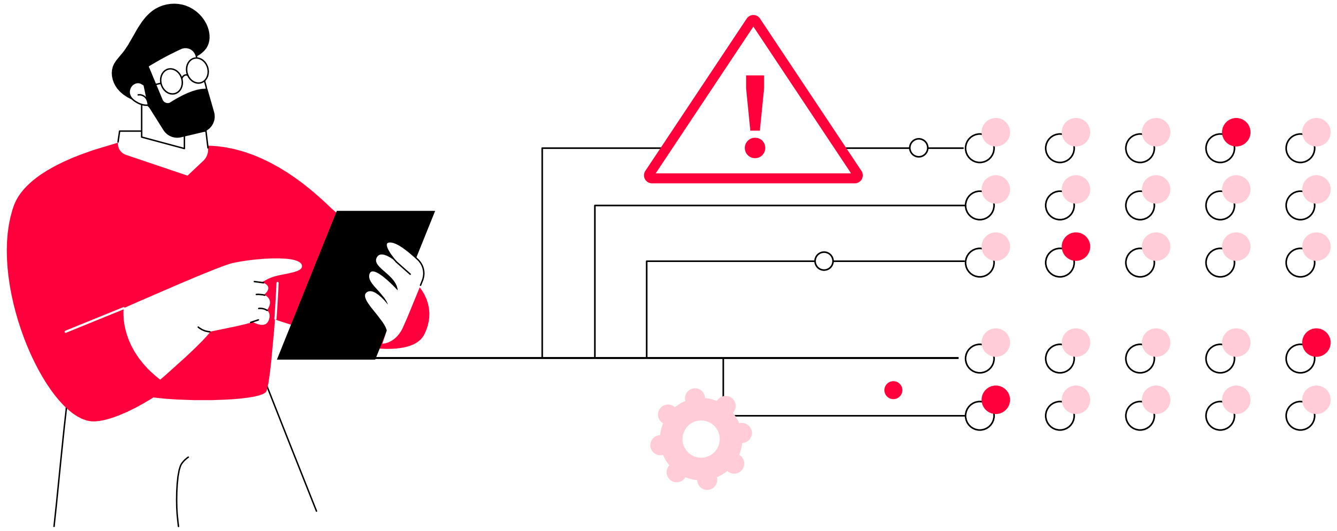
Generative AI providers could consider:

- ✓ **Providing transparency on the sources of data being used** to train their models, and how such data is being collected (i.e. using web crawlers).
- ✓ **Ensuring they have lawfully collected any personal data** or copyrighted/IP-protected works to train their model.
- ✓ **Introducing mechanisms to filter out illegal and harmful content** (including disinformation) from training data, in line with the GARM suitability floor.
- ✓ **Ensuring human review of the data** being collected to train the model to avoid unrepresentative or biased datasets.

Algorithmic Processing

AI technologies have typically been referred to as ‘black boxes’; impenetrable systems which arrive at conclusions or decisions without providing any explanation as to how those were reached. Generative AI is no different, and a lack of clarity remains about the internal workings of how content and answers are derived and generated.

History has shown that the black-box systems typically prioritise the business outcomes of those who create them, and can have severe, harmful and often unforeseen consequences on individuals and society. This is particularly the case as it pertains to social media recommendation algorithms.



As a result, regulators across the world are considering introducing rules which would require AI developers not only to be more transparent about how their algorithms work, but also carry out comprehensive impact assessments to understand any potential risks of their models to the fundamental rights of individuals and society and to be accountable for these.

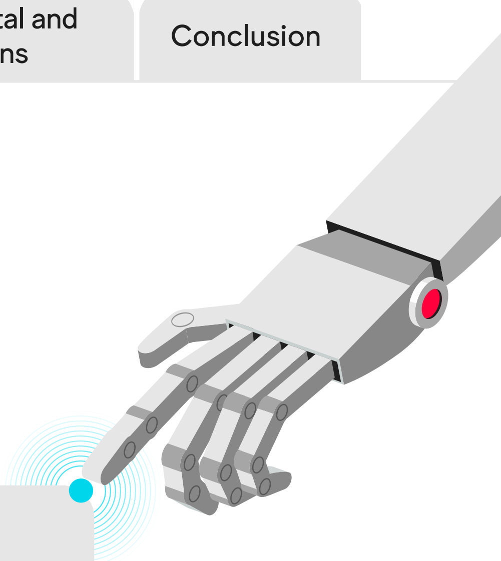
As the true implications of generative AI on users and society are still relatively unknown, it is important for brands to understand how the systems they are using work, and the potential risks these raise more broadly. Using tools which could be considered as having a disregard for people’s fundamental rights or society and democracy more generally could bring reputational ramifications for brands

Diversity, equity and inclusion

While ensuring representation in data sets being used to train the model is important to address potential biases in outcomes, it is also important that diversity, equity and inclusion are built into generative AI models by design.

Many have attributed the existing biases within AI models to the lack of diversity within the industry. As with previous waves of technological advancements, the gender, racial, ethnic biases in generative AI are caused by the inclusion of them at every stage of the AI life cycle, from development, to training and to deployment. The problem, according to the World Economic Forum, is that the field remains male-dominated with only 22% of AI professionals being women. Case in point: the OpenAI executive team are predominantly white and male, with only one female executive.

Recommendations: Algorithmic Processing



Brands could consider:

- **Seeking clarity from generative AI providers regarding how their algorithmic systems are developed and built**, what processes are used to derive outcomes and whether there is human oversight.
- **Understanding whether the generative AI provider has assessed the broader impact of the use of their tool.**

Generative AI providers could consider:

- **Providing meaningful transparency** about how their algorithmic systems work and how decisions and content are generated.
- **Ensuring meaningful representation and diversity** is built into generative AI models by design and into company leadership and decision-making.
- **Carrying out regular impact assessments** on their algorithmic systems and the steps taken to address risks.
- **Carrying out independent audits** of their algorithmic systems and make these publicly available to users.

Protection of the end user

AI models generate content in response to a user prompting the tool with questions, information or data. In most cases, the model keeps a record of the data fed into the tool by the user and in some cases (such as publicly accessible, free third-party tools), this data may then be used to further train the model or to generate outputs for other users. This can raise a number of considerations for brands when it comes to company confidentiality and data protection.

Company confidentiality and security

Models generate content in a response to a user prompting the tool with questions, information or data and in many cases the provider will keep records of the data and use it for their own further purposes. As a result, any proprietary or confidential information fed into the tool by users may then be used to train the model and be disclosed to other users. Often, it is then impossible for users of the tool to retrieve or delete the data they have input into the system.

This could result in brands inadvertently revealing confidential proprietary information (either their own data or data of their partners), without having control over what happens to it and how it is used. In a marketing context, this could mean that newly developed trademarked material (logos and other brand assets) and aspects of campaigns are inadvertently being released ahead of schedule or being shared with competitors.

Numerous companies including Samsung Electronics, Deutsche Bank, JPMorgan Chase, Apple, Amazon have introduced restrictions on the use of generative AI tools by employees after staff had uploaded sensitive code and proprietary information to tools such as ChatGPT, GitHub CoPilot and more. Many companies are now also requiring that teams first conduct AI-specific trainings on how and what types of data should be input into these tools before allowing their use.

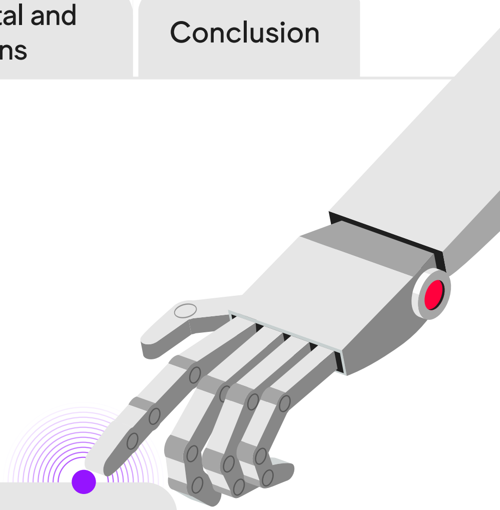
In response to the concerns around company confidentiality, some AI tool providers now offer enterprise versions of their tools which allow companies control over the information they share with the tool and how it is used. For example, enterprise solutions may allow users to disable the data they input into the tool from being used to further train the AI models, enable brands to keep inputs confidential and allow brands to maintain exclusive rights over outputs.

Data protection and privacy

For certain use cases, brands may (unintentionally or intentionally) be input personal data into an AI tool to generate an output and are therefore ‘sharing’ personal data with a third party (the generative AI provider). This could include data they have collected directly or indirectly from individuals from various online and offline sources, or of individuals who they work with (such as employees, supply chain partners and even actors used for previous marketing campaigns).

In many jurisdictions globally, there are strict legal requirements and conditions for the use and sharing of personal data, oftentimes accompanied by robust transparency requirements. Brands should therefore ensure that they are being transparent about how they plan on using personal data for generative AI purposes at the point of collection, and ensure that they are processing, using and sharing such data in accordance with local laws. This is particularly pertinent in cases where the generative AI provider is using the data prompted into the tool by its users for further purposes.

Recommendations: Protection of the end user



Brands could consider:

- ✓ **Using only enterprise versions of generative AI tools** which allow business users to maintain confidentiality of inputs, particularly when sensitive or personal data is being used.
- ✓ **Seeking clarity as to whether the generative AI provider is looking for rights/licensing** to further use information input into the tool.
- ✓ **Informing teams on potential risks** of using free and publicly available generative AI tools and how these can be used in a safe way.
- ✓ **Training teams on types of data** that can and cannot be input into the tool and the risks of sharing sensitive or commercial business data.
- ✓ **Considering imposing limits** on prompting generative AI tools with personal data, in particular sensitive personal data, unless it is necessary for the particular use case.
- ✓ **Putting in place data processing agreements** with generative AI providers to ensure roles and responsibilities for personal data collection and use are clear.

Generative AI providers could consider:

- ✓ **Providing enterprise solutions to business users** which enable them to remain in control of the data they input into the tool and commit to maintaining the confidentiality of inputs.
- ✓ **Ensuring they have security measures in place** such as SOC 2 Type 2 to protect the security of information of users.
- ✓ **Providing transparency about the rights/licensing** they are seeking to further use the information fed into the tool by users.

Safety of outputs

Generative AI tools can produce synthetic content at scale and often in ways which make it difficult to distinguish from human-created content. As outputs are highly dependent on the data that was used to train their systems, this could mean that user-generated content and outputs closely resemble copyright and IP protected works, that they contain personal data, or that they include harmful content such as disinformation.

This also means that generative AI tools could be leveraged by bad actors to spread illegal and harmful content at scale, increasing the risk of brands' ads appearing next to and inadvertently funding content.

Intellectual Property and Copyright

Ownership of generative AI outputs

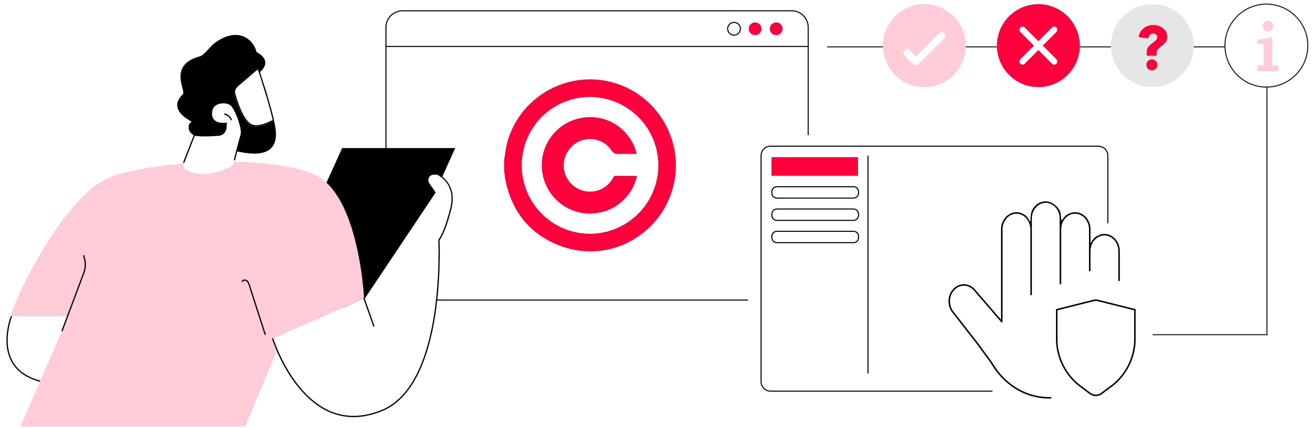
Many AI providers do not transfer the rights over the content generated through the use of the tools (outputs) to the end user, instead using them for their own purposes, such as to train their models. This could mean that brands have no control over how the assets they generate are further deployed and could result in brand-generated content being replicated (intentionally or unintentionally) in the outputs of other end users (including brands' own competitors).

However, even in cases where tools do transfer the rights in outputs to the end-user, the brand may in fact not own any valid IP rights over the content. Copyright law is traditionally granted to humans, but AI-generated content has blurred the lines of authorship. Questions have arisen as to whether there can be any copyright ownership for content generated by machines and, if not, what level of human influence or input is necessary for ownership over such content.

Risks of 'brandjacking'

Generative AI tools present a more efficient and scalable means through which malicious actors can 'brandjack'; use a well-known brand's name, logo, identity or other intellectual property to impersonate a brand or mislead or defraud individuals. For example, a generative AI output could appear branded (carrying a specific identity, logo, brand mascot or other element that distinguishes it from other brand) and advance a particular message (i.e. support for a particular political or societal cause), when the company had no involvement in the development of the content or its messaging. For example, an AI-generated 'deep fake' of singer Taylor Swift was used for a Le Creuset giveaway ad scam.

There are still no universal standards for watermarking AI-generated works or for providing disclosures on the provenance of such content. If 'brandjacking' content is circulated online, including social media platforms, this can therefore have numerous reputational and legal repercussions for brands.



Risks of infringing third-party IP and copyright

Generative AI tools can produce outputs which closely resemble or replicate existing works, particularly when the models are trained on IP and copyright-protected works. However, many generative AI providers free themselves of any liability for copyright or IP infringements in all content generated from the use of their tool. This increases the risks that brands could be held legally and financially responsible for any inadvertent third-party IP or copyright infringements.

There have already been a number of lawsuits brought against generative AI providers over claims that their tools generate outputs replicating their styles, thereby infringing their copyright. For example, Universal Music Group has filed a lawsuit against a music generator, claiming that the model unlawfully generates identical or near-identical copies of iconic songs.

Most of the copyright-related lawsuits have been directed to the generative AI providers, rather than to (business) users of the tool, but this could change as generative AI starts featuring more prominently in brands' marketing campaigns.

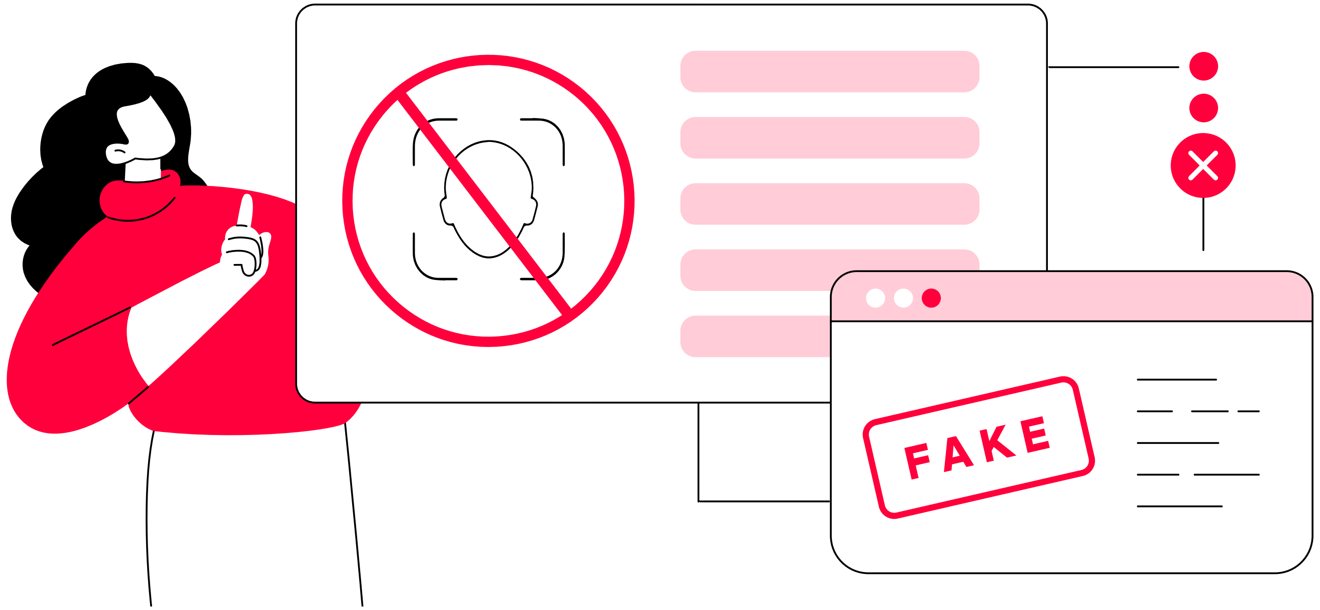
In a response to these lawsuits, certain AI providers have put in place some limited indemnity protections for users of their services. For example, under Google Cloud's 'Generated Output Indemnity', Google reportedly indemnifies users for copyright lawsuits over outputs generated using its models. However, these are tied to specific conditions and products which will warrant further examination.

Data protection and privacy

Content containing personal data of individuals and 'deep fakes'

When generative AI training data or prompts input into the tool by the user contain personal information, there is a possibility that outputs also include personal data. As outlined previously, in certain jurisdictions across the world, brands may need to have an appropriate legal basis for the use and sharing of this data. As AI-generated content proliferates online, users may find their likeness being replicated through 'deep fakes'; digital content purporting to depict a real person or scenario, whether through video, image or audio, and often in a way which is indistinguishable from authentic content. For example, an image generator may be trained on images of real people found on public social media pages and could therefore feature in brand generative AI outputs.

A Wired investigation revealed how an individual found her likeness 'deep faked' into multiple ads on marketplaces across the world. This may mean that brands using synthetic people in their ad creative may unintentionally feature real people who have no say in how their images are being used, raising ethical risks.



Reliability, safety and integrity

➔ (Un)intentional spread of mis/disinformation

Generative AI tools are not credible sources of information and are often prone to ‘hallucinations’; inadvertently generating inaccuracies with a significant degree of confidence. As vividly put forward by the MIT Technology Review, “AI language models are notorious bullshitters, often presenting falsehoods as factors”. Research by Stanford University has found that AI-generated content is only factually correct 25% of the time.

This means that disinformation or harmful content could appear in brands’ own generative AI outputs, and by extension, their ad creative or brand assets. By way of example, misinformation ended up in a Google ad promoting its new generative AI chatbot, Bard. The ad demonstrates Bard claiming that the telescope “took the very first pictures of a planet outside of our own solar system”. However, scientists were quick to point out that this was incorrect, as the very first image of an exoplanet was taken 14 years prior to the telescope.

This error reportedly resulted in Google losing $100 billion in market value.

➔ Abuse by malicious actors

Generative AI represents a more accessible and efficient way to (intentionally) produce harmful content and disinformation at scale. According to the Anti-Defamation League (ADL), Americans are concerned that generative AI will be manipulated by bad actors to spread hateful and harmful content, calling into questions the robustness of tech companies’ trust and safety efforts.

For example, Microsoft Bing’s AI image generator (powered by OpenAI) has been used to create numerous images promoting Nazi propaganda, designed to inflame opinion surrounding the conflict in Israel and Gaza. The Internet Watch Foundation, a nonprofit eradicating child exploitation online, published a report detailing the growing presence of AI-generated child sexual abuse material.

In addition, human rights advocacy group Freedom House claims that generative AI is being used to “sow doubt, smear [political] opponents or influence the public debate”. With more than half of the world’s population participating in an election in 2024, the widespread use generative AI for political purposes raises the risk of an overall decline in information integrity, potentially siphoning money away from quality news and content and threatening societal and democratic wellbeing.

Some AI providers have started taking steps to address these risks. OpenAI has prohibited the use of its tools to impersonate political candidates or for political campaigning in a bid to tackle election misinformation. Meta has also announced new steps to label AI-generated images and disclose video and audio content.

More broadly, this may increase the risk of brands’ ads appearing next to and inadvertently funding content which is harmful, to the detriment of brand safety efforts. Research by fact-checking organisation NewsGuard demonstrates how generative AI is resulting in the proliferation of ‘content farms’, websites that intentionally churn out disinformation or ‘clickbait’ to maximise revenues. And the research finds that 140 major brands are already funding these AI-generated content websites with their ads.

Some tech companies, agencies and publishers have come together and launched the ‘Coalition for Content Provenance and Authenticity’, an alliance focused on developing technical standards for certifying the source and history of media content to address the prevalence of AI-generated misinformation online.

WFA’s Global Alliance for Responsible Media, a brand led initiative aimed at removing harmful content from ad supported media, is collaborating with key industry stakeholders to ensure that advertising revenues are not acting as a financial incentive for the spread of harmful AI-generated content.

Diversity, equity and inclusion

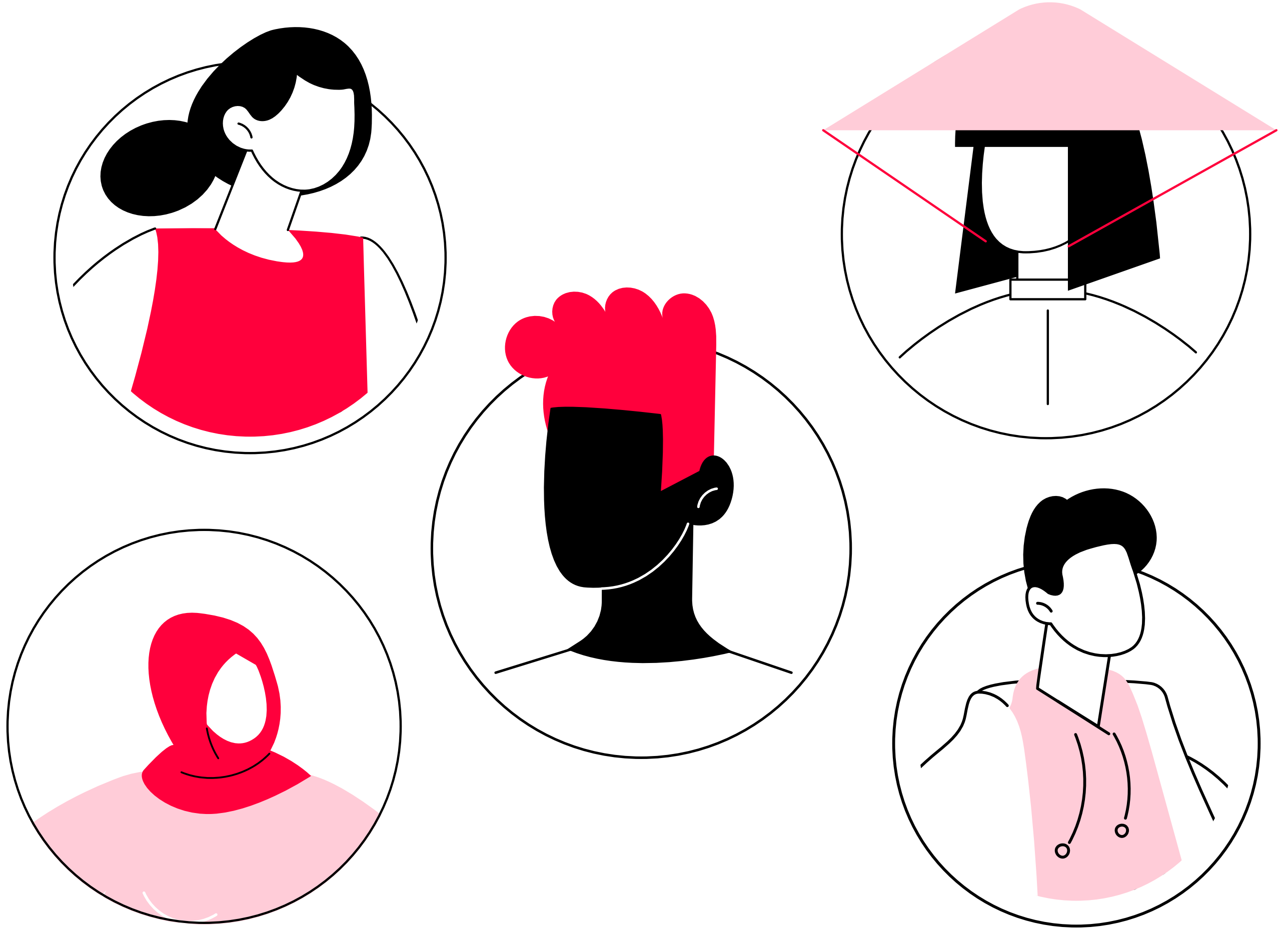
Amplification of existing bias and stereotyping

The UK's Equality and Human Rights Commission has expressed concerns that the use of AI risks "further exacerbating current gender and racial divides".

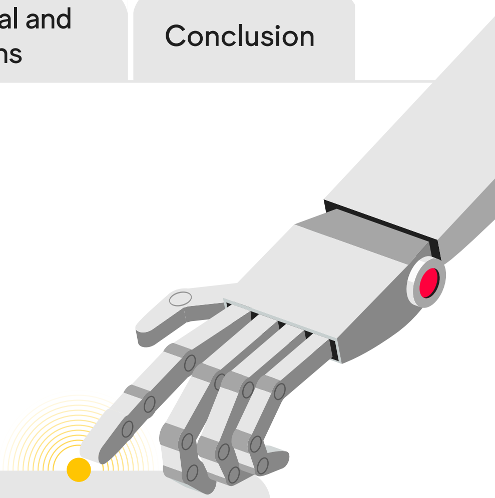
When brands use generative AI tools that are based on unrepresentative data sets, the content they create and use may reinforce gender, racial, ethnic and cultural stereotypes, which could then inadvertently feature in ad creative.

For example, research conducted by Bloomberg found that images generated with the keyword 'inmate' generated people with darker skin, even though people of colour make up less than half of the US prison population. And some gen AI tech companies such as OpenAI have acknowledged that their tools replicate stereotypes; revealing that prompting the word 'lawyer' results disproportionately in images of people who look like middle-aged Caucasian men and wear Western Dress.

As generative AI models become more advanced, the content they create will be increasingly difficult to distinguish from real and human-generated content, making it hard to know what's real. If these images depicting amplified stereotypes find their way back into future models as training data, next generation models could become even more biased, creating a snowball effect of compounding bias with potentially wide implications for society.



Recommendations: Safety of outputs



Brands could consider:

- ✓ **Seeking clarity** on how the AI provider plans on further using the outputs generated by their use of the tool (i.e. rights and licensing arrangements).
- ✓ **Ensuring human involvement and review of outputs** to identify whether assets closely resemble existing copyright or IP-protect works and whether they contain personal data to understand and assess legal obligations.
- ✓ **Identifying and removing** any content or elements depicting or reinforcing gender, racial, ethnic and cultural stereotypes or misinformation.
- ✓ **Avoiding the use of generative AI for 'deep fakes'** or creating synthetic humans or replicating real people in marketing creative unless they have explicit consent from the individual.
- ✓ **Supporting and participating in industry initiatives** such as the Global Alliance for Responsible Media, which is working to prevent advertising revenues from funding harmful AI-generated content.
- ✓ **Redirecting media away from AI-generated content** farms which intentionally churn out click bait and disinformation towards quality media.

Generative AI providers could consider:

- ✓ **Adopting and supporting the development of content provenance solutions** to drive standards for certifying the source of AI-generated media content and addressing AI-generated misinformation.
- ✓ **Taking steps to remove illegal and harmful content** in outputs, in line with the GARM suitability framework.
- ✓ **Providing transparency about the rights/licensing** they are seeking to further use the outputs generated by users.

Publishers & online platforms hosting AI-generated content could consider:

- ✓ **Ensuring they have mechanisms in place to prevent the monetisation of AI-generated harmful content** and misinformation (including 'brandjacking'), in line with the GARM safety floor and suitability framework.
- ✓ **Providing transparency on the performance of the tools** in place to demonetise and remove AI-generated harmful content.
- ✓ **Adopting and supporting the development of content provenance solutions** to drive standards for certifying the source of AI-generated media content and addressing AI-generated misinformation.

Broader environmental and societal considerations

Although concerns around IP and confidentiality, privacy, diversity equity and inclusion and brand safety represent more immediate risks for brands, there are a number of other ethical considerations that brands should take into account when procuring generative AI.

➔ Transparency and consumer harm

As mentioned previously, research has shown that consumers are concerned about how generative AI may be used for harmful purposes, including for the spread of ‘deep fakes’. And an IPSOS public poll found that over 40% said they would trust companies and brands less if they used AI more, including for advertising purposes.

The Federal Trade Commission (FTC) in the US, responsible for enforcing consumer protection laws, has explicitly stated that companies using ‘deep fakes’ in advertising could be ‘misleading’ consumers, and could therefore result in FTC enforcement action. And the recently adopted EU AI Act puts forward stringent transparency requirements on the use of AI in content, regardless of whether deep fakes feature.

Brands will need to consider how their use of AI may impact consumer perceptions of their brand, and the appropriate transparency and disclosure requirements necessary to build trust with key stakeholders.

➔ Impact on the creative industry and brand authenticity

Advances in generative AI technologies have facilitated the cloning and synthesis of people and voices through the use of material such as recordings, videos and scripts; enabling the replication of actors’ images, facial expressions, unique speech patterns, pronunciation and emotional range at scale.

In a marketing context, brands have been leveraging such tools already within their ad creative, replicating the voices and images of famous actors and synthetically altering them. The costs of such generative AI tools are generally low and therefore represent an attractive alternative to using real-life actors and voice artists. However, there are concerns about the unauthorised and uncompensated use of actors’ likeness in AI-generated synthetic content.

This also raises questions about how entire segments of the industry could be pushed out of work, replaced instead by machines. There have already been numerous cases where famous actors’ likeness has been used in online ‘deep-fake’ ads without permission or compensation.

It is precisely within this broader context that in July 2023, the Screen Actors Guild-American Federation of Television and Radio Artists Guild (Sag-Aftra) and the Writers Guild of America (WGA) launched a strike over concerns around companies paying them for one day’s work and then ‘owning’ their image, likeness, voice on any future project, without consent or compensation. They have also argued for more protection to avoid being the targets of ‘deep fake’ technologies for the purposes of advertising, sham songs, pornography and more.

More generally, given the scale and pace at which generative AI produces content based on existing works, investing too heavily in such technologies for creativity could result in a loss of originality, causing the saturation of similar-looking and repetitive material. It is widely acknowledged that one of the shortcomings of generative AI is its inability to incorporate and embody brand identity and authenticity. Relying too much on these tools could therefore dilute the uniqueness of their branding and creative content, making it harder for them to stand out from the crowd.

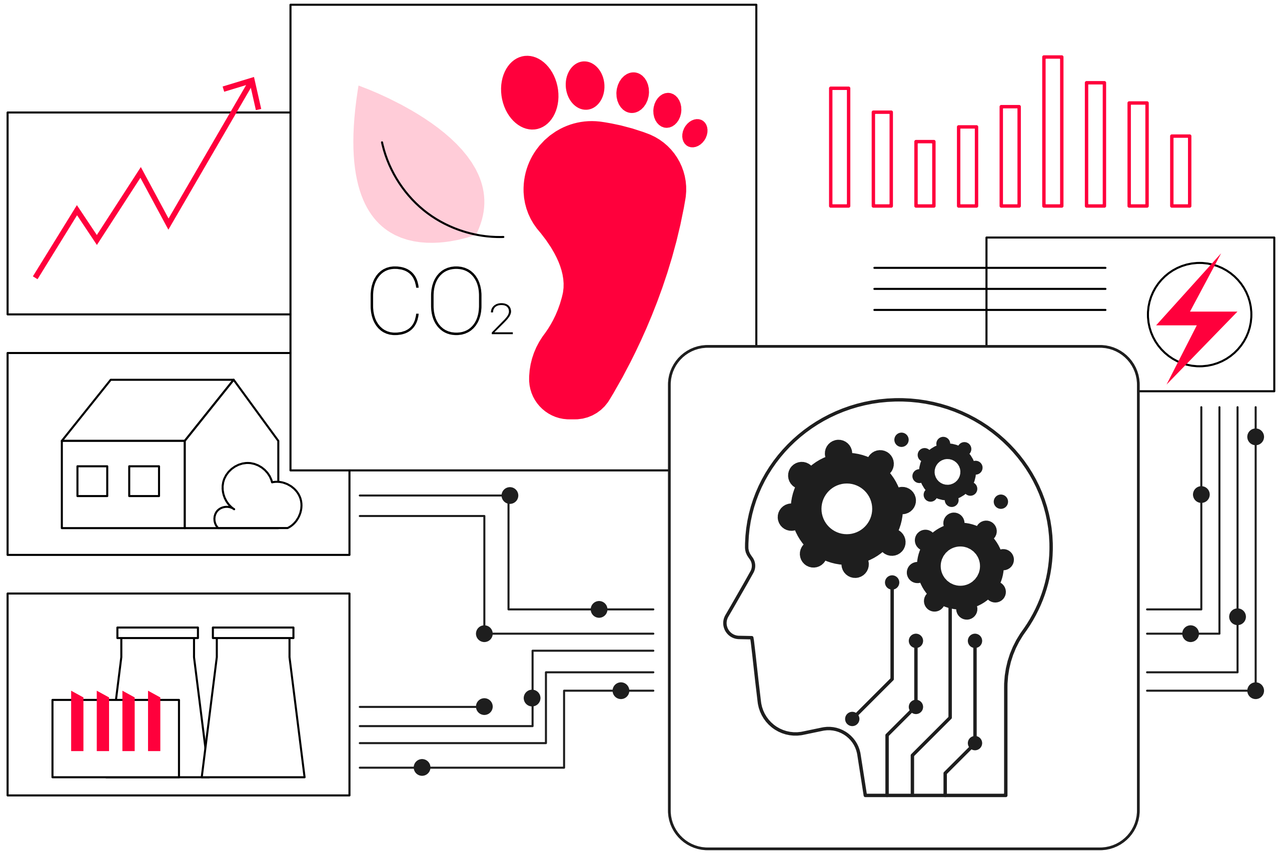
Environmental sustainability

Although the environmental costs and impact of generative AI are often overlooked, this technology relies on significant energy for development, training and deployment.

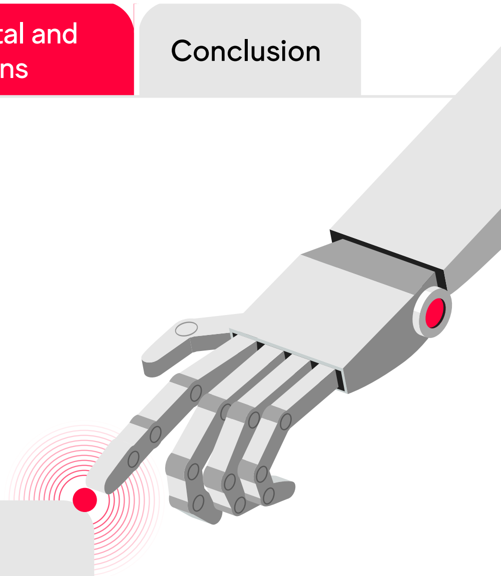
Research has typically focussed on quantifying the operational energy and carbon required to perform the training and development of machine learning models. According to research by the Massachusetts Institute for Technology, training just one AI model can emit more than 626,000 pounds of carbon dioxide, equivalent to nearly five times the lifetime emissions of an average car.

However, there has been a lack of research quantifying the resources required to deploy generative AI models at scale. Some estimate that the resources for deployment could amount to as much as 80% of total emissions of generative AI models, and are therefore much more significant than the resources used for training and development. Emissions for deployment will vary depending on the generative AI model; for example, some have found that an image generator could use up to 60 times more energy than a text generator.

Calculating and understanding the environmental impact of brands' use of generative AI remains a challenge, and more independent research is needed to drive transparency in this area.



Recommendations: Broader environmental and societal considerations

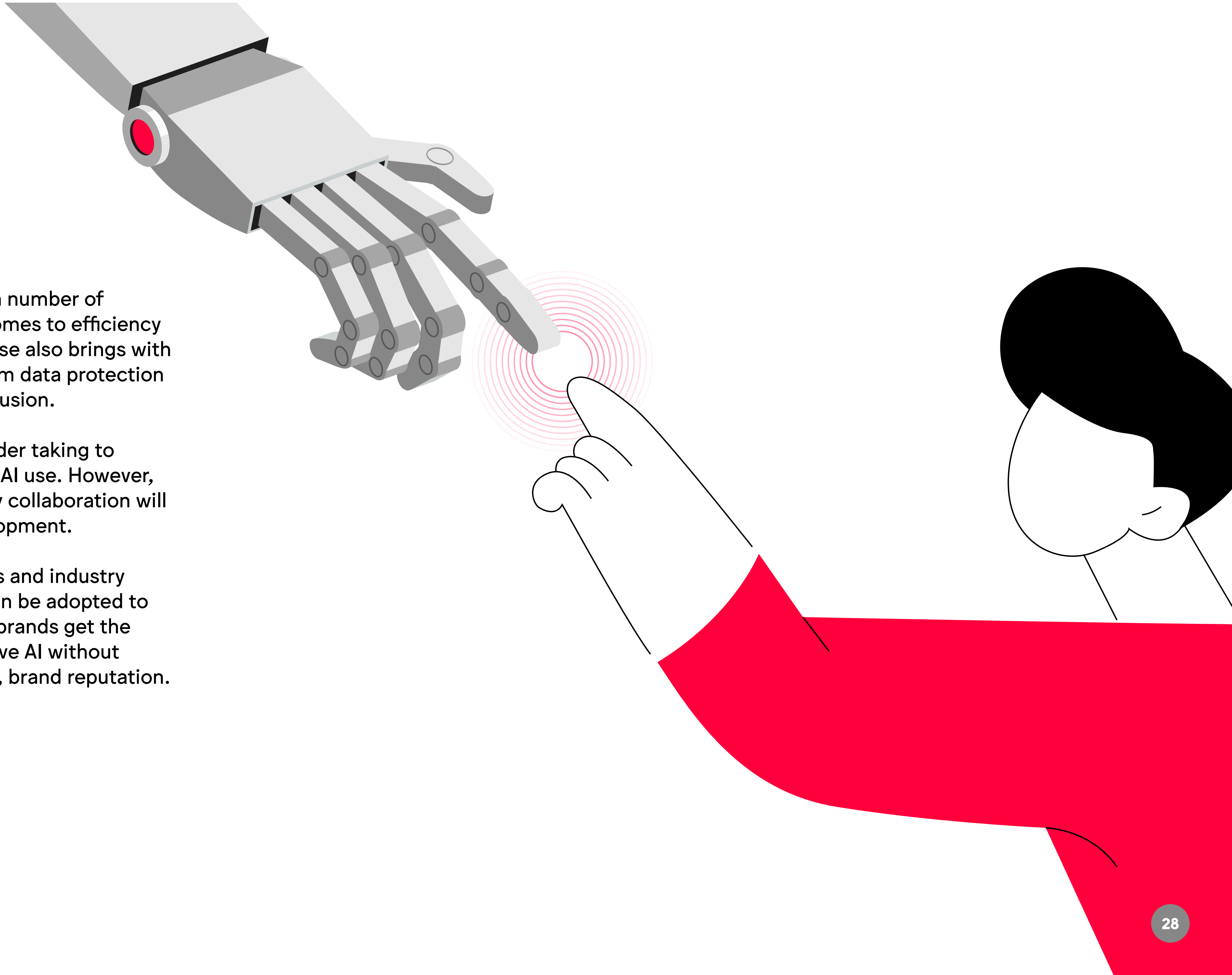


Brands could consider:

- **Supporting industry research** aimed at developing a model for calculating the carbon footprint of generative AI.
- **Incorporating and considering the environmental impact** of generative AI tools when contemplating their use internally.
- **Understanding how consumers perceive their use of generative AI in marketing,** and putting in place robust disclosure and transparency measures.
- **Considering using real actors and artists** over AI-generated synthetic versions.
- **Compensating actors and artists fairly** for the use of their voice and likeness for AI purposes.

Generative AI providers could consider:

- **Considering cleaning up value and resource chains** to limit environmental impact – from pollution and emissions to raw materials, depletion etc.
- **Taking steps to prevent creators' likeness from being used** without their permission or without appropriate compensation, particularly to train their models.



Conclusion

The use of generative AI in marketing will continue to bring a number of benefits and opportunities for brands, particularly when it comes to efficiency and productivity, creativity and effectiveness. However, its use also brings with it many considerations and potential challenges, ranging from data protection and privacy, to IP and copyright and diversity, equity and inclusion.

There are already a number of steps that brands could consider taking to mitigate some of this risk and propel responsible generative AI use. However, some issues are beyond a brands’ direct control and industry collaboration will be crucial in driving forward sustainable and ethical AI development.

In the coming months, WFA will be working with both brands and industry partners to develop practical solutions and measures that can be adopted to mitigate risk. We believe this work will be critical to helping brands get the assurance they need to leverage the possibilities of generative AI without compromising on trust, safety and inclusivity, and ultimately, brand reputation.

World Federation of Advertisers
London, Brussels, Singapore, New York

wfanet.org

info@wfanet.org

+32 2 502 57 40

twitter @wfamarketers
youtube.com/wfamarketers
linkedin.com/company/wfa

King & Spalding

Competition compliance policy

The purpose of the WFA is to represent the interests of advertisers and to act as a forum for legitimate contacts between members of the advertising industry. It is obviously the policy of the WFA that it will not be used by any company to further any anti-competitive or collusive conduct, or to engage in other activities that could violate any antitrust or competition law, regulation, rule or directives of any country or otherwise impair full and fair competition. The WFA carries out regular checks to make sure that this policy is being strictly adhered to.

As a condition of membership, members of the WFA acknowledge that their membership of the WFA is subject to the competition law rules and they agree to comply fully with those laws. Members agree that they will not use the WFA, directly or indirectly, (a) to reach or attempt to reach agreements or understandings with one or more of their competitors, (b) to obtain or attempt to obtain, or exchange or attempt to exchange, confidential or proprietary information regarding any other company other than in the context of a bona fide business or (c) to further any anti-competitive or collusive conduct, or to engage in other activities that could violate any antitrust or competition law, regulation, rule or directives of any country or otherwise.

Please note that the recommendations included in this document are merely meant as suggestions or proposals. They are not binding in any way whatsoever and members are free to depart from them.