Digital Regulation Cooperation Forum

The Future of Synthetic Media

A Summary of Stakeholder Views on the Future Development of Synthetic Media and its Implications for DRCF Regulators

Publication date: 26 November 2024

CMA Competition & Markets Authority
Ofcom
ico. Information Commissioner's Office
FCA FINANCIAL CONDUCT AUTHORITY

Contents

1. Executive Summary

This paper explores how synthetic media might develop, what its uses might be, and potential regulatory implications over the next 3-5 years.

For the purposes of the project, we defined synthetic media as "an umbrella term for video, image, text, or audio that has been generated in whole or partly by artificial intelligence (AI) algorithms". Definitions of synthetic media and deepfakes varied between our stakeholders and we explore this in Section 2.2. While we have not adopted a common definition for deepfakes, we agree that these are a subset of synthetic media.

For this research, the Digital Regulation Cooperation Forum (DRCF) held discussions with a range of stakeholders, including other regulators, Government, industry, academia, and civil society. This paper provides a summary of those stakeholders' views on synthetic media and its implications for DRCF regulators, it does not represent a policy position of the DRCF or its member regulators.

Technological Development

Technology developments and their implications are explored in *Section 3*.

There was consensus amongst our stakeholders that:

- over the next 3-5 years, synthetic media will become more widely integrated in online content and services;
- the technologies that create synthetic media, such as AI chatbots and audiovisual AI generators, will become **more** sophisticated and easier to access and adopt for a wide range of users; and
- synthetic media will become harder to distinguish from other content.

Opportunities and Risks

Stakeholders emphasised that synthetic media could present both opportunities and risks, both of which require regulatory consideration. Opportunities, which are explored in greater detail in *Section 4.1*, include:

- **Creative content production:** Synthetic media can be used to streamline content creation, increase the quality of outputs, lower production costs, and democratise the production of creative outputs.
- **Personalisation:** Synthetic media may enable the personalisation of consumer journeys by tailoring recommendations based on individual preferences.
- **Synthetic avatars:** Digital representations modelled to look and behave like real humans can be used to enhance consumer experiences in entertainment, work, and even in grief counselling.
- **Digital twins:** Virtual representations of physical objects or systems which mirror their real-time conditions can be applied across industries. For example, these can be used in healthcare to personalise treatments, or in industries for product prototyping.
- **The creation of synthetic datasets:** Using algorithmically generated datasets, organisations can perform processing with a reduced risk of exposing real personal data.

Stakeholders believed that it is likely that bad actors will continue to employ synthetic media for malicious ends. Such activities will likely entail more targeted, harmful activities resulting in regulatory challenges. Beneficial use of synthetic media is also likely to multiply and increase. Risks, which are explored in greater detail in *Section 4.2*, include:

- **Harmful/illegal content creation:** Synthetic media could facilitate the mass creation and distribution of harmful and illegal content, including synthetic non-consensual sexual images, child sexual abuse material (CSAM), and hate speech. The spread of such content can lead to harm for individuals and society more broadly.
- **Psychological impact:** Beyond content that is deliberately harmful or illegal, some synthetic media may have secondary psychological effects that influence those that consume the media, particularly where those individuals may already be vulnerable.
- **Circumvention of authentication**: Bad actors might use synthetic media to circumvent security and identity checks. This can enable scams and fraud at both a personal and organisational level, posing a significant threat to consumers, industries, and markets.
- **Disinformation:** Synthetic media could be used to produce highly convincing disinformation, including political deepfakes. Such content has the potential to destabilise democracy and markets, and to erode trust in media.
- **Misleading consumers:** The creation, use and publication of misleading synthetic advertisements, false endorsements, or synthetic products could negatively impact both consumers and markets. This could breach existing consumer law.
- **Copyright infringement:** The use of synthetic media trained on existing intellectual property (IP) may have legal and ethical implications for copyright protection and fair use. Such models can generate 'copycat' content, which diminishes the exclusivity and profitability of original IP.
- **Inaccuracies and a lack of transparency:** Where synthetic media is being generated by black-box systems which those using them do not fully understand, there is the risk that the media generated may be unintentionally inaccurate or misleading.

The Future of Regulation

As synthetic media might impact across many online sectors and industries, regulatory collaboration will be critical to ensure effective individual, and market protection, as well as ensuring that innovation is enabled and can contribute to economic growth. A holistic effort from regulators, Government, academia, industry, and civil society will be important. Regulatory considerations are explored in *Section 6*.

Stakeholders agreed that regulators will continue to have a substantial role to play in relation to the emergence of synthetic media. Regulators already have a role under existing regimes in certain circumstances – for example synthetic media that could be illegal or harmful to children may be in scope of the Online Safety Act 2023, and Ofcom already has broad Media Literacy duties. Further, where businesses are engaged in commercial practices concerning the creation, use or publication of synthetic media to promote, sell or supply products to or from consumers, they will have responsibilities under consumer law. Platform operators that publish or display such content also have responsibilities under consumer law, whether or not they sell or supply products to consumers themselves.

Many stakeholders argued that additional legislation and policy initiatives may be required to keep pace with technological change. Regulators should both be alert to the harms amplified by synthetic

media and foster environments which enable potential benefits for individuals, markets and wider economic benefits.

Conclusions and Next Steps

There was agreement amongst the stakeholders interviewed by the DRCF that synthetic media and deepfakes are likely to improve in fidelity and ease of access. This would result in the creation of synthetic media that would be increasingly difficult to identify and discern from media which had been created without the use of artificial intelligence. DRCF member regulators will need to understand how their remits might be impacted. Stakeholders expressed both optimism around how the technology could be deployed beneficially and concerns about the risks it could present. They agreed that the role of online regulators was singularly important to ensure positive outcomes for individuals.

Regulatory considerations include responsible innovation to derive the benefits from synthetic media, the use of personal data within synthetic media, the protection of consumers and markets, security and resilience, and mitigating threats to democracy and society more broadly. *Section 7* explores the next steps that the DRCF will take in relation to synthetic media both in the short and long term.

Moving forward, DRCF member regulators will:

- Follow the development of synthetic media and deepfakes technologies.
- Continue to engage and bring together industry, Government, academia, and others on the subject.
- Continue to consider synthetic media where relevant to their respective remits, whilst also exploring opportunities for further collaborative efforts through the DRCF.

2. Introduction

The DRCF and HSET team

The DRCF was established to ensure coherence between the regulatory regimes of its member regulators (the Competition and Markets Authority (CMA), Financial Conduct Authority (FCA), Information Commissioner's Office (ICO), and Ofcom), to work together on complex challenges across the digital landscape and develop capabilities for the future.

The DRCF Horizon Scanning and Emerging Technology team (HSET) is the leading cross-regulatory voice on emerging technologies and trends in digital markets. We take a proactive approach to understanding the potential benefits, risks, and regulatory implications of emerging technologies. We subsequently provide actionable insights to regulators, other public bodies, government, parliament, industry, and the public.

Definitions

For the purposes of this project, we adopted Ofcom's definition of synthetic media: "an umbrella term for video, image, text or voice that has been generated in whole or partly by artificial intelligence (AI) algorithms".[1] We have not set out to find a common definition for deepfakes but agree that deepfakes are a subset of synthetic media.

Through our stakeholder engagement with industry, academia, and government agencies, the majority broadly agreed with our definition of synthetic media. Slight divergences included:

- whether synthetic media must be generated by AI;
- the types of media that should be included in the definition;
  - whether the definition should include text because people experience text and visual content very differently
  - whether the term "voice" should be amended to "audio" to represent the different sounds that people want to make when producing synthetic media; and
- whether the definition of synthetic media should include consideration of its intent.

Most stakeholders did not have a separate definition for the concept of deepfakes[2] but agreed that deepfakes are a subset of synthetic media. Only a few used the two words interchangeably. Some of those who characterised deepfakes as a subset of synthetic media used characteristics such as intent (whether the content is misleading or intended for manipulation—meaning you cannot have 'good' deepfakes) and the quality of content (such as likeness) to distinguish between the two concepts. Some stakeholders considered that a deepfake should be a recreation of a real person, citing the origin of deepfakes in face swapping and dubbing, but others felt deepfakes are a broader category, giving examples such as fake scenes from a warzone being deepfakes without a real person as the subject.

[1] Ofcom definition of Synthetic Media can be found in its Note to Broadcasters publication

[2] Though there are some definitions. For example, Ofcom defines deepfakes as "audio-visual content that has been generated or manipulated using AI, and that misrepresents someone or something" - see https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/deepfake-defences/

Why Synthetic Media?

External stakeholders urged us to explore this topic as part of our 2024/25 workplan, and discussions with colleagues across the DRCF indicated that this is an area of clear cross-regulatory interest. We discuss the future of regulation and the cross-regulatory implications throughout the paper; however, a short summary of each member regulator's interest and work to date is provided below:

CMA:

- The CMA helps people, businesses and the UK economy by promoting competitive markets and tackling unfair behaviour. The CMA is interested in how the market for synthetic media might develop, and how the emergence of synthetic media might impact competition in established markets.
- The CMA is further interested in scenarios where synthetic media may be used to defraud, mislead or manipulate consumers and the creation, use and publication of such content may result in breaches of consumer protection law by one or more actors in the supply chain – including the operators of online platforms where material is published. The CMA has conducted several consumer enforcement investigations into the presence or facilitation of economically harmful illegal content on online platforms. By this, we mean content originating from third parties and involving an unfair commercial practice under the Consumer Protection from Unfair Trading Regulations 2008 (CPRs). For example, the CMA has opened enforcement cases in relation to fake consumer reviews of products being displayed on major platforms like Amazon and Google.
- The CMA has been considering the potential impact of AI and related matters on competition and consumer protection issues for a number of years including the 2021 paper, 'Algorithms: How they can reduce competition and harm consumers', and an extensive programme of research into AI foundation models, products and services. AI foundation models, with their vast training on diverse data, can enable the generation of synthetic media by producing realistic and contextually accurate content. The CMA's work into AI foundation models includes the initial review into foundation models in September 2023, an update paper and technical update report in April 2024. The CMA also developed a set of principles to guide the development of the market towards positive outcomes, promoting innovation and growth while ensuring fair competition and consumer protection. The CMA is continuing to monitor developments in AI related markets, and we are considering opportunities where competition could unlock barriers to investment and innovation across the AI value chain – from infrastructure and development, to release and deployment.

FCA:

- The FCA is interested in the risks and opportunities synthetic media may drive in the UK's financial services industry and is committed to protecting consumers, enhancing market integrity and promoting competition in the interests of consumers. The FCA has already supported around 200 firms through its innovation services, particularly in machine learning, natural language processing, and generative AI. Coupled with its recently launched AI Lab, the FCA is well positioned to develop firsthand insights into emerging AI use cases.
- The FCA recognises that this is a fast-moving area. We are interested in how synthetic media might drive beneficial and responsible innovation and help to address challenges in financial services, such as enhancing financial inclusion, preventing financial crime and improving financial risk management. The FCA is also interested in the risks synthetic media could create, such as disinformation causing market volatility and synthetic financial promotions misleading or defrauding consumers. These risks and benefits are likely to be considered in use cases or events in the AI Lab.

- The FCA's work on synthetic data - a privacy-preserving technique that can be used to develop advanced modelling techniques and train AI models without compromising individual privacy or breaching data protection law – is another significant aspect of its broader AI strategy. This includes publishing an Organisation for Economic Co-operation and Development (OECD) blog post: 'Can synthetic data enable data sharing in financial services?'. Furthermore, there have been commitments to deepen the FCA's understanding of use cases. The Synthetic Data Expert Group, constituted in 2023, published its report on Using Synthetic Data in Financial Services, sharing knowledge of current use cases and addressing technical challenges.
- The FCA has also been exploring synthetic data use for combating financial fraud and crime. In September 2023, the FCA launched an Authorised Push Payment (APP) synthetic dataset to develop products and services that can minimise fraud in partnership with City of London Corporation. The FCA is working on a synthetic data project that aims to foster innovation in the detection of money laundering, as mentioned in its AI Update.

ICO:

- As the independent authority for data protection and freedom of information, the ICO's interest in synthetic media and deepfake technologies centers around how people's personal data is processed within those technologies. This interest encompasses both mitigating harms from malicious uses of the technologies, and supporting the development of innovative positive uses, in line with privacy by design and default.
- Any organisation which is processing the personal data of individuals either to generate synthetic media, or to identify it will need to understand their obligations under data protection law, including how the rights of individuals are exercised.
- Data protection by design and default is a legal requirement for organisations looking to develop systems and products which use personal data. The ICO has published support on how to meet that requirement.
- Beginning in early 2024, the ICO launched a series of call for views as part of a consultation on how data protection law should apply to the development and use of generative AI models. As generative AI is developed and deployed in ways that are distinct from simpler AI models used for classification or prediction objectives, the ICO has sought feedback from stakeholders on issues including the accuracy principle, interaction with data subject's rights, and purpose limitation. The input received through the call for views will be used to update our guidance on AI and other products.
- The data protection implications of methods used to identify synthetic media and deepfakes will feature in an upcoming Technology Horizons Report to be published in Q1 of 2025.
- The ICO has published guidance for organisations considering the use of synthetic data as a privacy-enhancing technology.

Ofcom:

- Ofcom has an interest in synthetic media from an Online Safety perspective. When the new duties under the Online Safety Act 2023 come into force next year, regulated services like social media firms and search engines will have to comply with risk assessment duties and safety duties in relation to illegal content and content that is harmful to children, which may include some types of deepfake content.
- Ofcom's illegal harms and children's safety codes will set out the measures that are recommended for services to take, some of which could be relevant to addressing harms associated with deepfake content. Ofcom has also considered how the harms of deceptive deepfakes can be mitigated in a recent discussion paper. Ofcom has also set out how the

Online Safety Act 2023 will apply to Generative AI and chatbots in an open letter to UK online service providers.

- Further, Ofcom has media literacy duties around supporting people to understand technology and to protect themselves, which has links to synthetic media, for example synthetic mis/disinformation. Ofcom published a set of Best Practice Design Principles for Media Literacy that sets out how platforms can promote media literacy through on-platform interventions. Ofcom has also published papers on understanding generative AI in the context of media literacy.
- In relation to the broadcasting sector, Ofcom has set out its considerations in its Note to Broadcasters. This note will also guide Ofcom's approach to use of synthetic media by services that will be subject to the new Video-On-Demand Code, under new duties granted by the Media Act 2024.
- Synthetic media may also have implications for the proliferation and development of scams and fraud across the communications sector.

Approach

To provide actionable insights, the DRCF conducted desk research and held discussions with a range of stakeholders, including other regulators, Government, industry, academia, and civil society. We held workshops with experts across each member regulator and developed scenarios (as set out in Annex 1) to inform our thinking on potential futures. This paper collates our research and foresight activities, and provides insight into definitions, use cases, future scenarios, and regulatory considerations relating to synthetic media and deepfakes.

3. Technological Development

The stakeholders we interviewed mostly stated that synthetic media technology will evolve towards a greater level of sophistication and become harder to differentiate and detect over the next 3-5 years. Synthetic media will be more prevalent online as technology becomes cheaper and more accessible. As such, stakeholders expect to see greater use of and weaponisation of synthetic media, as well as the rise of 'cheapfakes'[3] which will potentially raise further misinformation challenges.

Sophistication and Likeness

The level of sophistication and likeness of synthetic media is expected to increase significantly in the coming years, with progress likely particularly in voice technology. Stakeholders suggested that many of the current limitations like unnatural audio and mouth movements in deepfakes may no longer exist as the technology evolves. Synthetic voice, which is already relatively advanced, is expected to further improve, especially in voice replication and interactive voiceovers. Stakeholders noted that synthetic videos and images are less sophisticated than voice technology, mainly because generating high-resolution images and videos requires significant computing power, though this quality gap may close in the medium term. As a result, synthetic content may be more difficult to detect and could become nearly indistinguishable from other content.

Accessibility

One of the most notable trends in synthetic media technology development is its growing ease of access. Stakeholders noted that current users of the technology face some, albeit low, barriers-to-entry like login requirements and paywalls, and may need specific technical knowledge to navigate high-quality synthetic media generation. It is anticipated these obstacles will be reduced or removed in the future, with synthetic media becoming more easily accessible for day-to-day use via generative AI mediums, in-app and on-device generative models. Technical knowledge will no longer be required, and the cost will be reduced or non-existent.

Overall, stakeholders expect that generating synthetic media will become cheaper, faster and easier, and its use will be more widespread. An interesting implication, from a different perspective, was that other, 'non-synthetic' content may become more valued and expensive in the future if it becomes rarer.

Weaponisation

Since synthetic media technology is expected to be cheaper and more accessible, we may see the increased weaponisation of the technology, particularly for more targeted disinformation activities. Stakeholders argued that such activities will be less likely with larger synthetic media generation platforms which may be subject to strict regulatory measures, but emerging smaller open-source platforms may enable this activity because they can be harder to target and regulate. It is likely we will also see the rise of cheapfakes, and several stakeholders warned that the ease of creating and sharing large volumes of cheapfakes poses a bigger challenge for misinformation than sophisticated deepfakes.

[3] Refers to low effort and easily produced content spreading quickly before being debunked.

4.1 Opportunities

Through our stakeholder engagement, we identified four broad areas where synthetic media may be used to the benefit of individuals, businesses and society more generally including through the potential contribution to economic growth. These opportunities, and the associated regulatory considerations, are explored below. We note that some of the applications discussed in this section may also raise risks. We discuss risks, and the associated regulatory considerations, in Section 4.2.

4.1.1 Synthetic Media for Creative Content Production

"So far, synthetic media has democratised creative execution – people used to need specific skills to create synthetic content, now they do not necessarily".

– Academic

Overview

Synthetic media is transforming the way creative content is produced across multiple sectors. It has the potential to speed up content creation, increase the quality of outputs, lower production costs and democratise creative execution. Key areas identified during research and stakeholder engagement include:

- **Entertainment content production:** Synthetic media can be used to streamline film, TV and online video production by generating digital environments, virtual characters and special effects more efficiently. An emerging area is text-to-video for film and video production.
- **Gaming:** Synthetic media has been used in the gaming industry. It is expected to speed up the development of voiceovers, non-player characters (NPCs), and virtual environments and enable more immersive and dynamic gaming experiences.
- **Art and culture:** Synthetic media is said to be democratising the creation of digital art and personalised designs as people no longer need specific skills to create digital content with synthetic media tools readily available. Synthetic media also has several promising applications in the cultural sector, such as multi-language translation for audio-visual content and recreating historical figure chatbots to engage with museum visitors.
- **Marketing and advertising:** Synthetic media will potentially enable advertisers to generate highly complex content including targeted and personalised advertisements, product reviews and even entire marketing campaigns in a fast and low-cost manner.

Regulatory Considerations

i. Democratisation of content creation

Stakeholders suggested that the increased adoption of synthetic media in creative content production is likely to lower market entry and change the competitive dynamics of markets. Synthetic media's ability to potentially lower production costs and improve the quality of digital content may increase opportunities for smaller creators to produce high-quality audiovisual content and compete with major studios, which may encourage new content platforms to thrive and put competitive pressure on incumbents. However, there are still concerns that smaller players could be acquired by larger technology firms, which may ultimately reduce the number of competitors and content diversity leading to fewer choices for consumers. Ofcom, the FCA and the CMA may

particularly need to monitor changing competition dynamics in synthetic media markets to protect consumer choice and fair market competition.[4]

Stakeholders also argued that industry should consider the balance of using synthetic media for efficiency gains against supporting original content creation for sustainable growth of the sector. Potential risks to the content sector are explored in Section 4.2.

ii. Marketing, advertising, and consumer protection

To the extent that such media is created, used or published by businesses in connection with the promotion, sale or supply of products to or from consumers, the CPRs are likely to apply to those businesses. For example, as described above, advertisers and other businesses who create or use online avatars to promote advertised products to consumers must ensure that they do not mislead consumers by action or omission and businesses who operate online platforms which publish content from third parties must take such appropriate steps as are necessary to prevent and remove false and misleading content from publication.

4.1.2 Synthetic Media for Personalisation

"There are many positive use cases of synthetic media - speech to text and text to speech tools makes content more accessible for people with disabilities, for example, and text to image and video are also great for translating content into sign language".

– Intellectual Property Expert

Overview

Synthetic media can enable personalisation of a consumer journey by offering tailored content and recommendations based on information gathered on an individual's circumstances and preferences. For example, this may include things like personalised landing pages, documents or 'how-to' videos. This has the potential to elevate consumer engagement and satisfaction by streamlining consumer experiences and encouraging further engagement through tailored recommendations. Synthetic media can also be used to enhance accessibility and digital inclusion through the creation of targeted support tools. Key areas identified during research and stakeholder engagement include:

- **Personalised education and upskilling:** This refers to content for education and upskilling that adapts to individual learning styles. It can be delivered through tools like chatbots or immersive learning experiences that replicate real-world scenarios.
- **Personalised finance:** This refers to AI-generated personalised content and information for financial products or services.
- **Personalised tools for accessibility support:** This mainly refers to customised assistive technologies like speech synthesis or adaptive interfaces for improving user access. An example is creating synthetic audio for those losing their voices.

[4] See section 4.2.3 on Misleading Consumers

Regulatory Considerations

i. Opportunities for digital literacy and inclusion

Synthetic media might be used to improve digital/media literacy and skills in different subject areas. For instance, stakeholders discussed bespoke interactive educational content that could be developed to support media literacy initiatives, relevant to Ofcom's broad media literacy duties. The same might be done with important but complex subjects like promoting financial literacy as a part of the FCA's efforts in financial inclusion, and providing specific risk training for regulated firms and personalised professional upskilling paths. Going beyond personalised training, in the future, individuals may be able to receive real-time educational information relating to specific products or risks. This could be particularly helpful for vulnerable populations and for bridging the digital skills and literacy divide.

Synthetic media tools tailored for accessibility requirements may improve digital inclusion by serving groups who are more likely to be excluded and closing gaps with more cost-effective solutions. This might include developing text-to-audio applications for those who are blind, text-to-video for those who are deaf, or personalising user-friendly interfaces for consumers with limited digital skills. However, in these scenarios the consumer journey and experience would largely lack human elements which could be detrimental, especially for more vulnerable demographics. In some situations, the human aspect or engagement may be important for emotional support, empathy, and care. In some cases, it may also be possible that being encouraged or forced to use AI chatbots may result in worse outcomes for consumers, particularly if the chatbots are low quality or cannot offer the same services or options as a human operator. This may particularly affect vulnerable users, particularly if they are less able or confident to push back and request a human agent.

ii. Personal data processing

As personalised experiences are generally created using large amounts of personal data, there are data protection implications. Whether for personalised education, financial advice or assistive technologies, accurate and effective personalisation relies on the collection and use of information such as students' learning habits, or a consumer's financial details. The data used may come from vulnerable groups like children, making data breaches an increased concern. Developers of personalised experiences will need to ensure that processing personal data for that purpose is compliant with data protection law. For specific applications like financial services, stakeholders noted that the FCA may also need to have clear requirements as accountability would likely rest with the firms to handle the data appropriately.

Data protection laws require that the processing of personal data – which includes the collection and storing of personal data – must be fair, transparent and lawful. For personal data to be processed lawfully, organisations are required to have a lawful basis for that processing.[5]

As well as potentially breaching data protection legislation, a failure by a business to abide by data protection principles may, depending on the circumstances, constitute an unfair commercial practice under the CPRs. For example, a failure by an online platform operator to be transparent about the collection of consumers' personal data and the fact that it may be used for commercial purposes (e.g. targeted advertising, or sharing personal data with other businesses), or otherwise failing to give consumers an informed choice about how their information will be used, may be considered a misleading commercial practice under the CPRs.

[5] https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/lawful-basis/a-guide-to-lawful-basis/

Businesses and advertisers who create or use personalised advertising content when engaging with consumers should take steps to ensure that such advertising is not misleading by action or omission, to avoid infringing the CPRs. Platform operators who publish or otherwise make third party advertising available should also ensure that they have appropriate systems and processes in place to prevent and remove false or misleading information from publication. Where they do not, they risk infringing the CPRs.

4.1.3 Synthetic Media for Creative Likenesses

"In the future, we may see deepfake avatars performing multiple jobs as it becomes so cheap, realistic and effective to do. This can impact the economy and job market with both positive and negative consequences".

– Scientific Research Academy

Overview

Synthetic media can be used to create digital representations, or avatars, modelled to look and behave like real humans. Synthetic avatars are already used in media and gaming industries and stakeholders view them as promising and sophisticated use cases. It was also expected that the likeness of avatars would largely improve. Avatars may provide benefits such as emotional support, or improving consumer experience, and they will likely be used predominantly for envisaging different futures rather than creating deepfakes of the past. Key areas identified during research and stakeholder engagement include:

- **Grief counselling:** This refers to AI avatars, also commonly referred to as 'deadbots', that simulate deceased people based on a collection of their information. Deadbots are intended to offer comfort to those who are grieving by allowing them to imitate interactions with a digital version of departed loved ones.[6]
- **Entertainment personas:** These are digital characters representing real or fictional individuals designed to better entertain and engage audiences, mostly in interactive or immersive environments.
- **Avatars for work:** This refers to human-like avatars performing jobs across the economy, ranging from playing instructional roles to attending virtual meetings on behalf of real humans.

Regulatory Considerations

i. Privacy protection and cyber resilience

A significant regulatory issue with synthetic avatars is the need to use large amount of personal data to train AI models to make avatars more human-like. This raises privacy and cybersecurity considerations.

Stakeholders were concerned about the potential misrepresentation of individuals' digital likeness through unauthorised use of their data, and in the case of grief counselling, there were concerns that it was unclear who has the right to authorise the use of a deceased person's data for synthetic software training. Some stakeholders thought that an individual's digital likenesses should only be used with their explicit permission, whereas others noted the potential need for consent from their

[6] Many stakeholders considered this use case to be a risk rather than an opportunity due to potential psychological impact. These risks are explored in Section 4.2.7.

next of kin if deceased. Others suggested that children should not be able to access such avatars, and that age verification conditions should be required for both the creation of, and access to synthetic avatars. The general consensus was that regulators could explore providing clear guidelines on consent mechanisms for personal data use for creating synthetic avatars. Consent as a basis for processing personal information within these systems may not be suitable, however, since consent can be withdrawn and there may be a power imbalance between those providing the service and those seeking to use it. Developers should use an appropriate basis for processing to fit the purposes for which the data was collected.[7]

ii. Labour market impact and accountability gap

Stakeholders suggested that if avatars for work became more prevalent and could take on more sophisticated tasks, they could have an impact on labour markets and employment opportunities.

Stakeholders also suggested that if in future regulators engage with regulated entities 'AI Avatars' to provide information or respond to compliance requests, there may be more complicated liability and accountability issues. Unlike bots, which typically present information without a human-like presence, avatars can create a stronger perception of authority and reliability, potentially leading users to have more trust in their responses. It could be unclear who should be accountable for false information or failures in meeting compliance requirements; this could be the developers, the company using the avatars, or the critical third-party involved in some of the work. It was noted that this level of technological advancement and societal change is not expected to fully materialise in the medium term (next 3 to 5 years), but that regulators considering a future where synthetic avatars are a part of workforces in multiple industries, should consider updating existing frameworks.

4.1.4 Synthetic Media for the Creation of Digital Twins

"Synthetic media unlocks the potential of easier prototyping of products – organisations will be able to iterate through different examples quicker by prototyping digital models quickly via natural language prompts."

– Research Institution

Overview

Synthetic media could be used to create and update digital twins more easily. Digital twins are virtual representations of physical objects, processes, systems or even biological entities like a person that mirrors characteristics and behaviours. Unlike avatars, digital twins are dynamic, precise functional models focused on simulating and reflecting the real-time state and conditions of the object they represent.

Digital twins are currently primarily focused on industrial applications, but stakeholders expected a shift towards more varied and sophisticated uses in the future. Stakeholders noted applications in healthcare—where digital twins can model a body and help doctors personalise treatments or monitor health conditions—and in wider industries, digital twins could potentially be used for easier product prototyping by creating a digital twin of a specific product and iterating digital models through natural language prompts.

[7] Information Commissioner's Office, Lawful basis guidance, https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/lawful-basis/

Regulatory Considerations

Digital twins may have several regulatory implications:

- Digital twins create opportunities for real-time monitoring and simulation. They could allow users in sectors such as manufacturing to anticipate challenges, improve accuracy, and address risks proactively. By modelling complex systems and visualising outcomes, digital twins enhance transparency and support collaborative decision-making. This technology also has the potential to promote inclusion by making data insights more accessible, which bridge gaps for diverse stakeholders and enable data-driven strategies.
- If a digital twin requires the processing of special category data, such as genetic and biometric data, there are extra considerations under data protection law.
- The accuracy of digital twins will be crucial to allow for informed decision making regardless of whether in health or other sectors. Some stakeholders raised the potential for digital twin model updates to cause unexpected changes in how errors are detected and managed. Model changes can affect the reliability and accuracy of digital twins in simulating real-world behaviours and could potentially lead to incorrect decisions or responses. Stakeholders raised concerns around consumer protection and liability if and when these risks materialise.
- There may be competitive considerations for future digital twin markets specifically in healthcare. If larger companies hold more health data, it may be harder for smaller entrants to enter the market. It may also be difficult for consumers to switch digital twin provider once they have one, especially in relation to data portability and interoperability. Stakeholders argued that consumers should be given sufficient choice, and portability of health information and interoperability of systems could be key solutions.
- As described above, to the extent that such media is being used in connection with the promotion, sale or supply of products to or from consumers, consumer protection law is likely to apply to those businesses who create, use or make such media available to consumers – including online platform operators.

4.1.5 Synthetic Training Data

"More sector targeted datasets will emerge giving smaller businesses a potential competitive advantage over LLM providers, as smaller models produce more accurate outputs."

– Research Institution

Overview

The use of synthetic content to train AI models is emerging as a way to enhance the representativeness and accuracy of AI outputs. Synthetic training data can be generated to ensure AI models are appropriate to a diverse range of scenarios and conditions without needing to use the data of real individuals. It could also help enhance the fidelity of synthetic media outputs. This approach could be particularly beneficial for creating images, text, and other media types where real-world data may be limited or biased. It is important to bear in mind that whilst the closer a synthetic data set is to the real-world data on which it is based, the more likely it is to be useful as a substitute, but also present risks around inadvertent exposure of that real-world data.

Synthetic training data is particularly valuable in reinforcement learning, where AI "agents" learn to interact with graphical user interfaces and other environments. This includes applications in safety-critical areas such as algorithmic auditing and data augmentation for rare events. By simulating a

wide range of scenarios, synthetic training data can help AI systems develop robust and reliable decision-making capabilities.

Regulatory Considerations

i. Competition

As the technology evolves, more sector-targeted datasets could emerge. These specialised datasets may have the potential to provide smaller businesses with a competitive edge over large language model (LLM) providers by producing more accurate and contextually relevant outputs. This shift could democratise access to high-quality AI tools and foster innovation in industries.

ii. Ensuring data quality and safety

One of the primary concerns with synthetic training data is that despite its ability to make samples more representative, it may also be used in the opposite direction. It is crucial to ensure synthetic training data does not contain harmful or biased content. One stakeholder suggested that organisations must implement stringent measures to filter out inappropriate material from their training datasets. This is essential to prevent the propagation of harmful stereotypes or misinformation through AI-generated content.

As described in the CMA's initial review into foundation models, businesses which supply or license AI models in connection with the promotion, sale or supply of products to or from consumers have responsibilities under the CPRs. While it can be difficult to pinpoint responsibility for a particular failure, businesses developing AI models and those in the downstream supply chain – including those that incorporate models in consumer facing products or services – should consider carefully whether they have satisfied their obligations under consumer law. For example, businesses which use synthetic content to train AI models should ensure that the model is not trained to produce false or misleading content and should take such appropriate steps as are necessary to prevent their models from being used by others to harm consumers' economic interests. Businesses should also keep this under review as practices, technology and the law continue to develop.

4.2 Risks

The consensus among stakeholders was that, as synthetic media tools become universally accessible and more user-friendly, there will be an increase in bad actors using these tools for malicious purposes. We identified seven broad categories of risk through our research. The regulatory implications of each of these are explored in this section.

Many of these harmful applications are or will be addressed to some degree by existing legislation or regulation, for example where they fall within scope of Ofcom's Online Safety regime. However, there is a consensus among stakeholders that mass uptake of synthetic media has the potential to exacerbate existing harms, allowing bad actors to undertake harmful activities at pace and scale. Stakeholders stated that regulators will therefore need to adapt to ensure their regulatory guidance keeps pace with technological advancements, the public are protected, and digital markets work well for individuals, businesses and the wider economy.

Across these areas of risk, mitigation and detection methods may be necessary to reduce the harm. Media literacy initiatives to support users to protect themselves from harm may also be an important approach – Ofcom has duties in this area. Mitigation, detection and media literacy are discussed in Section 5.

4.2.1 Circumvention of Authentication

'Currently you need a lot of money and time to effectively use synthetic media for fraud. As the tech becomes cheaper and easier to use, this type of fraud will become more widespread'.

- *Tech Industry Advocate*

Overview

Stakeholders flagged several ways synthetic media might be used to circumvent authentication systems with potentially harmful results. This is typically done by imitating an existing person to either bypass automated authentication systems or persuade a third party to act on their behalf. Synthetic media could be used to generate entirely new identities which can be added to authentication systems, allowing bad actors to access confidential systems and information.

This could cause harm for both individuals and businesses. Synthetic identities could be used to bypass Know Your Customer (KYC) Checks,[8] which are used by businesses to verify the identity of clients to mitigate illegal activities, and liveness checks,[9] which enhance security and prevent fraud in industries that manage sensitive data. By using convincing synthetic identities to bypass security measures, bad actors could directly access confidential and sensitive data, funds, and resources.

Authorised Push Payment (APP) scams are another area where synthetic media could be used to circumvent authentication, where bad actors deceive individuals into authorising payments to

[8] KYC (Know Your Customer) checks are processes used by businesses to verify the identity of their clients, helping to prevent activities like money laundering and fraud. For example, banks use KYC checks to ensure that the person opening an account is who they claim to be.

[9] Liveness checks are security measures used to ensure that a person is physically present during a verification process, helping to prevent identity fraud. For example, during online banking, a liveness check might require the user to blink or turn their head to confirm they are not using a static image.

fraudulent accounts. These scams rely on deception to persuade victims to willingly transfer funds, and synthetic media enables perpetrators to create more realistic and convincing deceptions.[10]

Stakeholders argued that the increasing sophistication and availability of synthetic media will likely lead to more frequent instances of harmful applications over time, posing significant challenges for individuals, industries, and regulators, as well as potentially damaging trust in digital markets and their potential growth.

Regulatory Considerations

i. Protection from scams and fraud

The misuse of synthetic media to circumvent authentication checks could lead to more sophisticated scams (which may be indistinguishable from legitimate activities) and fraud. These might be implemented at a large scale, and effectively target individuals and businesses even with robust protection systems in place.

For individuals, this could lead to financial loss, data loss, and hacking of personal accounts; vulnerable users may be particularly susceptible to synthetic media enabled scams. For businesses these fraudulent activities could lead to mass data leaks, financial losses or increased risks of operational disruption through cyber-attacks, which can contribute to a devaluation of trust in established firms and systems.

Protection from scams and fraud is already a fundamental aspect of many regulatory regimes. However, stakeholders noted that as scams and fraud using synthetic media become more commonplace, regulators will need to ensure their frameworks keep pace and ensure that regulated firms have appropriate measures in place to identify synthetic profiles and mitigate associated harms. As described above, each business in the supply chain for synthetic media is likely to have obligations under consumer law where such media is used in connection with the promotion, sale or supply of products to or from consumers e.g. online platforms which publish or make third party content available to consumers.

ii. Security and resilience

Some stakeholders noted that the ability of bad actors to bypass authentication using synthetic media could pose a serious threat to the security and resilience of key infrastructure. Critical services like telecoms, cloud storage, finance, utilities, and healthcare may be vulnerable to security breaches, which could result in service disruption, data breaches and privacy risks, economic damage, and the erosion of public trust.

The ICO is responsible for ensuring that organisations implement adequate security measures to protect personal data from breaches. Similarly, the FCA, Ofcom and the ICO have responsibilities to ensure that aspects of financial services, communications networks and internet infrastructure are adequately protected from cyber threats and infrastructure failures. The increased threat that synthetic media might pose for security systems will need to be considered in carrying out these duties.

[10] For an example of the type of deception possible, see: CNN, Finance worker pays out $25 million after video call with deepfake 'chief financial officer', 2024, https://edition.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk/index.html

iii. Protection of children

Stakeholders raised concerns that the ability of bad actors to falsely authenticate their identity could pose risks to children whose online safety is dependent on robust authentication systems.

Children are particularly at risk because they can be misled into believing they are interacting with someone their own age if bad actors can falsely authenticate as peers. This may increase the prevalence and ease of online grooming or exposure to harmful content. Similarly, synthetic media may enable children to circumvent age authentication and access inappropriate content online.

The Online Safety Act 2023 places duties on platforms to take steps to protect users from illegal content and content that is harmful to children. Ofcom is considering the risks posed by deepfake content as it continues to implement its Online Safety regime.[11]

Additionally, children are likely to be considered vulnerable consumers under the CPRs and businesses should consider the potential impacts on children where their commercial practice reaches or is addressed to them e.g. where synthetic media is being used to promote products which are likely to be used by or appeal to young children, businesses which create or use such media to advertise products must ensure that children (or their parents) are not misled by the content of such advertising.

iv. Market evolution

Platforms may require more personal data to be provided for their authentication systems in order to mitigate against the malicious use of synthetic media, with implications for current data protection and consumer protection regimes.

As authentication software and tools evolve to account for potential security risks from synthetic media, smaller firms may not have resources to invest in state-of-the-art systems in response to the increased threat. These smaller firms risk not effectively protecting their users and could be at a significant disadvantage in comparison to larger firms if they are unable to protect consumers to the accepted industry standard.

Additionally, synthetic media may allow for the widespread application of bad-faith business practices, such as using fake profiles to artificially boost platform user bases. Businesses who implement these practices may have an unfair or unwarranted perception of market dominance as a result.

The CMA may need to explore the potential implications in the context of its new consumer protection powers established by the Digital Markets, Competition and Consumers Act 2024.

[11] See: Ofcom, Deepfake Defences: Mitigating the Harms of Deceptive Deepfakes, 2024, https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/deepfake-defences/

4.2.2 Disinformation

'Bad actors are watching televised local parole board meetings in the US and using synthetic tools to mimic them and create misleading content. If this is already happening at a local level, it's hard to imagine how things might evolve in the near future…'

- Authenticity Specialist

Overview

Synthetic media can be used to create misleading or deceptive content, which could result in political or ideological disruption and eroding trust in media. Such content might include synthetically altered video, audio, still images, or entirely artificial outputs. For example, 'deepfakes' of national politicians have been shared on social platforms and can be created to either push a specific political agenda, or to create confusion and undermine trust more generally.[12] Some stakeholders argued that such misleading content could influence voter perceptions and election outcomes.

The continued use of synthetic media in this way may risk causing harm to democratic processes and to ensuring a well-informed public. Additionally, as synthetic media becomes more accessible, it will be easier for bad actors to create disinformation that disrupts politics and democratic processes at a local level, as well as on a national and global scale. Local bodies may have insufficient infrastructure, funds, and technology to effectively implement detection measures, making it harder to identify and disprove synthetic disinformation at the local level.

Synthetic disinformation is not limited to politics; it could be used to proliferate misconceptions across all sectors including health, finance, public safety, and criminal justice. Highly convincing falsified information on these subjects can lead to significant harms for individuals, businesses, and society more broadly. For instance, stakeholders suggested that we must be cognisant of the risks of synthetic media being used to mislead financial markets with the intention of taking advantage of resulting asset price movements, with its improved ability to manipulate markets.

Regulatory Considerations

i. Authenticity and trust

Synthetic disinformation can be highly convincing and difficult to distinguish from authentic content, particularly when it has been used to mimic the style and branding of trusted sources. Consumers may struggle to differentiate between real and false information, which ultimately could undermine trust in media and institutions more generally. Providing support to consumers to identify and avoid disinformation will be critical to protect trust in accurate and authentic media, and to ensure that positive outcomes from such media are delivered to the public.

Similarly, synthetic media may complicate the process of authenticating sources taken from social media or anonymous whistleblowers, with implications for broadcasters and media outlets. Media firms may have to implement more stringent verification standards to ensure the authenticity of news content, in line with standards for broadcast content, which Ofcom oversees.

---

[12] See for example: BBC, How AI and deepfakes are changing politics, 2024, https://www.bbc.co.uk/reel/video/p0hkflt4/how-ai-and-deepfakes-are-changing-politics. However, a number of stakeholders argued that deepfakes have had a more limited impact on the 'year of elections' than might have been expected or feared.

ii. Privacy and targeting

Disinformation campaigns often exploit personal data to target or mimic specific individuals, raising significant privacy concerns and potentially resulting in severe repercussions including defamation.

In order to process personal data lawfully, an organisation must have an appropriate basis for doing so. It is unlikely that an organisation seeking to disinform would have a valid lawful basis to process personal data to create a product or service for that purpose. Similarly, it is unlikely that organisations seeking to disinform would meet transparency requirements under data protection law. In addition, as described above, businesses may also infringe the CPRs where they do not abide by data protection principles in connection with the promotion, sale or supply of products to or from consumers.

iii. Societal disruption

Stakeholders suggested that synthetic disinformation could be used to cause widespread societal disruption leading to tangible harm for individuals and markets. For example, bad actors could use synthetic media to mislead financial markets and commit market abuse under the FCA's Market Abuse Regime, including through creating false content that aims to boost the price of a stock as part of a pump-and-dump scheme. Such conduct could also have wider implications for market stability. Political disinformation could lead to democratic disruption, influencing voter perceptions and election outcomes. Insofar as this disruption stems from a foreign actor, this may have implications for Ofcom, given the foreign interference offence (FIO) offence in the Online Safety Act 2023. More generally, media literacy may be key to reducing the impact, which is relevant to Ofcom given its media literacy duties, including in relation to mis/disinformation.

While some of these disruptions are already addressed by current regulatory regimes, the mass proliferation of convincing synthetic disinformation will undoubtedly increase the burden on parties responsible for identifying and removing such content.

4.2.3 Misleading Consumers

'The handling of advertisements that mislead consumers is already well established, so synthetic misrepresentation would be treated as a subset of that.'

- Advertising expert

Overview

Stakeholders suggested that synthetic media has the potential to disrupt commerce by misleading consumers about goods or services (which may be by action or omission). Artificially generated or enhanced advertisements and promotions could mislead consumers and create fake expectations of products, events, or services. Similarly, synthetic media can be used to create false endorsements by public figures. Synthetic media could also be part of the product itself, for example synthetically produced recipes, patterns, or instructions. These could be advertised as legitimate but may be impossible or nonsensical to use once purchased.

Commercial practices that are misleading can financially harm consumers, put legitimate businesses at a disadvantage, and lead to the erosion of trust in advertisements more generally. False endorsements can cause reputational damage to impersonated individuals, as well as harm to consumers and markets. As synthetic media improves and is scaled, it may become increasingly difficult for even savvy consumers to identify synthetic or enhanced products, amplifying these issues.

As described above, all businesses in the supply chain for advertising which concerns the promotion, sale or supply of products to or from consumers are likely to have obligations under consumer law.

Regulatory Considerations

i. Consumer protection

When synthetic media is used to create false or misleading advertisements, or to create falsified products or services, it has significant negative implications for consumers. Consumers may be misled into purchasing or investing in products they may ordinarily not have, leaving them at a financial disadvantage and eroding their trust in future advertisements. This may also create an unlevel playing field for businesses that are not misleading consumers. Businesses which create or use advertising to promote products to consumers must ensure that the advert is not misleading by action or omission, and businesses which publish or make third party advertising available to consumers – e.g. online platforms – must take such steps as are necessary to prevent and remove false and misleading advertising from publication – where they fail to do this, they risk infringing consumer law.

The CMA, Ofcom, ICO, FCA and Trading Standards/DETINI in Northern Ireland all regulate to protect the rights of individuals, including through the CPRs. Current advertising standards prohibit content that misleads consumers. There could be challenges with enforcement where the source of the material is not clear, in which case enforcers may need to consider using powers to require content to be removed. However, the role of platform operators will continue to be an important aspect of effective enforcement – and the CMA expects platform operators to abide by their responsibilities under consumer law.

Where businesses use synthetic media to create or cause the publication of false, deceptive or other misleading content they risk infringing consumer law. The CPRs prohibit commercial practices which mislead by action or omission, and which cause or are likely to cause the average consumer to take a different decision as a result. Accordingly, where advertisers or other businesses create or use false or misleading synthetic media to promote, sell or supply products to or from consumers, this could breach the CPRs. Further, the CMA has taken consumer enforcement action concerning unlabelled advertising by 'influencers' who are paid or otherwise rewarded for talking about consumer products on online platforms such as Instagram and is investigating issues with the publication of fake reviews by third parties on Google's and Amazon's sites. Platform operators need to take appropriate steps to prevent and remove fake and misleading third-party content from publication. Where they do not, they risk infringing the 'general prohibition' in the CPRs which requires businesses to abide by the requirements of professional diligence (meaning honest market practice and good faith in the business's field of activity). Where a business such as an online platform operator contravenes these requirements and this distorts consumer behaviour, they may infringe the CPRs. The CMA published compliance principles for social media platforms that addresses the requirements in the area of unlabelled advertising. The FCA also published guidance on financial promotions on social media, which highlights financial promotions on all advertising channels should be fair, clear and not misleading, and support consumer understanding.

In addition, the FCA's Consumer Duty requires firms to play a greater and more proactive role in delivering good outcomes for retail customers, including (in some circumstances) those who are not direct clients of the firm. Firms are required to act in good faith, avoid causing foreseeable harm, and enable and support retail customers to pursue their financial objectives. The FCA's Principles for Business are also relevant. Where firms are not conducting retail market business and the Consumer Duty does not apply, firms need to pay due regard to the interests of their customers and treat them fairly (Principle 6) and communicate information in a way that is clear, fair and not

misleading (Principle 7). Under the Principles for Business, the FCA has guidance which sets out what firms should be doing to treat customers in vulnerable circumstances fairly, with the aim that vulnerable consumers experience outcomes as good as those for other consumers.

ii. Market protection

If misleading synthetic advertisements and endorsements become more widespread it could significantly disrupt commercial markets. Online advertising could become saturated with misleading advertisements if they can be produced quickly and at scale, meaning consumers may not have sufficient exposure to legitimate advertisements. This could lead to a reduction in consumer engagement with legitimate businesses, reducing incentives for companies to offer quality goods and services.

Similarly, the practice of astroturfing—the creation of synthetic support and enthusiasm for a product—can undermine fairness, make investigation more complex, and distort competition by reducing rewards for businesses that offer quality products. This practice is already likely to be prohibited under the CPRs as a misleading commercial practice and businesses – including advertisers – who create or use false and misleading content to promote products to consumers are at risk of infringing the law. Note that Government has recently introduced new 'banned practices' as part of the Digital Markets, Competition and Consumers Act 2024 which are expected to come into force in 2025. The new legislation will make it illegal, in all circumstances and irrespective of the potential impact on consumers, for businesses to submit or commission others to submit fake reviews or concealed incentivized reviews. Businesses who publish this content – including online platforms – will also have an express legal responsibility to take such steps as are necessary to prevent and remove this false and misleading content from publication. The CMA and other public enforcers such as Local Authority Trading Standards services will have the power to enforce the new prohibitions when they come into force in 2025.

Synthetic media also enables copycat firms to duplicate effective advertising practices and techniques. This may lead to consumer confusion, infringing consumer law, and put good faith firms at a competitive disadvantage.

4.2.4 Copyright Infringement and Creative Incentives

'Original UK content is being used to train synthetic models without authorisation or attribution, and this practice is devaluing UK content.'

- UK Content Creator

Overview

Stakeholders noted legal and ethical concerns around copyright infringement and fair use when synthetic media is trained on existing intellectual property (IP). Synthetic media can convincingly imitate original works, which may be considered infringement of intellectual property of the original creator. Synthetic media can be used in this way to mimic various content, including written text, designs, music, and artistic works, with implications for a variety of sectors.

Improvement in synthetic media capabilities may enable the production of more convincing imitations at pace and scale. Creators and creative markets may face increasing challenges in protecting their intellectual property rights and the value of original works as a result. Some stakeholders argued that, even if the synthetic media is not similar to an original work, it may still be an infringement if it is produced by a model trained on that original work.

Regulatory Considerations

i. 'Copycat' content

Some stakeholders noted that UK media faces significant challenges with the proliferation of synthetic replications of creative works, particularly as the software used to produce synthetic media can be trained with original content to produce similar or identical outputs. This 'copycat' content could fragment audiences by offering niche, tailored versions of recognisable IP, driving audiences away from legitimate content and potentially reducing market share and influence of the original creator.

This practice could diminish the exclusivity and profitability of original IP, reducing incentives for individuals and firms to produce and invest in new creative works. Consequently, facilitating growth and investment in the UK creative industries may be challenging for regulators, as could ensuring audiences have access to a wide range of high-quality, original content.

Stakeholders raised similar concerns in relation to synthetic media that covers topics of societal, cultural, or historical importance. For example, chatbots replicating historical figures at heritage sites can enhance visitor experiences, however they may also lead to a potential loss in traditional art and culture by oversimplifying complex historical narratives and reducing audiences' engagement with authentic sources of information. This has broad implications for cultural and economic incentive structures.

ii. Unfair commercial practices

Some stakeholders suggested synthetic media could be similarly implemented to replicate business models, domains, and contract text, allowing bad actors to effectively imitate legitimate businesses.

This practice could distort markets by creating confusion among consumers, who may struggle to distinguish between genuine and fake businesses. For legitimate businesses, it could undermine trust, reduce market share, and discourage investment in innovation due to an increased risk of imitation and unfair competition. If 'bad actors' replicate legitimate business models and in fact mislead or are likely to mislead consumers in their practices, then this may breach consumer protection law. As described above, all businesses in the supply chain for the promotion, sale, or

supply of products to or from consumers are already likely to have obligations in this space under consumer protection law – including advertisers and online platforms which publish or otherwise make false and misleading content available.

A potential future implication is whether producers of non-synthetic media will still be incentivised to produce original content given the low costs of replications, or equally they may have to increase the price of their outputs. One stakeholder discussed this in reference to chatbots replicating historical figures.

4.2.5 Harmful/Illegal Content Creation

'Image abuse is widely accessible: most users of the internet can create this in 5 minutes…'

- Tech Expert

Overview

Synthetic media can be used by bad actors to harass, defame, or otherwise harm individuals. For example, stakeholders noted that 'deepfake' technology has been used to create nonconsensual sexual images using a victim's likeness. The resulting content may be used for extortion, blackmail, or to cause reputational damage. Synthetic media can also be used to produce artificial child sexual abuse material (CSAM), exacerbating the individual and societal harms caused by such content.

Similarly, synthetic media can be used to create other forms of illegal content such as terrorist content and hate speech. These outputs may defame or falsely implicate innocent parties, whilst contributing to wider harms.

Regulatory Considerations

i. Implementing preventative measures

Synthetic media has the potential to raise the risk of harm from illegal content and the risk of harm to children – for example by increasing the volume of such content available. The Online Safety Act 2023 requires regulated user-to-user services like social media platforms and regulated search services to, among other things, carry out risk assessments in relation to illegal content or content that is harmful to children on their services; and to take appropriate measures to address these risks. This may include consideration of AI-generated deepfake content, where it is in scope of the regime. Ofcom will issue Codes of Practice, which set out the measures that services – both large and small – can take to ensure compliance.

Ofcom will seek to ensure that regulated services take the necessary steps to protect their users from harms associated with deepfakes where this is in scope of the regime, including by taking appropriate enforcement action.

One thing that stakeholders noted is that, if varying approaches to protecting consumers emerge, those with the infrastructure and resources to introduce more robust standards of safety could gain market dominance, leading to a competitive advantage. It will be essential to consider how small and novel platforms can securely enforce consumer protection measures and compete effectively with larger firms.

4.2.6 System Errors and Lack of Transparency

Overview

The nature of the models used to produce synthetic media can lead to inaccuracies, which pose risks even when this technology is used by 'good' actors. Stakeholders noted risks including the creation of unintentional results due to AI model 'hallucinations'[13] a lack of understanding of how data is used, and decisions or recommendations that are unexplainable due to the black-box nature of many AI systems.[14]

For example, AI-generated financial advice could lead to suboptimal or risky financial decisions if users are unaware of how the AI is processing their information or if the AI system cannot explain its recommendations. Similarly, algorithms might employ dark patterns to exploit user biases and encourage or otherwise manipulate consumers to make decisions that optimise commercial gains for a firm rather than what might be in the best interest of the consumer, could mislead consumers and may breach consumer law. In education, a personalised learning path may be biased or produce unfair outcomes if the algorithms are not transparent or well-understood.

Regulatory Considerations

i. Transparency and accuracy of software

Actors in the synthetic media supply chain, including advertisers, content creators, and platforms have responsibilities under existing consumer law to protect consumers from harm, which may inadvertently be breached by system errors. To anticipate and mitigate these breaches, it is vital that algorithms are transparent, explainable, and accurate. The ICO and the Alan Turing institute have created guidance for organisations to help them explain the processes, services and decisions delivered or assisted by AI to the individuals affected by those decisions.

Any practice that materially distorts or is likely to distort the economic behaviour of the average consumer – including using synthetic media - could breach the CPRs' requirements of professional diligence, as well as being a misleading commercial practice. This is the case regardless of whether the practice was intended to mislead consumers. Accordingly, businesses which create or use such systems should take such appropriate steps as are necessary to ensure that the design of their systems does not distort consumers' economic decision-making – for example, by hiding information which consumers need to take informed decisions about products or providing it in an opaque or confusing manner to the consumer. Further, the application of pressure by a business to a consumer in a way which significantly impairs or is likely to significantly impair their freedom of choice in relation to a product could amount to an aggressive practice under the CPRs e.g. pressuring consumers into taking quick purchasing decisions.

[13] An AI hallucination is when an artificial intelligence system generates information or answers that are incorrect or nonsensical, despite appearing confident and plausible.

[14] This refers to the difficulty in understanding or explaining how AI systems make decisions or predictions, due to their complex and opaque internal workings.

4.2.7 Psychological Impact

Overview

A major risk, particularly in relation to synthetic avatars, is the potential psychological impact on users. This is particularly pertinent when such avatars are used in sensitive circumstances, such as for grief counselling. Stakeholders acknowledged that as technology evolves, the improved accuracy of synthetic services that simulate deceased people, also known as 'deadbots' could provide meaningful bereavement support. However, in the longer-term users may start to view these avatars as 'real' disrupting the natural grieving process and potentially leading to emotional dependency. Similar issues apply to entertainment personas, where avatars may blur the line between reality and digital recreation, leading to psychological harm for users.[15]

Regulatory Considerations

i. Transparency and consent

Some stakeholders suggested that regulators may be well positioned to ensure that service providers are transparent to their consumers about avatar capabilities and limitations and the fact that synthetic avatars are not real people. To unlock the benefits of avatars it will be crucial that users fully understand they are interacting with artificially generated characters to avoid psychological harm. Thus, regulatory frameworks may need to evolve to account for the increased difficulty in identifying synthetic avatars and promoting and protecting media literacy will be critical in realising the benefits of this technology.

Stakeholders also suggested that regulators may have a role to play in ensuring that service providers adequately protect users from psychological harm whilst engaging with their services.

ii. Linkages to Online Safety

Ofcom has also set out how the Online Safety Act 2023 will apply to generative AI and chatbots in an open letter to UK online service providers.

[15] See, for example: Blake Montgomery, Mother says AI chatbot led her son to kill himself in lawsuit against its maker, The Guardian, 2024, https://www.theguardian.com/technology/2024/oct/23/character-ai-chatbot-sewell-setzer-death

5. Detection of Synthetic Media[16]

Our stakeholders expected that the future of synthetic media and deepfakes would involve an increase in the availability of tools for content generation, and the fidelity of that content. AI generated or manipulated media would form a large proportion of the media that people see and share, and it would be increasingly difficult to distinguish from other, 'non-synthetic', media.

The majority of stakeholders argued that some form of system for distinguishing the objectively real from the artificial or manipulated will be needed, particularly as creation tools are democratised and used for a range of purposes. This view is reflected in legislative changes in the US and Europe. For example, in 2023 US President Biden mandated that a system be developed to guarantee the veracity of official content (to protect against mis-and-disinformation promulgated by bad actors). Similarly, the EU AI Act - which came into force in August 2024 - requires providers and deployers of certain AI systems to be transparent that outputs from those systems are detectable as artificially generated or manipulated.

Several methods are proposed to identify synthetic media when it is created. These include:

- Watermarking
- Data provenance

Other techniques focus on identifying synthetic media when it is received or shared. These include:

- Reputation-type-measures
- Automated detection through the use of software (including AI)

Stakeholders also talked about the importance of awareness and education, which focused predominantly on media literacy.

5.1 Watermarking

Watermarking involves adding an imperceptible mark to content that signals whether it is synthetic or genuine. In the case of images and videos, one means of doing this is by making minute alterations to the pixels of the content, in a way that cannot be seen by the naked eye. This could also involve encoding information within the metadata of the file.

Watermarking allows the recipient of the media to be alerted if it is synthetic, because the information is inserted at the point of creation. For example, Google's SynthID for text and Veo for video both embed characteristic features into the media. In future, a browser extension might trigger a warning to say that metadata shows a piece of media was artificially generated.

There was a consensus among stakeholders that watermarking was unlikely to be the dominant method of denoting and identifying synthetic media in the long term. They believed this was partly because watermarking might impact the quality of the media itself, and partly because it is likely that those who utilise synthetic media and deepfakes for malicious purposes would seek to avoid watermarking. For example, using tools that do not add watermarks, or using methods to remove watermarking (through compressing the file, changing the encoding etc).

[16] In this section we focus on what we heard from stakeholders interviewed. Ofcom's paper on 'Deepfake Defences' provides further detail on some of these measures: Ofcom, Deepfake Defences: Mitigating the Harms of Deceptive Deepfakes, 2024, https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/deepfake-defences/

Stakeholders also noted that watermarks might need to be updated to maintain applicability as models advanced, and that by changing even 'non-synthetic' media to add the watermark it has now become distinct from the original and may qualify as synthetic media (as altered by AI).

5.2 Data Provenance

Stakeholders highlighted the Content Authenticity Initiative's effort in promoting the C2PA open standard of data provenance. The C2PA was formed through an alliance between Adobe, Arm, Intel, Microsoft and Truepic. Its purpose is certifying the source and history of media content.

By holding information about a piece of content separate to that content, users can verify if the version they received is the same as the original, or if it has been manipulated. Since the certification information about the media is held separately from the file itself, data provenance measures do not share the same issues as watermarking around circumvention and removal. Media can be certified as either artificially generated or real, and when file information is compared to the certification information, people are better able to decide whether that media can be trusted.

Regulatory Considerations

Stakeholders argued that, for at-creation measures to be useful and successful, there needs to be a critical mass of acceptance. If the alert from any identification system is not recognised by the receiver, then regardless of the assertion that media is real or altered, it will be treated the same.

As the market for detection and identification systems capable of highlighting when received media has been altered or created by AI grows, the CMA will have a role in ensuring that incumbent large technologies firms don't leverage their market position to unduly favour their own systems.

For each of the digital regulators, any creation of open standards and codes for watermarking or data provenance would be impactful, as they may have direct relationships with relevant legislation and regulatory regimes.

For those organisations developing identification and detection platforms, then there are several important considerations around the use of personal data (which the ICO regulates) with those platforms. An example might be a certification or data provenance system which processes the name and geolocation of the creator of the synthetic media. Development in line with the principle of data protection by design and default is a legal requirement under the UK GDPR. This means that data protection must be a design consideration, and that controllers will need to implement appropriate technical and organisational measures, such as pseudonymisation, which are designed to implement data-protection principles, such as data minimisation, in an effective manner and to integrate the necessary safeguards into the processing to meet the requirements of this Regulation and protect the rights of data subjects.

5.3 Behavioural Analysis

Firms that operate platforms, particularly if they store and facilitate the sharing of media, could take a role in examining the nature of accounts which might be using synthetic media or deepfakes in malicious or illegal ways. Considerations may include, how long an account has been established, which other accounts it engages with, and even how "viral" its content is. Platforms can collect aggregate information about the behaviour, history, and actions associated with an account or source of a piece of media in order to establish whether it should be trusted. The reputation of accounts or media might be determined by considering previous behaviour or connections with

other accounts or media that have acted in a way contrary to what is acceptable, contrary to the standard the platform expects. Platforms can then use behaviour or reputation to restrict the access those accounts have to perform certain actions, such as sharing media, or interacting with accounts which don't follow them. Our stakeholders were unsure as to the possibility of identifying the origin of malicious synthetic media or deepfakes across platforms, since it would require co-operation between those platforms.

Most of our stakeholders believed there is not enough impetus for content distribution platforms to engage in this type of inspection. Whilst some stakeholders believed it is possible to do when certain thresholds were met, they also suggested that action was unlikely unless platforms are compelled to act.

Regulatory Considerations

These measures have been used for other types of content and in different platforms. Indeed, moderation of content and mitigation of harms online are key issues for online regulation.

Since these types of identification and detection measures are platform-level, there are key issues around motivating those platforms to act. Ofcom is committed to doing its part to curtail the circulation of this malicious content. There will be circumstances where the duties on services under the Online Safety Act 2023 will require consideration of some types of deepfake (though not all types). Behavioural analysis may be one possible approach services could choose to take.

5.4 Automated Detection through the use of Software (including AI)

Automated detection systems operate at the point that media is received. They inspect it and create an estimate of how likely it is to be synthetically created or altered. These detection systems (often themselves using AI) examine key identifiers in the media to determine whether it has undergone modification. These might include looking for inconsistencies in lighting, looking for boundaries where two sources may have been blended, or differences in mouth movements versus the sound being generated. By combining multiple analyses, an aggregate "risk" score can be generated. By their nature, detection systems operate in an arms race with creation software. As creation software improves, so too must the detection software to remain effective.

Regulatory Considerations

Organisations who offer synthetic media and deepfake detection either as a software product or a live service could see improved effectiveness through the sharing of data and datasets. This would support the work of the FCA in addressing consumer harms and scams and relates to the ICO's code of practice for data sharing.

Any processing of personal data where it relates to an identifiable natural person within automated detection systems will need to be done in a manner which is compliant with data protection law, which the ICO regulates.

As described above, the CMA has taken a number of consumer enforcement cases concerning economically harmful illegal content on online platforms. Where an online platform publishes or otherwise makes available content originating from third parties, the CMA considers the operator of the platform must take such appropriate steps are necessary to prevent and remove false and misleading content – including via suitable proactive detection and removal measures. While there is

no ‘one-size-fits-all’ approach, the CMA considers that, in practice, there are certain measures which all platforms which publish third party content should take.[17]

5.5 Media Literacy and Education

Ultimately, what matters is whether people can trust that what they’re reading, seeing or hearing is a fair representation of reality. For a video, what matters is whether what is depicted actually happened, and whether this is clear to viewers. Identification and classification systems are therefore designed to either make it clear that something is (or likely to be) real or artificial, so that individuals can be equipped with enough knowledge to determine what is real or not. If, in future, there are instances where there is no clear and obvious verdict either way, people may need to change their attitudes and trust when receiving or viewing media. This represents a fundamental change in how people engage online.

Whilst some stakeholders suggested that such ‘media literacy' this places the emphasis on the individual (rather than the platform) to inspect and detect synthetic media, there was widespread agreement that education will be needed for people to be able to understand the issues around synthetic media and deepfakes, and to an extent anticipate them. They also argued that education would need to be kept up to date to maintain effectiveness as models change.

It is worth noting that media literacy can also be supported by the platforms. Ofcom has published a set of Best Practice Design Principles for Media Literacy that sets out how platforms can promote media literacy through on-platform interventions. Ofcom has also previously published papers on understanding generative AI in the context of media literacy.

Regulatory Considerations

Ofcom has an important role as the broadcast media regulator and online safety regulator. As set out in its Three-Year Media Literacy Strategy, Ofcom has a number of duties in relation to Media Literacy, including around building awareness, encouraging technology and systems which allow users of regulated services to protect themselves and others online, and publishing a media literacy strategy and annual statement.

Ofcom will seek to fulfil its media literacy duties through a combination of research, evidence, and evaluation, engaging and working with platforms, and working and delivering through a range of partnerships.

[17] See: Competition and Markets Authority, Consultation on improving price transparency and product information for consumers – response from the Competition and Markets Authority (CMA), https://assets.publishing.service.gov.uk/media/65294a96697260000dccf7e0/CMA_response_to_price_transparency_consultation.pdf

6. Future of Regulation

The increased prevalence of synthetic media and deepfakes may present new challenges for regulators, as well as policy questions for Government and Parliament. Through our research we gathered insights into how stakeholders felt that regulators should respond to the emergence of these new technologies and the potential implications for digital markets.

Regulators and Government have a substantial role to play

Stakeholders agreed that regulators have a substantial role to play in enabling the opportunities and mitigating the risks that will arise with the emergence of synthetic media, with many encouraging strong regulatory interest and action. Many stakeholders noted that now may be an opportune time for regulators to act, due to the recent growth of AI, and the passing of pivotal legislation – including the Online Safety Act 2023 and the Digital Markets, Competition and Consumers Act 2024.

While stakeholders welcomed recent developments in legislation and policy, many highlighted that, in time, additional legislation and policy may be required to keep pace with the rate of technological change. Some also suggested that current legislative frameworks will require updating in response to the emergence of some practices. For example, legislation which outlines powers to enforce against smaller companies whose headquarters are not in the UK, but target UK customers.

Stakeholders welcomed plans for the new Data (Use and Access) Bill and the introduction of new legislation relating to AI, saying that they believed the proposed Bills have the potential to 'fill in the gaps' in relation to existing legislation, and account for future developments of AI and related technologies. Stakeholders also highlighted the critical role of regulators in using their expertise to inform the development of relevant legislation, guidance, and codes of conduct, and in exercising the powers derived from existing and forthcoming legislation.

Furthermore, stakeholders emphasised the risks associated with regulators failing to act on harms or failing to act rapidly enough. Stakeholders argued that the Government may need to look further ahead when designing legislation–potentially going as far as considering the technology landscape 20 years in the future. Some gave the example of the Online Safety Act 2023 coming into force around 20 years following the emergence of major social media platforms like Facebook/Meta, as an example of perceived delayed action. Finally, numerous stakeholders note that bad actors are unlikely to comply with regulatory measures, highlighting that swift and effective enforcement action is critical to set precedents and deter future malicious practices.

Regulators must collaborate

To ensure effective enforcement—alongside other regulatory duties such as effective policy design, horizon scanning, identification of harms, etc—stakeholders highlighted the importance of regulatory collaboration.

The DRCF is a key vehicle to ensuring continued collaboration. DRCF member regulators already collaborate on several workstreams with relevance to AI. This includes exploring the AI assurance landscape, conducting research into the consumer use of generative AI and providing an informal advice service for queries on AI and Digital Services through the DRCF AI and Digital Hub. This paper highlights how synthetic media falls under the remit of several regulators and stakeholders confirmed that understanding emerging technologies such as synthetic media requires a cross-regulatory approach. This helps to ensure regulatory coherence, whilst delivering the best outcomes

for the public. The emergence and growth of synthetic media will have an impact across society and the areas impacted are much wider than those within the remits of the DRCF member regulators. For example, synthetic media may impact issues as varied as workers' rights, and intellectual property. Therefore, stakeholders argued that collaboration between regulators, government, academia, industry and civil society is essential. This could also include working with international bodies. Stakeholders also argued that clarity on each regulator's role in relation to synthetic media is also important, for both the regulators themselves and for third parties and to help ensure regulatory responses are swift and coherent.

Regulators must be alert to harms – but also embrace the positives

Stakeholders overwhelmingly outlined how synthetic media, and deepfakes in particular, have potential to cause great harm at a societal and individual level. Regulators should be prepared to tackle these harms through existing and future regulatory frameworks. However, several stakeholders also outlined the positive use cases for synthetic media, noting that these should be factored into regulatory approaches, particularly to ensure that emerging technologies such as synthetic media can contribute positively to economic growth.

As set out in Section 4, synthetic media is expected to give rise to both risks and opportunities. Stakeholders argued that when regulators consider interventions and approaches to synthetic media, they should consider proportionality. During any assessment of proportionality, the benefits of the technology to consumers, society, and the economy should be considered. A fair and balanced approach to regulation will benefit the economy, businesses, and people—and should build confidence in the decisions of regulators.

7. Conclusion and Next Steps

While synthetic media is still a fairly new technology, as it becomes more sophisticated and widespread, regulators and the Government will need to ensure they are equipped to deal with the potential challenges as well as ensuring the benefits of the technology to people, businesses and the economy are realised.

The growth of synthetic media may raise some significant policy issues for Government, including clarifying the circumstances in which the creation of synthetic media of an identifiable individual is likely to be permissible without their consent; whether the media should be watermarked or otherwise flagged; whether existing laws on misrepresentation, and to tackle fraud, are adequate to address associated risks; and where liability should fall in the case of harms, to name a few. Addressing the above could allow people to be better protected and enable positive uses of the technology to materialise in a safe way, maximisation economic and societal benefits, while mitigating the risks.

Through our research, we have highlighted a number of these opportunities and risks, and explored how the effective detection and mitigation of synthetic media and deepfakes might be achieved. However, Government policy and legislation evolves, and several regulatory considerations may be relevant in the future. These include the protection of consumers and markets through measures such as transparency and consent mechanisms; the use of sensitive personal data particularly in the development of personalised content; security and resilience risks to financial products; and threats to democracy and society more broadly, including trust in the media.

In the short term DRCF member regulators will need to consider how to effectively apply existing regulatory frameworks to maximise the economic and societal benefits while minimising risks, and in the longer term consider if and how regulatory frameworks may need to adapt. Ultimately the DRCF member regulators, through the HSET programme and beyond, will:

- Closely follow the development of synthetic media and deepfakes as they continue to evolve.
- Continue to engage with industry, government, academia, and others on the subject.
- Bring together stakeholders working in this and related areas through our events and projects.
- Scan the horizon collectively and individually to ensure we remain on the front foot with synthetic media and other emerging technologies.

Taking account of the regulatory considerations discussed in this paper, each DRCF regulator will continue to carry out work on synthetic media, where appropriate, in their respective remits, whilst also exploring opportunities for further collaborative efforts through the DRCF.

Annex: Scenario Analysis

Using the insights from our stakeholder engagement and workshops we developed a set of scenarios to explore how synthetic media could evolve, and the associated implications for regulators. Stakeholders expressed a wide range of views, from optimistic to highly negative, about the future of synthetic media and its impact. Each scenario has been designed as a fictional summary of the future a set of interviewees viewed as plausible in 3 – 5 years. Considering plausible futures, as set out in these scenarios, helps us to consider the regulatory implications of various possible futures.

Scenario 1: Ineffective mitigation leads to greater emphasis on media literacy.

i. Landscape Evolution

There is continued investment in and development of AI products. As a result, synthetic media can be generated quickly, easily, and at a low cost. High-quality synthetic media becomes commonplace across all markets and is regularly produced by legitimate parties and bad actors at local and global levels.

Despite ongoing research and investment in mitigation techniques, such as watermarking and provenance, these are not universally adopted by tech companies and platforms. Consequently, there is no uniform approach to labelling or identifying synthetic media, leaving users inadequately informed and unprotected from potentially harmful content.

To counteract the lack of effective technological mitigation, there is an increased focus on media literacy. Educational programs and resources are introduced and promoted by the Government and third-party organisations to empower users to make informed decisions about content authenticity and take precautions when engaging with potentially harmful outputs, while current regulatory approaches continue to address the most harmful or illegal forms of content.

ii. Consumer Attitudes

Users are increasingly exposed to synthetic media of high quality and are unable to easily discern the authenticity of content. However, the prevalence of media literacy initiatives allows users to approach content with caution despite the lack of clear authenticity indicators. Users are more aware of the risks of synthetically produced or augmented content, and are less susceptible to mis/disinformation, commercial misrepresentation, fraud, and scams. They are also empowered to flag potentially harmful or false content and have the confidence to add community notes where possible.

Media literacy initiatives provide users with a robust understanding of the landscape and how to navigate it. Consequently, users have the confidence to engage with a variety of content including potentially beneficial forms of synthetic media, and experience positive outcomes from this engagement.

iii. Market Outcomes

The emphasis on media literacy education over technological mitigation means that, while users are generally able to discern when content might be synthetic, systems and processes are not. As a result, authentication systems remain vulnerable to circumvention by bad actors and threats to security and resilience persist. Despite adequate protection at the individual level, critical infrastructure remains at risk and large-scale harms like data breaches and service disruptions continue to impact consumers and markets.

Effective implementation of media literacy initiatives is not a silver bullet for harm and some users still experience negative outcomes despite being able to critically assess content. This is particularly salient for content such as nonconsensual sexual imagery or hate speech. Although users can discern when such content is falsified, they may still suffer emotional distress at having been exposed to it. It is therefore critical that harmful content continues to be addressed through existing regimes.

Scenario 2: Successful mitigation reduces societal harms

i. Landscape Evolution

There is an increase in the quality and prevalence of synthetic media across all platforms due to continued investment and development. Synthetic media is commonplace across all markets and is used by both legitimate and bad actors.

To mitigate harms caused by the increasing quality and prevalence of synthetic media, tech companies and platforms adopt transparency software that can accurately and consistently determine the origin of content. These standards of provenance tracking are uniformly adopted across both major industry players and smaller open-source models and outpace circumvention techniques employed by bad actors. Their implementation is supported and enforced through the establishment of industry alliances and standards.

Synthetic content is clearly labelled and identifiable across platforms, allowing users to benefit from positive use cases whilst making informed decisions about which potentially harmful content is inauthentic. Similarly, authentication and moderation systems can easily and accurately detect synthetic content.

ii. Consumer Attitudes

Users are increasingly exposed to synthetic media of high quality, but the uniform implementation of effective provenance tracking provides them with up-front detail on the origin of content. Consequently, users can make more informed decisions when interacting with synthetic media online. Most users recognise the benefits of synthetic media and feel comfortable engaging with online content because clear and consistent labelling enhances its transparency and perceived safety.

However, many users become over-reliant on labelling when determining whether content is trustworthy. Users begin to assume that all unlabeled content is authentic and reliable, and consequently become more susceptible to traditional scams or mis/disinformation.

iii. Market Outcomes

The uniform adoption of technological mitigation techniques means that platforms and systems can quickly and effectively identify synthetic content, allowing for resilient security systems and improved safety of infrastructure. Authentication systems can prevent synthetic disruption, and the likelihood of mass data leaks or fraudulent activities caused by synthetic interference is reduced.

Since synthetic content is identifiable across platforms, users are less susceptible to synthetic commercial misrepresentation, audiovisual scams, and mis/disinformation, as they are informed up-front when this content has been synthetically produced or altered. Consumers are also protected from synthetic scams that rely on existing infrastructure like telecoms networks, as provenance tracking is adopted across industries and operators can detect synthetic activity. They are, however, left vulnerable to traditional scams and misinformation due to an over-reliance on labelling when making decisions about trustworthy content.

Scenario 3: Stagnation of development

i. Landscape Evolution

Developments in synthetic media slow due to rising costs for tech companies, limited availability of high-quality training data, and increasing public scepticism.

While current positive use cases are still implemented effectively, there is a reduced incentive to develop new beneficial use cases due to slowing tech development and innovation. Bad actors continue to produce harmful synthetic outputs, but stagnation of development means that harms remain broadly similar.

There is continued research and investment in media literacy and novel mitigation techniques, however there is less perceived need for their implementation. There is a reduced focus on cross-industry and international collaboration, and existing regimes are seen as sufficient solutions.

The most concerning harms presented by synthetic media, such as scams, fraud, disinformation, the creation of nonconsensual sexual imagery, are seen to be effectively mitigated by existing regimes. While harms do occur, these are not necessarily considered to be worse when inflicted through the application of synthetic media. Consequently, harmful content is addressed regardless of its origin, and those responsible for consumer protection are trusted to effectively identify, label, or remove harmful content without the implementation of novel mitigation techniques.

ii. Consumer Attitudes

Users are aware of the potential benefits of synthetic media, as well as the risks of encountering false or misleading content. They regularly come across synthetic media, however stagnated tech development means that users can sometimes tell when content has been synthetically produced or manipulated.

Some users are misled or misinformed by harmful content they encounter, but this content is generally identified and removed where necessary.

iii. Market Outcomes

Synthetic outputs have little impact on markets and consumers beyond that which their authentic counterparts do. Harmful synthetic outputs cause disruption but are identified and addressed using the same approaches as are implemented for non-synthetic content.

While each instance of harmful content can be addressed, it is difficult for those responsible for enforcement to identify perpetrators and prevent reoffending due to underdeveloped and underused provenance tracking and a lack of multilateral cooperation. Approaches vary according to territory and platform, and consumers are only effectively protected when they engage with trusted, good-faith market players. Consequently, some users are exposed to harmful synthetic content and mis/disinformation remains impactful.

Scenario 4: Erosion of trust in media

i. Landscape Evolution

Synthetic media is generated quickly, easily, and at a low cost, and users become accustomed to seeing synthetic content across all media platforms. There is significant public awareness of the potential risks and harms brough on by synthetic media use, and less emphasis on its positive applications.

Bad actors continue to find novel ways to misuse synthetic media and harmful content and use cases become increasingly prevalent, outweighing any perceived benefits. Bad actors also develop novel techniques for circumventing mitigation efforts, and current frameworks for consumer and market protection are unfit to address the novel harms infiltrating the market.

Media literacy efforts aim to equip users for this high-risk media environment, but most users are left concerned for their well-being when interacting with content.

ii. Consumer Attitudes

There is increased scepticism among users when engaging with online content and services due to the perceived risks arising from the proliferation of harmful synthetic content and a lack of consistent, effective mitigation.

This scepticism leads to an erosion of trust in media in general, and users become less likely to interact with any potentially synthetic content. Consequently, many users are excluded from potentially beneficial services and content, and markets are negatively impacted.

iii. Market Outcomes

The erosion of trust in media makes consumers less likely to engage with legitimate content and media sources. Users gravitate towards content that reinforces their existing views and beliefs, and accurate news and traditional media is unable to reach a mass audience. This has significant implications for democracy, with users regularly making decisions based on false or misleading content.

Similarly, the proliferation of harmful synthetic media reduces consumer faith in existing infrastructure and markets. Consumers are more likely to take commercial risks as they are unable to differentiate between legitimate businesses and endorsements and synthetic mimicries.

Scenario 5: Lack of sufficient mitigation leads to amplification of harms.

i. Landscape Evolution

Synthetic media is generated quickly, easily, and at a low cost, and becomes highly sophisticated and indistinguishable from genuine content. It is common across platforms and is used for both beneficial and harmful purposes.

There is continued research and investment in technological mitigation techniques, but these are not foolproof due to a lack of universal adoption and advanced circumvention methods employed by bad actors. Similarly, current regulatory efforts are unable to identify and remove harmful synthetic content because of the increasing sophistication of outputs and applications effectively and consistently.

Media literacy is not prioritised, and users are largely unaware of the potential risks or impact of harmful synthetic content as a result. There is a proliferation of harmful synthetic content including scams, fraud, and mis/disinformation, and consumers are not equipped or empowered to identify potentially detrimental outputs.

ii. Consumer Attitudes

Consumers are frequently exposed to synthetic content, but don't have the media literacy skills and awareness to identify or assess potentially harmful synthetic outputs. Many users gravitate towards content that confirms their existing beliefs regardless of authenticity, and there is increasing fragmentation of views and experiences.

As a result of this high-risk environment, users frequently fall victim to sophisticated synthetic scams, convincing mis/disinformation, and commercial misrepresentation. This in turn leads to significant political disruption and market destabilisation.

iii. Market Outcomes

The proliferation of harmful synthetic media causes mass disruption. Bad actors can effectively circumvent authentication systems, leading to increasing instances of data theft and cyberfraud. There is insufficient infrastructure to effectively mitigate these harms, and markets and services are left in disarray.

Bad actors are also able to easily harm individuals through scams or the creation of harmful content. The ease with which they can create highly convincing synthetic content leads to increasing instances of personal attacks such as the creation of defamatory or sexualised deepfakes. There is a lack of effective detection or provenance tracking methods available at the local level, meaning there is limited recourse for victims.

Users are not empowered to identify potentially synthetic content, and increasingly fall victim to highly convincing scams. They are also unable to determine when content might be misinformative, which leads to further harms for society.