![Ofcom logo — making communications work for everyone]

Your response

Please refer to the sub-questions or prompts in the annex to our call for evidence.

Question	Your response
Question 1: Please provide a description introducing your organisation, service or interest in Online Safety.	*Is this response confidential? – N*
	Synthesia is an AI video platform that allows clients to generate video by simply typing text. Videos can be generated at scale, directly in the browser and in minutes. Without the need for cameras, microphones or studios. These videos can be used for training, marketing, sales and more. Our platform lets users create content in over 60 languages within minutes.[1]
	Synthesia is a UK company and was founded in 2017 by a team of AI researchers and entrepreneurs from UCL, Stanford, TUM and Cambridge. Synthesia's web-based platform is now used by more than 10,000 companies worldwide, ranging from SMEs to those in the FTSE100.
	Synthetic media describes video, image, text, or voice that has been fully or partially generated by computers. Synthesia uses proprietary AI, developed in-house, to help generate such synthetic audiovisual content.[2]
	Advances in technology mean that the creation of new synthetic media will increase, and so it is important that a combination of both ethics and education forms a part of the digital and

---

[1] https://www.synthesia.io/about
[2] https://www.synthesia.io/post/the-future-of-synthetic-media

regulatory landscape when we consider these developments. For instance, we believe that education on so-called "deepfakes" is important. Increasingly the term "deepfake" is being used to describe media that is manipulated to make it look like an individual is doing or saying something which they have not said or done. These videos are not always malicious, and are often created using traditional editing tools, rather than AI. Synthesia will shortly be launching a public educational resource on reported "deepfakes", and will be pleased to meet with Ofcom to explain and demonstrate this resource.

Question 2: Can you provide any evidence relating to the presence or quantity of illegal content on user-to-user and search services? **IMPORTANT: Under this question, we are not seeking links to or copies/screenshots of content that is illegal to hold, such as child sexual abuse. Deliberately viewing such images may be a criminal offence and will be reported to the police.**	*Is this response confidential? – N*
Question 3: How do you currently assess the risk of harm to individuals in the UK from illegal content presented by your service?	*Is this response confidential? – N* When it comes to content moderation, Synthesia does not share the Silicon Valley ethos of the internet, and internet platforms, being a "public town square" where users can say anything, and where anything goes. We know that we are a pioneering company which is at the forefront of a new kind of media, and we are aware of the responsibility that this brings. And so we have made sure that ethics, and our approach to content, is an integral part

| | of the company.[3] For this reason, we have developed a content moderation framework.[4]

Synthesia took the decision early on to require all content to go through an internal screening process before being released to our clients. Further, we do not allow the reenactment of an individual without their explicit recorded consent. As we continue to scale we intend to integrate Know Your Customer checks as part of our wider onboarding process.

We have invested heavily into content moderation and we use a mixture of in-house human moderation, along with machinebased tools, to ensure that only appropriate content that meets our Terms of Service reaches our customers. This approach allows us to cover the scale of content, through the rollout of our machine tools, while also providing the room for us to judge context, which will usually require a human moderator. |
|---|---|
| **Question 4: What are your governance, accountability and decision-making structures for user and platform safety?** | *Is this response confidential? – N* |
| **Question 5: What can providers of online services do to enhance the clarity and accessibility of terms of service and public policy statements?** | *Is this response confidential? – N* |
| **Question 6: How do your terms of service or public policy statements treat illegal content? How are these terms of service maintained and how much resource is dedicated to this?** | *Is this response confidential? – N*

We enforce a strict Terms of Service for the use of Synthesia's platform which sets out clearly what we expect of our customers and how we expect Synthesia to be used. These terms go further than the minimum legal standards required by law and offers us the flexibility to ensure that all problematic content is not |

---

[3] https://vriparbelli.medium.com/a-framework-for-ethical-synthetic-media-a109e00ec450
[4] https://www.synthesia.io/post/content-moderation

	approved and therefore not released to clients or the wider world.[5]

	Synthesia's Terms of Service explicitly prohibit:
	  a) Violations of national or international laws or regulations.
	  b) Exploiting or harming minors, or by exposing them to inappropriate content.
	  c) For the purpose of adult entertainment.
	  d) Impersonation of another person or entity.
	  e) To infringe upon the rights of others or act in an obscene, defamatory, immoral threatening, fraudulent, bullying discriminating or harmful manner.
	
	Our Terms of Service are regularly updated. Our most recent version was published in July 2022, but we issued previous versions in March 2022, February 2022 and November 2021.

Question 7: What can providers of online services do to enhance the transparency, accessibility, ease of use and users' awareness of their reporting and complaints mechanisms?	*Is this response confidential? – N*
Question 8: If your service has *reporting or flagging* mechanisms in place for illegal content, or users who post illegal content, how are these processes designed and maintained?	*Is this response confidential? – N*
Question 9: If your service has a *complaints* mechanism in place, how are these processes designed and maintained?	*Is this response confidential? – N*

[5] https://www.synthesia.io/terms/terms-of-service

Question 10: What action does your service take in response to *reports* or *complaints*?	*Is this response confidential? – N*
Question 11: Could improvements be made to content moderation to deliver greater protection for users, without unduly restricting user activity? If so, what?	*Is this response confidential? – N*
Question 12: What automated moderation systems do you have in place around illegal content?	*Is this response confidential? – N* Please refer to our answer to question 3 above. When it comes to content moderation, Synthesia does not share the Silicon Valley ethos of the internet, and internet platforms, being a "public town square" where users can say anything, and where anything goes. We know that we are a pioneering company which is at the forefront of a new kind of media, and we are aware of the responsibility that this brings. And so we have made sure that ethics, and our approach to content, is an integral part of the company. For this reason, we have developed a content moderation framework.[6] Synthesia took the decision early on to require all content to go through an internal screening process before being released to our clients. Further, we do not allow the reenactment of an individual without their explicit recorded consent. As we continue to scale we intend to integrate Know Your Customer checks as part of our wider onboarding process. We have invested heavily into content moderation and we use a mixture of in-house human moderation, along with machine-based tools, to ensure that only appropriate content that meets our Terms of Service reaches our

---

[6] https://www.synthesia.io/post/content-moderation

	customers. This approach allows us to cover the scale of content, through the rollout of our machine tools, while also providing the room for us to judge context, which will usually require a human moderator.[7]
Question 13: How do you use human moderators to identify and assess illegal content?	*Is this response confidential? – N* Please refer to our answer to questions 3 and 12 above.
Question 14: How are sanctions or restrictions around access (including to both the service and to particular content) applied by providers of online services?	*Is this response confidential? – N*
Question 15: In what instances is illegal content removed from your service?	*Is this response confidential? – N* Please refer to our answer to question 6 above. Any content violating our Terms of Service, including illegal content, is removed as soon as it is detected. Additionally, Synthesia has different restrictions which apply to content using either our stock or other Custom avatars. For example, when using our stock avatars we have restricted use conditions to protect the human actors that the stock AI avatars are based on, which limit the content that can be created. Restricted cases focus on expressions of opinion and age-restricted products and services that the human actors might not want to associate with. For example, we do not allow the creation of: content about certain regulated goods and services (like tobacco or gambling), sharing first-person experiences, and content about religion, race, politics, gender and sexuality when using our stock avatars.[8]

---

[7] https://www.synthesia.io/post/content-moderation
[8] https://help.synthesia.io/en/articles/6341934-content-moderation

Question 16: Do you use other tools to reduce the visibility and impact of illegal content?	*Is this response confidential? – N*
Question 17: What other sanctions or disincentives do you employ against users who post illegal content?	*Is this response confidential? – N*
Question 18: Are there any functionalities or design features which evidence suggests can effectively prevent harm, and could or should be deployed more widely by industry?	*Is this response confidential? – N*
Question 19: To what extent does your service encompass functionalities or features designed to mitigate the risk or impact of harm from illegal content?	*Is this response confidential? – N*
Question 20: How do you support the safety and wellbeing of your users as regards illegal content?	*Is this response confidential? – N*
Question 21: How do you mitigate any risks posed by the design of algorithms that support the function of your service (e.g. search engines, or social and content recommender systems), with reference to illegal content specifically?	*Is this response confidential? – N*
Question 22: What age assurance and age verification technologies are available to platforms, and what is the impact and cost of using them?	*Is this response confidential? – N*
Question 23: Can you identify factors which might indicate that a service is likely to attract child users?	*Is this response confidential? – N*

Question 24: Does your service use any age assurance or age verification tools or related technologies to verify or estimate the age of users?	*Is this response confidential? – N*
Question 25: If it is not possible for children to access your service, or a part of it, how do you ensure this?	*Is this response confidential? – N*
Question 26: What information do you have about the age of your users?	*Is this response confidential? – N*
Question 27: For purposes of transparency, what type of information is useful/not useful? Why?	*Is this response confidential? – N*
Question 28: Other than those in this document, are you aware of other measures available for mitigating risk and harm from illegal content?	*Is this response confidential? – Y / N (delete as appropriate)*

Please complete this form in full and return to OS-CFE@ofcom.org.uk