Your response

Questions for industry	Your response
Question 1: Are you providing a UK-established service that is likely to meet the AVMSD definition of a VSP? **Please provide details of the service where relevant. The establishment criteria under the AVMSD are set out in annex 5.**	Confidential? –N No
Question 2: Is your service able to identify users based in specific countries and do you provide customised User Interfaces (UI), User Experience (UX) functionality or interaction based on perceived age and location of users?	Confidential? –N n/a
Question 3: How does your service develop and enforce policies for what is and is not acceptable on your service? (including through Ts&Cs, community standards, and acceptable use policies) **In particular, please provide information explaining:** • **what these policies are and whether they cover the categories of harm listed in the AVMSD (protection of minors, incitement to hatred and violence, and content constituting a criminal offence – specifically Child Sexual Exploitation and Abuse, terrorist material, racism and xenophobia);** • **how your service assesses the risk of harm to its users;** • **how users of the service are made aware of Ts&Cs and acceptable use policies; and** • **how you test user awareness and engagement with Ts&Cs.**	Confidential? –N n/a
Question 4: How are your Ts&Cs (or community standards/ acceptable use policies) implemented? **In particular, please provide information explaining:**	Confidential? –N n/a

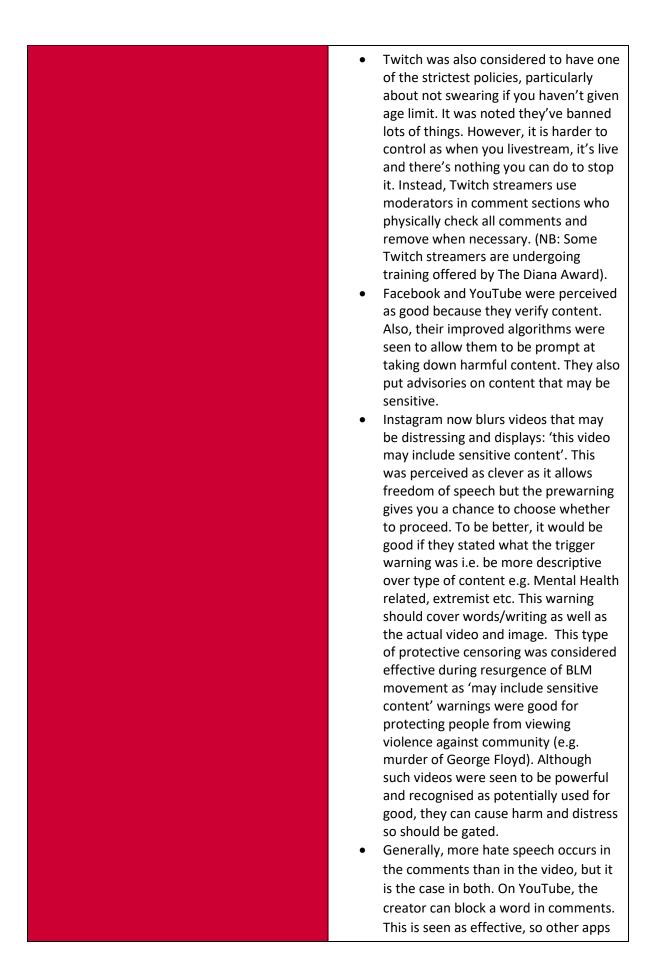|	
• **what systems are in place to identify harmful content or content that may breach your standards and whether these operate on a proactive (e.g. active monitoring of content) or reactive (e.g. in response to reports or flags) basis;** • **the role of human and automated processes and content moderation systems; and** • **how you assess the effectiveness and impact of these mechanisms/ processes.**	
Question 5: Does your service have advertising rules? **In particular, please provide information about any advertising rules your platform has, whether they cover the areas in the AVMS Directive, and how these are enforced. See Annex 5 for a copy of the AVMSD provisions.**	Confidential? –N n/a
Question 6: How far is advertising that appears on your service under your direct control, i.e. marketed, sold or arranged by the platform? **Please provide details of how advertising is marketed, sold and arranged to illustrate your answer.**	Confidential? –N n/a
Question 7: What mechanisms do you have in place to establish whether videos uploaded by users contain advertising, and how are these mechanisms designed, enforced, and assessed for effectiveness?	Confidential? –N n/a
Question 8: Does your service have any reporting or flagging mechanisms in place (human or automated)? **In particular, please provide information explaining:** • **what the mechanisms entail and how they are designed;**	Confidential? –N n/a

- how users are made aware of reporting and flagging mechanisms; - how you test user awareness and engagement with these mechanisms; - how these mechanisms lead to further action, and what are the set of actions taken based on the reported harm; - how services check that any action taken is proportionate and takes into account Article 10 of the European Convention of Human Rights (freedom of expression); - how users (and content creators) are informed as to whether any action has been taken as a result of material they or others have reported or flagged; - whether there is any mechanism for users (including uploaders) to dispute the outcome of any decision regarding content that has been reported or flagged; and - any relevant statistics in relation to internal or external KPIs or targets for response.	
Question 9: Does your service allow users to rate different types of content on your platform? **Please provide details of any rating system and what happens as a result of viewer ratings.**	Confidential? –N n/a
Question 10: Does your service use any age assurance or age verification tools or related technologies to verify the age of users? **In particular, please provide information explaining:** - how your age assurance policies have been developed and what age group(s) they are intended to protect; - how these are implemented and enforced; - how these are assessed for effectiveness or impact; and - if the service is tailored to meet age-appropriate needs (for example, by restricting specific content to specific users), how this works.	Confidential? –N n/a

Question 11: Does your service have any parental control mechanisms in place?
 	
In particular, please provide information explaining:
   • **how these tools have been developed;**
   • **what restrictions they allow;**
   • **how widely they are used; and**
   • **how users of the service, and parents/ guardians if not users themselves, are made aware of and encouraged to use the parental control mechanisms that are available.**
	Confidential? –N
	
	n/a

Question 12: Does your service have a complaints mechanism in place? Please describe this, including how users of your service can access it and what types of complaint they can make.
 	
In particular, please provide information explaining:
   • **any time limits for dealing with complaints;**
   • **how complainants are informed about the outcomes of complaints;**
   • **any appeals processes, how they work, and whether they are independent from the complaints processes; and**
   • **the proportion of complaints which get disputed or appealed.**
	Confidential? –N
	
	n/a

Question 13: What media literacy tools and measures are available on your service?
 	
In particular, please provide any relevant information about:
   • **how you raise awareness of media literacy tools and measures on your service;**
   • **how you assess the effectiveness of any media literacy tools and measures provided on your service; and**
   • **how media literacy considerations, such as your users' ability to understand and respond to the content available to them feature in your thinking about how you design**
	Confidential? –N
	
	n/a

and deliver your services, for example in the user interfaces, flagging content and use of nudges.	
Question 14: Do you publish transparency reports with information about user safety metrics? **Please provide any specific evidence and examples of reports, information around the categorisation and measurements used for internal and external reporting purposes, and whether you have measures in place to report at country/ regional level and track performance over time.**	Confidential? –N n/a
Question 15: What processes and procedures do you have in place to measure the impact and effectiveness of safety tools or protection measures? **If not already captured elsewhere in your response, please provide information relevant to all of the measures listed above explaining:** • **how you test and review user awareness and engagement with each measure (including any analysis or research that you would be willing to share with Ofcom);** • **how often policies and protection measures are reviewed, and what triggers a review; and** • **how you test the impact of policies on users and the business more generally, such as how you balance the costs and benefits of new tools.**	Confidential? –N n/a
Question 16: How do you assess and mitigate the risk of inadvertent removal of legal or non-harmful content? **In particular, please provide any information on:** • **how freedom of expression is taken into account during this assessment;** • **how appeals are handled and what proportion are successful; and**	Confidential? –N n/a

• audits of automated removal systems and, if you have them, any metrics that relate to their effectiveness.	
Question 17: Have you previously implemented any measures which have fallen short of expectations and what was your response to this? **Please provide evidence to support your answer wherever possible.**	Confidential? –N n/a
Question 18: How does your service develop expertise and train staff around different types of harm? (e.g. do you have any partnerships in place?)	Confidential? –N n/a

Questions for all stakeholders	**Your response**
Question 19: What examples are there of effective use and implementation of any of the measures listed in article 28(b)(3) the AVMSD 2018? **The measures are terms and conditions, flagging and reporting mechanisms, age verification systems, rating systems, parental control systems, easy-to-access complaints functions, and the provision of media literacy measures and tools. Please provide evidence and specific examples to support your answer.**	Confidential? –N In a focus group with beneficiaries of The Diana Award, 5 young people aged 13-18 in England gave feedback on the following from their personal experience as users of VSPs: • Most sites now have a report button which is becoming more available to use. The best functions are buttons where you can report content quickly and give the reason. Snapchat was seen as good as can block, report and remove as friend. • YouTube was considered to have good community guidelines, and they are perceived to be better at taking down extremist content since the Christchurch massacre. YouTube's strike system was suggested to be employed by more VSPs.

- Twitch was also considered to have one of the strictest policies, particularly about not swearing if you haven't given age limit. It was noted they've banned lots of things. However, it is harder to control as when you livestream, it's live and there's nothing you can do to stop it. Instead, Twitch streamers use moderators in comment sections who physically check all comments and remove when necessary. (NB: Some Twitch streamers are undergoing training offered by The Diana Award).
- Facebook and YouTube were perceived as good because they verify content. Also, their improved algorithms were seen to allow them to be prompt at taking down harmful content. They also put advisories on content that may be sensitive.
- Instagram now blurs videos that may be distressing and displays: 'this video may include sensitive content'. This was perceived as clever as it allows freedom of speech but the prewarning gives you a chance to choose whether to proceed. To be better, it would be good if they stated what the trigger warning was i.e. be more descriptive over type of content e.g. Mental Health related, extremist etc. This warning should cover words/writing as well as the actual video and image. This type of protective censoring was considered effective during resurgence of BLM movement as 'may include sensitive content' warnings were good for protecting people from viewing violence against community (e.g. murder of George Floyd). Although such videos were seen to be powerful and recognised as potentially used for good, they can cause harm and distress so should be gated.
- Generally, more hate speech occurs in the comments than in the video, but it is the case in both. On YouTube, the creator can block a word in comments. This is seen as effective, so other apps

	should have a censor/screening for horrible comments. Certain words could be banned, or mean posts sent straight to a person to review for harmful/bullying content.
Question 20: What examples are there of measures which have fallen short of expectations regarding users' protection and why? **Please provide evidence to support your answer wherever possible.**	Confidential? –N In a focus group with beneficiaries of The Diana Award, 5 young people aged 13-18 in England gave feedback on the following from their personal experience as users of VSPs: • Clearer community guidelines are needed to encourage better content to be uploaded. For example, TikTok was considered to have really vague guidelines. • It was strongly felt that there should be several steps before a video can be posted ("more hoops to jump through"). There are currently no steps to encourage thinking before uploading. If it's a public account, the user should have to do more to post – e.g. the platform could ask questions about what this content contains. In particular, it's really easy to post on TikTok compared to YouTube and TikTok has many examples of harm (see below). • VSPs should put trigger warnings at the start of things. Trigger warnings could perhaps go at top of posts (where you put location etc – this is a strategy used by some creators) or you should have to slide past the trigger warning before swiping to content. NB: Instagram jumps to the second slide if you've already scrolled past, so could do with 2 TW slides. • VSPs should approve content before its posted. Public accounts in particular need moderation more so than private accounts. • A quick reaction is needed from the companies when guidelines are broken. TikTok could be quicker at taking down content that may be upsetting to young people, particularly as content comes

up so quickly which means no chance to get off the video before seeing.

- On TikTok especially, upsetting content is often disguised by splicing in less harmful content to the video i.e. Innocent images are put before harmful video content to avoid first image being of the harmful content. This lulls users into a false sense of safety and prevents videos from being moderated quickly. Therefore, moderators need to watch the whole video when reviewing, not just the opening.

- On mental health, there needs to be better regulation of content that glorifies mental illness; education on the topic is helpful but glorification can be harmful to people.

- If content might involve children, YouTube will block comments. However, for some content creators, not all videos include children and their comments are amazing, positive spaces, but still all their videos have had comments blocked because of just some videos with children. Needs more nuance.

- TikTok does have a 3-strike policy, but creators find a loophole easily, so the blocking system allows for some users to come back on the platform under a different name. There should be a system where if blocked a certain amount of times then their account should be removed or at least reviewed. Should lead to proactive steps being taken.

| **Question 21: What indicators of potential harm should Ofcom be aware of as part of its ongoing monitoring and compliance activities on VSP services?** **Please provide evidence to support your answer wherever possible.** | Confidential? –N

Outside of the platforms themselves, support hotlines used by young people may give an indicator of where and how harm is occurring. The Diana Award provides a 24/7 crisis messenger service, where young people text their concerns and receive guidance. Other organisations provide different services e.g. Childline. An indicator of patterns across sites could be gained from the number or type of incidents recorded in these sessions and their associated VSP. |

Question 22: The AVMSD 2018 requires VSPs to take appropriate measures to protect minors from content which 'may impair their physical, mental or moral development'. Which types of content do you consider relevant under this? Which measures do you consider most appropriate to protect minors? **Please provide evidence to support your answer wherever possible, including any age-related considerations.**	Confidential? –N In a focus group with beneficiaries of The Diana Award, 5 young people aged 13-18 in England gave feedback on the following from their personal experience as users of VSPs: • Extremist content was considered a harm that children needed protecting from • They have seen a lot of hate speech, i.e. slurs used against people on race, sexuality or gender. These haven't been taken down quickly. Can take the form of comments or be featured within the videos which used slurs against people. • On TikTok, there was a viral video of a suicide, which came from Facebook Live. Lots of things like this occur and can be distressing. Many young people had to delete TikTok due to that video being shared widely. For one of their friends, the video caused panic attacks. Some parents made children delete TikTok as a result. • Also, on TikTok in particular, there is seen to be a trend of glamorising mental illness and specifically eating disorders. Ill people need to be protected from this. TikTok most culpable for glorifying many different mental illnesses. • On bullying, there was another TikTok video of a Mum showing her son a picture of someone with a facial difference and her son's reaction of disgust. This went viral, and prompted young people to fear: what if one person shows someone their picture and they react this way? One of the young people never installed TikTok due to fear of experiencing bullying.
Question 23: What challenges might VSP providers face in the practical and proportionate adoption of measures that Ofcom should be aware of?	Confidential? – Y / N n/a

We would be particularly interested in your reasoning of the factors relevant to the assessment of practicality and proportionality.	
Question 24: How should VSPs balance their users' rights to freedom of expression, and what metrics should they use to monitor this? What role do you see for a regulator?	Confidential? –N In a focus group with beneficiaries of The Diana Award, 5 young people aged 13-18 in England gave feedback on the following from their personal experience as users of VSPs: • TikTok was generally considered to have the weakest safety measures but the young people recognised that it allows for expression of creativity. • Warnings on videos were considered an appropriate way to allow posts to be made while empowering users to have choice over what they see (see above examples e.g. Instagram).
Question 25: How should VSPs provide for an out of court redress mechanism for the impartial settlement of disputes between users and VSP providers? (see paragraph 2.32 and article 28(b)(7) in annex 5). **Please provide evidence or analysis to support your answer wherever possible, including consideration on how this requirement could be met in an effective and proportionate way.**	Confidential? – Y / N n/a
Question 26: How might Ofcom best support VSPs to continue to innovate to keep users safe?	Confidential? –N The Diana Award recommends that Ofcom and VSPs endeavour to include young people's voice in every aspect of the work that they do, to ensure that platforms' community standards are designed with young people's well-being in mind. As their willingness to contribute to these discussions shows, young people should be treated as powerful allies and included at each stage of the conversation. We recommend listening to as a broad a range of young people as possible to ensure that any guidelines formulated are sensitive to the diverse needs and experiences of those young people.

Question 27: How can Ofcom best support businesses to comply with the new requirements?	Confidential? – Y / N n/a
Question 28: Do you have any views on the set of principles set out in paragraph 2.49 (protection and assurance, freedom of expression, adaptability over time, transparency, robust enforcement, independence and proportionality), and balancing the tensions that may sometimes occur between them?	Confidential? – Y / N n/a

Please complete this form in full and return to VSPRegulation@ofcom.org.uk.