Your response

Please refer to the sub-questions or prompts in the Annex of our Call for Evidence.

Ofcom Question	*Ofcom Sub-Questions/Prompts*	OnlyFans Response
Question 1: Please provide a description introducing your organisation, service or interest in Online Safety.	For providers of online services, please provide information about: • The type of service and functionalities you provide;	**Confidential? – ¥ / N** OnlyFans is a unique and inclusive subscription-based platform that gives Creators[1] from all genres the ability to connect directly with their Fans[2] (together, "Users"), monetise their content, and own their intellectual property. Due to its inclusive content policy, OnlyFans is an 18+ platform. OnlyFans is unlike other social media platforms because OnlyFans does not allow third-party ads or seek to monetise its Users' personal data in any way. OnlyFans' mission is to be the safest platform in the world and is wholeheartedly supportive of the primary aim of the Online Safety Bill which is to protect people online.
	• Number of users globally, and in the UK;	As at the end of 2021, OnlyFans had approximately 188 million Fans and 2 million Creators, globally.
	• Global and UK revenues; and	OnlyFans Global revenue for 2021 was US $932 million. OnlyFans UK and Europe revenue for 2021 was US $126,017,000.

---

[1] A "Creator" is a User who has set up his or her OnlyFans account to be able to post video or photo content on OnlyFans, to be viewed by other Creators and/or Fans.

[2] A "Fan" is a User who has set up his or her account in order to follow particular Creators and is able to view those Creators' video or photo content upon subscription, using a payment card.

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	• Your business models and revenue generation. Please indicate where this information is confidential.	OnlyFans business model is different from other social media companies because of the way we built the platform and the User experience. In particular: • OnlyFans is only accessible to Creators and Fans who are over the age of 18 and we have robust procedures in place to keep people who are under 18 off the OnlyFans platform. • To access content on the OnlyFans platform, you must open an account. OnlyFans has a comprehensive onboarding and identity verification process for our Users, unlike other social media platforms which are easily accessible to the public (including minors) and beset by anonymous members. This has the effect of discouraging the posting of illegal content, hate speech, cyberbullying, and trolling. • OnlyFans is not a single-subscription/all-content-access platform. Instead, Fans must subscribe to a Creator to view that Creator's video or photo content. In the UK, a Fan must open an account, including completing an age assurance process and provide a valid payment card before the subscription is activated. • OnlyFans does not have a traditional search function. Unlike other social media platforms that allow a User to search for a topic across an entire platform you can only search for an OnlyFans Creator on the platform by using their specific OnlyFans handle. Fans find Creators on other platforms (e.g., an athlete, TV star, or influencer) and then choose to also follow them on OnlyFans. • Due to OnlyFans subscription model, the vast majority of our content is behind a paywall (save for bios, avatars, and certain text) – i.e., verified account holders must provide payment card details to view any photo or video content, even if the relevant Creator is not charging for that part of

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
		the content. One hundred (100) percent of the sexually explicit images and videos on OnlyFans are behind a paywall and because of this, Users cannot stumble across sexually explicit images and videos like they can on other platforms.
		• OnlyFans Creators elect which monetisation tools they utilise. Within the parameters set out in OnlyFans' Terms of Service (including, Terms of Use for all Users, Terms of Use for Fans, Terms of Use for Creators, Acceptable Use Policy, Referral Program Terms, Complaints Policy, Platform to Business Regulation Terms, and Community Guidelines, together defined for the purposes of this CFE as the "Terms of Service"), Creators set the subscription prices, determine which content is "pay-to-unlock", set the price for custom content and interactions, and maintain control over the content which they create.
		• OnlyFans' subscription-based model means that content published exclusively on OnlyFans does not "go viral," in contrast to other social media platforms. The subscription-based model has the effect of limiting the distribution of content posted.
		• OnlyFans screens all content on its platform including direct messages between Users on the platform. In addition, Fans cannot connect with Creators without first having subscribed to, and engaged with, their content. This means that if someone does behave unlawfully, inappropriately, or in any way which violates our Terms of Service (including the Acceptable Use Policy) in direct messages, we determine who is behind the messages and take appropriate action against that individual. This also radically reduces cyberbullying and trolling.
		• OnlyFans does not monetise our Fans' or Creators' data for its own benefit, in direct contrast to the advertising-based revenue generation business models of other, prominent, social media platforms. We do not

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
		sell any Fan or Creator data to any third parties, including advertisers. Our Fans and Creators have autonomy over the composition of their newsfeeds and the content they choose to view. Our method for generating revenue is transparent and clear to all Fans and Creators. We also regularly discuss it in public forums. • OnlyFans utilises a comprehensive blend of state-of-the-art technologies and extensive human, manual review of content in furtherance of our mission to be the safest platform in the world. In addition to technological review, all content on the platform is screened within 24 hours of being posted by our trained team of content reviewers, who make up more than eighty (80) percent of our workforce. Additionally, all Creators are reverified on a rolling cycle every thirty (30) days. This process is described in detail at Question 6 below.
Question 2: Can you provide any evidence relating to the presence or quantity of illegal content on user-to-user and search services?	We are particularly interested in evidence about how this might vary across different services or types of service, or across services with particular users, features or functionalities.	**Confidential? – ¥ / N** OnlyFans has a zero-tolerance policy for illegal content and any breach of the Terms of Service related to illegal content results in a swift, automatic, and permanent ban from the platform. OnlyFans continuously scans its platform to prevent the posting of illegal content. OnlyFans' Mission, Vision and Values place respect, safety, and integrity at their core. The commitment to ensuring that the platform is a safe and inclusive space is integral to the operation of the OnlyFans' business. We publish information concerning our commitment to protect our community from illegal content and criminal activity on our website. We are particularly focused on ensuring that our platform is never used to enable human trafficking, hate speech, and the publication or sharing of CSAM, but we are vigilant against all forms of illegal content and activity.

Ofcom Question	*Ofcom Sub-Questions/Prompts*	OnlyFans Response
		OnlyFans' business model is materially different from many other social media platforms and these differences explain the relatively smaller amount of illegal content on the platform. OnlyFans has in place robust identity verification processes for Users, unlike other platforms which are open to the public and can be used anonymously. Every Creator must provide true and accurate information about their identity before they can use OnlyFans, and Fans must provide a valid email address and a valid payment card. These measures ensure that OnlyFans has information about who is using its platform and how they are using it. As Fans and Creators are not anonymous, they know that if they do something illegal (or otherwise violate the Terms of Service), OnlyFans will be able to easily identify them and report them to law enforcement authorities, if appropriate.
		OnlyFans also monitors personal interactions between Creators and Fans, to prevent illegal material being disseminated that way.
		Whilst we track the numbers of posts that we remove from our platform because they violate our Terms of Service (including our Acceptable Use Policy), we do not track the nature of the violation (other than CSAM related violations). In any given month, as laid out in our monthly transparency reports, the number of posts that we identify as containing material that violates our Terms of Service (including our Acceptable Use Policy) is very low. For example, out of 20,423,107 posts uploaded to the platform in July 2022, we identified and proactively removed from the platform 36,452 posts that violated our Acceptable Use Policy. This constitutes fewer than 0.2% of all posts uploaded that month.
		CSAM
		Any incident of CSAM on our platform is taken incredibly seriously and is reported to the National Center for Missing & Exploited Children ("NCMEC"). OnlyFans aggressively targets and reports Users who create or try to use its

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
		platform to distribute illegal content and in particular reports all suspected incidents of CSAM to NCMEC. Within that population of posts identified as violating our Terms of Service (including our Acceptable Use Policy), the instances of CSAM are very small. Incidents of suspected CSAM make up less than 0.001% of all content submitted by Creators to be posted on OnlyFans. For example, in 2021, OnlyFans submitted 2,984 CyberTipline reports to NCMEC, out of a total of 151,911,634 total posts during the same period. OnlyFans has a page within its Safety and Transparency Center dedicated to its efforts to combat CSAM (see https://onlyfans.com/fightcsam). **Modern Slavery and Human Trafficking** OnlyFans is committed to taking all reasonable and appropriate steps to ensure that there are no modern slavery or human trafficking activities taking place on our platform. OnlyFans' Terms of Service prohibit Users from uploading, posting, displaying, or publishing content on the OnlyFans platform that: • shows, includes, or refers to any other individual unless written consent is obtained from each individual to use their name or images (or both) in the content; or • shows, promotes, advertises, or refers to escort services, sex trafficking, or prostitution. OnlyFans has taken the following steps to prevent, detect, and report modern slavery and human trafficking risks: • continued to roll out company-wide mandatory modern slavery and trafficking training;

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
		• issued company-wide policies and procedures as to acceptable business practices and ethical conduct; • reviewed our supply chain to identify and mitigate risks of modern slavery and human trafficking; • invested heavily in enhancing our approach to content moderation including hiring additional specialist resources; • appointed an independent Monitor to review our safety controls and compliance program to continue to enhance the safety of our platform (described in more detail below in the response to Question 3); and • increased our engagement with third parties that seek to combat modern slavery and human trafficking. OnlyFans has a page within its Safety and Transparency Center dedicated to its efforts to combat modern slavery and human trafficking (see https://onlyfans.com/transparency-center/preventing). **<u>Hate Speech</u>** OnlyFans' Terms of Service prohibit Users from uploading, posting, displaying, or publishing content that contains, promotes, advertises, or refers to hate speech. OnlyFans views hate speech to be any content that is intended to vilify, humiliate, dehumanise, exclude, attack, threaten, or incite hatred, fear of, or violence against, a group or individual based on race, ethnicity, national origin, immigration status, caste, religion, sex, gender identity or expression, sexual orientation, age, disability, serious disease, veteran status, or any other protected characteristic. OnlyFans has a page within its Safety and Transparency Center dedicated to its efforts to tackle hate speech (see https://onlyfans.com/transparency-center/tackling).

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
Question 3: How do you currently assess the risk of harm to individuals in the UK from illegal content presented by your service?	Please provide information about: • How risks from illegal content are identified (including any relevant internal processes, policies and documents);	**Confidential? – ¥ / N** OnlyFans regularly assesses its risk profile in order to adapt efficiently to evolving issues and develop protections for Users as the platform grows. OnlyFans adopts a multi-layered approach to its risk assessment process. First, it regularly assesses risk internally through trend analysis (utilising, e.g., the monthly metrics circulated to decision makers). OnlyFans also identifies the greatest risk areas for focus following an internal risk analysis of de-activated posts. Secondly, OnlyFans engaged an independent, external third-party team (the "Monitor" team), which is assessing risk associated with illegal content. More specifically, the Monitor team – a combination of lawyers and cybersecurity forensic analysts – is conducting a comprehensive 360-degree review of the Company's operations and internal processes and procedures to identify key risk areas and assess the effectiveness of the platform's efforts to mitigate those risks. The Monitor team regularly offers recommendations for control improvements to ensure OnlyFans' controls are appropriately designed to mitigate key risks. Ofcom can find further details regarding this third-party assurance in our Safety & Transparency Center – see https://onlyfans.com/transparency-center/commitment. Internal manuals have been developed to explain and codify the policies and procedures for content moderation. Clear sanctions are imposed in relation to any illegal content posted on the platform.
	• The extent to which you factor in evidence on users' behaviour and age to consider potential risks; and	When assessing the potential risk associated with allowing the sharing of lawful content on OnlyFans, the fact that OnlyFans is an 18+ platform is an important consideration. Accordingly, age is a significant factor considered in determining the potential risks when it comes to lawful content.

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	• The exacerbating and mitigating risk factors you consider (for example, relating to your users, business model or features and functionalities which may have an impact on the risk of harm).	OnlyFans is an 18+ platform and as such we have adopted an inclusive content policy which recognises and allows adult content to be shared on the platform subject to our Terms of Service. By virtue of our Terms of Service, Community Guidelines and Safety and Transparency Center, OnlyFans' Users have a clear and transparent understanding of the measures OnlyFans has in place to prevent illegal content and content deemed to be harmful from appearing on the platform. OnlyFans actively discourages bad actors from trying to abuse the platform in those ways. OnlyFans empowers Creators from all genres, including artists, musicians, glamour models, adult content creators, fitness professionals, athletes, and more. We do not classify or categorise Creators – we apply the same high levels of protections to all Creators. OnlyFans employs a series of successive and overlapping controls that create multiple opportunities to detect and prevent potential violations, including: • prohibiting anyone under the age of 18 from being a Fan or Creator – OnlyFans is an 18+ platform; • maintaining a rigorous onboarding process that eliminates the anonymity of Users, with a blend of state-of-the-art technologies and extensive human review; • maintaining a robust 24-hour Content Moderation team, so that all content undergoes both human and automated review; • maintaining a subscription-based model, where explicit content can only be accessed by verified 18+ Users who have subscribed to the content and provided a valid payment card; • ensuring private messages between Creators and Fans are also screened for illegal content and Terms of Service violations;

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
		• implementing a very limited platform wide search function limiting the discoverability of content and thereby obliging Fans to select the specific Creators they wish to follow; and • avoiding the use of algorithms to determine what content is shown to Users. As mentioned above at Question 3, the Monitor team conducts a comprehensive 360-degree review of OnlyFans' operations and internal processes and procedures to identify key risk areas and assess the effectiveness of the platform's efforts to mitigate those risks. The Monitor regularly offers recommendations for control improvements to ensure OnlyFans' controls are appropriately designed to mitigate key risks.
Question 4: What are your governance, accountability and decision-making structures for user and platform safety?	As part of your answer, please outline how different teams may consider user safety risks across different business functions such as product development, management engineering, public policy, safety, legal, business development and marketing. Please consider: • How are risks to user safety escalated and acted upon in your organisation? What senior management oversight is in place?	**Confidential? – Y / N**

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	• How are staff trained to understand how their own roles and responsibilities can potentially create risks to user safety?	

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	• How do you ensure consistency in consideration of user safety across teams?	
Question 5: What can providers of online services do to enhance the clarity and accessibility of terms of service and public policy statements?	Please submit evidence about what features make terms or policies clear and accessible.	**Confidential? – ¥ / N** OnlyFans' Terms of Service are accessible to the public without requiring an account or payment verification on OnlyFans' public facing website, prior to any login being required. See https://onlyfans.com/terms. The Terms of Service have been written and structured to be clear and accessible to all Users. Unlike a number of other platforms, we have attempted to keep our Terms of Service together in one place, not link to multiple pages, and ensure that the use of jargon and technical terms is kept to a minimum. We have expressly included explanations of any legal terms that we considered that Users may not readily understand. All Users must accept OnlyFans' Terms of Service (encompassing the Acceptable Use Policy) before they are able to establish an account and use the platform. OnlyFans has a Safety and Transparency Center (see https://onlyfans.com/transparency-center), where we clearly communicate the policies and procedures in a digestible, question-and-answer format.

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
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?	Please outline as part of your answer: • What you currently cover in your terms of service and public policy documents in relation to illegal content and why, including any reference to what is considered illegal content, any proactive measures to identify it, and sanctions applied to users who are in breach;	**Confidential? – ¥ / N** OnlyFans has zero tolerance for any illegal or unacceptable content. OnlyFans' public-facing Acceptable Use Policy (contained within the Terms of Service) prohibits content which is "*illegal, fraudulent, defamatory, hateful, discriminatory, threatening or harassing, or which encourages or promotes violence or any illegal activity*." Within two (2) minutes of an attempted upload onto the OnlyFans platform, all content is reviewed by state-of-the-art technologies to scan for the presence of explicit content, minors, weapons, as well as non-compliant terminology. This will prevent certain content from being posted on the OnlyFans platform. All content which passes this initial review is then also reviewed by a human content moderator within 24 hours of being posted onto the platform. This includes all content shared in direct messages. A significant proportion of OnlyFans' team is employed in roles related to User safety and enforcement of our Terms of Service. Users uploading illegal content will face sanctions including a swift, automatic, and permanent ban from the platform. Steps are taken to ensure that any banned User is prevented from opening a new OnlyFans account in the future, including IP and Device ID tracking.
	• Whether you consider, when drafting these documents, the accessibility needs of different user groups and affected persons, including children, parents or groups with certain characteristics which may put	OnlyFans is an 18+ platform. Accordingly, there is no need to consider accessibility needs of children, and for different User groups the need to consider accessibility is limited. We apply the same high level of controls for all Fans and Creators on the OnlyFans platform. Furthermore, all content, regardless of its nature, is subject to proactive review by our content moderation team.

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	them at a higher risk of harm from content on the service;	
	• Evidence of the process, time and any costs involved in developing these terms; and	Due to its subscription model, from its inception OnlyFans has always been an 18+ platform.
	• Whether you have any evidence about how users engage with your terms of service, or whether they understand what they mean in practice.	The Terms of Service are designed to be digestible and clear for Creators and Fans to understand. Creators and Fans can flag or report any content that violates OnlyFans' Terms of Service using a simple User reporting checklist. (Please see below at Question 8). OnlyFans' monthly transparency reports set out the number of posts removed (per month) for breaches of the Acceptable Use Policy (within the Terms of Service), along with the total number of posts on OnlyFans (per month). As an example, in June 2022, 0.16% of all posts were identified as violating the Acceptable Use Policy (and were removed), which indicates that the vast majority of Users comply with OnlyFans Terms of Service and Acceptable Use Policy. When OnlyFans Terms of Service are updated, Users are provided with advance notice, an opportunity to review the updated Terms of Service and where the updates are significant a summary of those updates, in addition to the full text, is also provided.
Question 7: What can providers of online services do to enhance the transparency, accessibility, ease of use and users' awareness of their	Please submit evidence about what features make user reporting and complaints systems effective, considering: • Reporting or complaints routes for registered users;	**Confidential? – ¥ / N** OnlyFans recognises the importance of ensuring the ability to report content for both Users and non-Users. As a result, it has set up clear and accessible reporting measures for both. Please see details below at Question 8.

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
reporting and complaints mechanisms?	• Reporting or complaints routes for non-registered users; and	Please see details below at Question 8.
	• Reporting routes for children and adults.	OnlyFans is an 18+ platform. Notwithstanding, reporting routes are the same for all Users and non-Users.
Question 8: If your service has <u>reporting or flagging</u> mechanisms in place for illegal content, or users who post illegal content, how are these processes designed and maintained?	We are interested in obtaining evidence on: • How users report content on your service (including the mechanisms' location and prominence for users, and any screengrabs you could provide) and whether this is separate from your complaints procedure;	**Confidential? – Y / N̶**

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	• Whether users need to create accounts to access reporting and flagging;	
	• What type of content or conduct users and non-users may flag or report, including any specific lists or categories;	

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	• The key choices and factors involved in designing these mechanisms;	
	• How you ensure that reporting and flagging is user-friendly and accessible;	
	• Whether any particular consideration is given to the needs of children, including children of different ages;	
	• Whether any particular consideration is given to the needs of different user groups, including users with certain characteristics which may put them at a higher risk of harm from content on the service (if	

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	so, which user groups you have considered);	
	• The cost involved in designing and maintaining these mechanisms;	
	• Whether your reporting and flagging mechanisms are effective, in terms of identifying illegal content, and how you determine this; and	
	• Whether you use trusted flaggers (and if their reports are handled differently).	
Question 9: If your service has a <u>complaints</u> mechanism in place, how are these processes designed and maintained?	We are interested in obtaining evidence on: • How users can complain about content on your service (including any screengrabs you could provide) and whether there is any separate procedure from your reporting mechanisms above;	**Confidential? – Y / N~~**

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	• The type of complaint that can be made, including any categories or lists (including if the same categories are used for complaints);	
	• Whether complaints can be made in cases when content is de-prioritised or taken down proactively by technology, or about freedom of expression and privacy issues;	
	• Who can make a complaint;	
	• Whether there are different procedures for different categories of users, including children or content creators vs. affected persons;	
	• The key choices and factors involved in designing your complaints procedure;	
	• The cost involved in designing and maintaining this process;	

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	• Whether any record of complaints is kept, and if so what; and	
	• Any appeals processes if complainants are unsatisfied with the outcome of the complaint, when these can be accessed, how they work, and whether they are independent from the complaints processes	
Question 10: What action does your service take in response to <u>reports</u> or <u>complaints</u>?	We are interested in obtaining evidence on: • Any variation in the number and actionability (i.e. the proportion that result in a takedown or other action in response) of reports or complaints across type of harm;	—
	• What proportion of reports are considered, and what proportion are acted upon;	

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	• Whether any reports or complaints are expedited or directed to specialist teams, the criteria for this, and the cost involved in facilitating this;	
	• If users identify illegal content, whether law enforcement or other third parties are informed;	
	• How the validity of reports is triaged and assessed, including how malicious reporting is identified, handled and/or mitigated;	

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	• How users and content creators are informed as to whether any and what action has been taken, and why, as a result of material they or others have reported or flagged;	

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response	

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	• What happens to the content while it is being assessed/processed; and	
	• Any internal or external timeframes or key performance indicators (KPIs) for acting on or determining reports or complaints.	
Question 11: Could improvements be made to content moderation to deliver greater protection for users, without unduly restricting user activity? If so, what?	Please provide relevant evidence explaining your response to this question. Please consider improvements in terms of user safety and user rights, as well as any relevant considerations around potential costs or cost drivers.	**Confidential? – Y / N**

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
Question 12: What automated moderation systems do you have in place around illegal content?	In particular, please consider: • Do you use hash-matching to identify any forms of known illegal content?	**Confidential? – ¥ / N** Before content is posted and available to Fans, OnlyFans conducts hashing of all images and hashed content that appears in cross-checked databases (internal and external) is blocked, prior to being posted on OnlyFans. In addition, OnlyFans works with third parties to identify and share IP addresses and other traceable data associated with CSAM video or photo content. When OnlyFans identifies previously unreported CSAM, it calculates a hash of the video or photo content, adds that information to its internal repository, and reports it to NCMEC and other third parties that have similar tools. OnlyFans' personnel regularly assess the quality of the hash technology and hash lists that we use.
	• Do you use any other automated techniques to identify any forms of illegal content?	OnlyFans also uses advanced image analysis software to automatically review all media content posted within the OnlyFans platform to identify images of persons who may be under 18 years old, images which contain extreme or violent imagery, or other images which may violate the Acceptable Use Policy (contained within the Terms of Service). For example, we are able to detect if there is more than one person in the content and confirm that person has provided consent to appear in content posted on OnlyFans. OnlyFans also transcribes and reviews all text and audio content against a list of words and phrases, which is maintained and updated by its content moderation team which may indicate that content is in breach of the Acceptable Use Policy.

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	• If you use in-house automated techniques, how are these developed and trained? What training data was used and how it was sourced?	OnlyFans' overall approach to content moderation has been developed in-house, and we continually seek to improve our approach. The tools we use are mostly developed by third parties, consistent with our desire to use best-in-class solutions that work for the particular set-up of the OnlyFans' platform.
	• Do you use any third-party datasets or providers and, if so, which?	█ Text scanning for non-compliant terminology. █: Scans media for content (i.e., what is present in the media) and demographic information (i.e., if a person or more than one person is present and detailed analysis of age, etc.). █ also performs OCR for images to text on a daily basis. **Ondato**: During the identity verification process, OnlyFans uses Ondato to analyse the identification and live image capture (i.e., selfie). These checks are followed by checks for sanctions, politically exposed persons, and adverse media screenings. All of this provides OnlyFans with information to evaluate any suspicions quickly. Ondato's technology can also detect spoofing attempts and confirms matches between identification documents and Users' faces. **Jumio**: A secondary system that backs up Ondato in case Ondato is technically unavailable. Jumio follows a similar process to Ondato, described above. **NCMEC Hash Database**: NCMEC's database of known bad hashes. █: Previews images and screenshots, and detects whether they contain explicit content, minors, and weapons. **Yoti** (certain jurisdictions, including the U.K): A combination of technologies, liveness anti-spoofing, and document authenticity checks to verify the age of

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
		potential Users. Yoti uses facial age estimation; if the facial age estimation fails, the User is submitted for an Ondato check, described above (the process referred to herein as "Age Assurance"). ███████ : Builds products, leads programmes, maintains existing resources, and develops awareness campaigns to tackle child sexual exploitation. OnlyFans utilises ████████████████' Hash Matching, CSAM Classifier, Reporting and Review Tools. ████████████████ : Technology that aids the finding and removal of images of child exploitation. **Others in process**: OnlyFans is also currently in the process of adding an ███ hash database and also ███ hash database.
	• If you use a third-party provider, did you consider there to be choice in the market? Why did you opt for your chosen provider?	OnlyFans beta tests numerous tools from a variety of providers to utilise industry-leading third-party technologies to maintain its world class content moderation process.
	• What happens to content once it is identified by automated means?	Any content which is flagged by the state-of-the-art technology tools used by OnlyFans is then escalated to the 24-hour content moderation team for in depth human moderation. The content moderation team is made up of trained manual reviewers who review content to ensure compliance with anti-fraud measures, identity verification, compliance with copyright law, and OnlyFans' Terms of Service.
	• Can you provide any evidence as to how effective these techniques are, in terms of	OnlyFans has a highly effective onboarding and content moderation processes to keep community members and Users safe. It also regularly

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	reducing harm to users, and how this may vary by harm?	tests existing tools and assesses new tools from other third-party service providers to ensure the best measures are deployed. As set out in the June 2022 Transparency report, over sixty (60) percent of applications for Creator accounts were rejected because they failed to meet the requirements of the account verification and vetting process, demonstrating the rigour with which OnlyFans approaches its onboarding process and its fidelity to policies and procedures that were carefully developed to protect community members.
	• How do you assess the performance (including wrongful takedown rates, precision and recall) of automated moderation techniques?	As each possible violation is reviewed on a case-by-case basis, there is a clear record of the decision-making process in relation to content takedown. Monthly reports from the automated technologies are reviewed by the team leads, to ensure any issues concerning false positives are identified and addressed.
	• What safeguards are employed to protect user privacy and/or to avoid over takedown?	Please see above at Question 14.
	• Are there certain types of automated techniques which have strengths over others, or which are better suited to certain harms or types of content over others?	Please see above at Question 12 for distinction between text and image scanning software.
	• What are the barriers and costs involved in deploying these techniques?	The cost is difficult to quantify based on the internal and external personnel involved, as well as the hours that go into each layer of these multi layered

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
		processes. As there is no all-encompassing package of automated tools, the cost is higher for a layered process of tools and human review. OnlyFans diligences each third-party to ensure, first and foremost, the vendor has the tools necessary to meet OnlyFans' Terms of Service and to protect Users, and secondly, if the relationship is financially viable. Certain vendors are tested in particular ways to determine if they can fulfil OnlyFans' needs, especially as the platform grows and develops. If a third-party has a technical issue, OnlyFans assesses whether additional quality control checks will be sufficient, or a new vendor is needed.
Question 13: How do you use human moderators to identify and assess illegal content?	In particular, please consider: • How do you determine the level of human moderation required by your platform, including by type of content?	**Confidential? – Y / N**
	• Are moderators employed by the service, outsourced, or made up of volunteers?	

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	• Are moderators vetted, and how?	
	• What coverage do moderators provide (e.g. on weekends or overnight, UK time)?	
	• What training and support is provided to moderators? For instance, are certain moderators specialised in certain harms or speech issues? Is training updated, and when? Are moderators trained to identify illegal content?	
	• If you make use of automated moderation, how do human moderators and automated systems work together, and what is their relative scale? How do you guard against automation bias?	
	• What are the costs involved in these practices? In the absence of specific costs, please provide indication of cost drivers (e.g.	

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	moderator location) and other relevant figures (e.g. number of moderators employed, how many items the platform moderates per day).	
	• How do you assess the accuracy and consistency of human moderation teams?	
Question 14: How are sanctions or restrictions around access (including to both the service and to particular content) applied by providers of online services?	Please provide evidence around the application and accuracy of sanctions/restrictions, and safeguards you consider should be in place to protect users' privacy and prevent unwarranted sanction.	**Confidential? – Y / N**

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
Question 15: In what instances is illegal content removed from your service?	Please outline the circumstances in which illegal content is removed and how this may differ by type of illegal content.	**Confidential? – Y / ~~N~~**
Question 16: Do you use other tools to reduce the visibility and impact of illegal content?	Please consider providing the following information: • The circumstances in which content is actioned and how;	**Confidential? – Y / ~~N~~**
	• Whether these tools are specified in public user terms or policies;	
	• Evidence of whether your users understand these tools' existence and application;	
	• What safeguards you use to ensure they are applied consistently and fairly;	
	• Evidence on the costs of developing and maintaining these tools; and	

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	• Evidence on the effectiveness of these tools, in terms of reducing harm.	
Question 17: What other sanctions or disincentives do you employ against users who post illegal content?	Please outline: • What type of sanctions do you apply and in what instances?	**Confidential? – Y / N**
	• Do you take steps to prevent users who have been banned from a service re-accessing it and, if so, what? What are the costs involved in this?	
	• Are users informed about sanctions and, if so, how?	
	• What evidence can you provide around their efficacy, in terms of reducing harm?	
	• What safeguards do you use to ensure users are sanctioned consistently and fairly?	
Question 18: Are there any functionalities or design features which	Please provide relevant evidence explaining your response to this question.	**Confidential? – Y / N**

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
evidence suggests can effectively prevent harm, and could or should be deployed more widely by industry?		
Question 19: To what extent does your service encompass functionalities or features designed to mitigate the risk or	In particular: • Can you identify any functionalities or features on your platform which you consider can effectively prevent harm?	**Confidential? – Y / N̶**

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
impact of harm from illegal content?		
	• What control do users have over what they are shown or is delivered to them?	
	• To what extent are the features or functionalities you identified reliant on other technology – for instance, age verification or ID verification mechanisms?	

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	• What costs or cost drivers are involved in developing these features or functionalities?	
	• Is safety incorporated into the product design and development process? To what extent do you consider evidence about user behaviour when developing features or functionalities intended to enhance user safety?	
	• How do you measure and what evidence can you provide around the impact and effectiveness of these techniques, in terms of reducing harm to users?	
	• How do you assess the impact of these techniques on users' privacy and minimise the risk of over restriction?	
Question 20: How do you support the safety and wellbeing of your	In particular:	**Confidential? – Y / N**

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
users as regards illegal content?	• Do you provide support through your platform (e.g. signposting to resources)?	
	• Do you provide support off-platform (e.g. funding or facilitating programmes)?	
	• How effective are these types of interventions, in terms of minimising harm from and impact of illegal content on	

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	users? What evidence do you have to show this?	
	• What are the costs involved in implementing the support measures you have described?	
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?	In particular: • If different from your risk assessment process outlined in response to Q3, how do you assess the risk from algorithms central to the function of your service?	**Confidential? – Y / N**
	• What safeguards do you have in place to mitigate the risks posed by algorithms (e.g. testing them before they are put into use, and monitoring their performance in real world settings)?	
	• What are the costs involved in implementing these safeguards? In the absence of specific costs, please provide indication of the key cost drivers.	**Confidential? – Y / N**

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	• How do you measure the effectiveness of these safeguards, in terms of reducing harm to users?	
	• What information can you provide to demonstrate the effectiveness of such safeguards?	
	• How do you assess the impact of these safeguards on users' privacy and minimise the risk of over restriction?	
Question 22: What age assurance and age verification technologies are available to platforms, and what is the impact and cost of using them?	In particular, please provide evidence explaining: • How these technologies can be assessed for effectiveness or impact on users' safety;	**Confidential? – Y / N̶**
	• How accurate these tools are in verifying the age of users, and effective in preventing children from accessing harmful content;	

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	• Steps that can be taken to mitigate any risk of bias or exclusion that may result from age assurance and age verification tools;	
	• The costs involved in implementing such technologies; and	
	• The safeguards necessary to ensure users' privacy and access to information is protected, and over restriction is avoided.	
Question 23: Can you identify factors which might indicate that a service is likely to attract child users?		**Confidential? – Y / 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?	In particular, please provide information explaining: • The nature of the age assurance or age verification tools you have in place, including whether you engage third party providers;	**Confidential? – Y / N**

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	• The costs involved in implementing such technologies;	
	• How accurate these tools are in verifying the age of users, and effective in preventing children from accessing harmful content;	

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	• How your age assurance policies have been developed and what age group(s) they are intended to protect;	
	• What kind of content/activity your age assurance policies and mechanisms are intended to protect children from;	
	• If the service is tailored to meet age-appropriate needs (for example, by restricting specific content to specific users), how this works;	
	• The steps you take to mitigate any risk of bias or exclusion that may result from the age assurance and age verification tools you have in place; and	
	• Any safeguards in place to ensure users' privacy and access to information is protected, and over restriction is avoided.	
Question 25: If it is not possible for children to access your service, or a part of it, how do you ensure this?		Confidential? – Y / N

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
Question 26: What information do you have about the age of your users?	Do you gather, or share with any third parties (and if so which), any information that can assist in estimating a user's age, either at the point a user first accesses the service or subsequently? If so, how do you do this and why?	**Confidential? – Y / N**
Question 27: For purposes of transparency, what type of information is useful/not useful? Why?	In particular, please consider: • Any evidence of public information positively or negatively affecting online user safety or behaviours, how this information is used, and by whom;	**Confidential? – Y / N**
	• What information platforms should make available, considering frequency, format and intended audiences;	

Ofcom Question	Ofcom Sub-Questions/Prompts	OnlyFans Response
	• What information Ofcom should make available through its transparency report, considering frequency, format, intended audiences and potential use cases by external stakeholders;	
	• The benefits and/or drawbacks of standardised information and metrics; and	
	• Any negative impacts or potential unintended consequences of publishing certain types of information, and how these may be mitigated.	
Question 28: Other than those in this document, are you aware of other measures available for mitigating risk and harm from illegal content?	We would be interested in any evidence you can provide on their efficacy, in terms of reducing harm to users, cost and impact on user rights and user experience.	**Confidential? – Y / N**

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