

---

# Ofcom's first year of video-sharing platform regulation

What we found

---

[Welsh translation available](#)

# Contents

---

## Section

1. Executive summary	3
----------------------	---

## Part A: Year 1 report

2. An introduction to VSP regulation	7
3. Our approach to VSP regulation in Year 1	10
4. Key findings in Year 1	12
5. Our strategic priorities in Year 2	21

## Part B: Platform reports

6. TikTok	29
7. Snapchat	46
8. Twitch	56
9. Vimeo	68
10. BitChute	76
11. Smaller VSPs	84
12. An introduction to adult VSPs	89
13. OnlyFans	93
14. Smaller adult VSPs	105

# 1. Executive summary

## Our first VSP report

This is Ofcom's first report on video-sharing platforms (VSPs) since being appointed as the statutory regulator for VSPs established in the UK. This report sets out our key findings from the first year of regulation, running from October 2021 to October 2022 (Year 1). Our findings cover VSPs that notified Ofcom they were in scope of the regime. We also set out our strategy and areas of focus for the next year of the regime (Year 2).

## We moved quickly to fully implement regulation of VSPs

Ofcom is one of the first regulators in Europe to implement and deliver a fully operational VSP Regime. Our regime was launched when we published our VSP Harms and Measures Guidance ([VSP Guidance](#)) and [Plan and Approach](#) in October 2021. At that time, we set out our four broad aims: **to raise standards, address non-compliance, increase transparency, and get industry and ourselves ready for the future Online Safety regime.**

We also set out five Year 1 strategic priorities. These focused on specific harms and user safety measures, with the overall goal of ensuring that VSPs put in place systems and processes to protect their users from harmful video material.

## Platforms have taken steps to comply with the new regime

We used our statutory powers to issue enforceable information requests to all notified VSPs. Our key Year 1 findings are that:

- **All platforms have safety measures in place, including rules on what kinds of video material are allowed.** Some platforms made changes to their measures in direct response to being regulated under the VSP Regime.
- **Platforms generally provided limited evidence on how well their safety measures are operating to protect users.** This creates difficulty in determining with any certainty whether VSPs' safety measures are working consistently and effectively.
- **More robust measures are needed to prevent children accessing pornography.** Some adult VSPs' access control measures are not sufficiently robust in stopping children accessing pornography. We expect them to make clear plans to implement more robust solutions and make these changes this year when possible.
- **Some platforms could be better equipped for regulation.** Some platforms are not sufficiently prepared and resourced for regulation. Going forward, we will be looking for platforms to improve and provide more comprehensive responses to Ofcom's information requests.
- **Platforms are not prioritising risk assessment processes,** which Ofcom believes are fundamental to proactively identifying and mitigating risks to user safety. Risk assessments will be a requirement on all regulated services under the Online Safety regime.

## We have driven positive change during our first year

In Year 1 we built a comprehensive baseline knowledge of the VSP sector through our programme of research, supervisory engagement, and information gathering. Highlights from the year include:

- **Notification:** We published our guidance on whether platforms need to notify and at the time of writing 19 UK-established VSPs had notified us.
- **Supervision:** Taking a continuous and open approach to engaging with platforms has allowed us to develop more effective regulatory relationships with the VSP providers in scope.
- **Information gathering:** We issued our first round of formal information requests to VSPs. Our ability to issue these requests and publish the information in regular transparency reports is one of our key levers for driving change across the industry.
- **Securing compliance:** We have engaged constructively with notified VSPs to assess compliance, including following the tragic Buffalo shooting to understand what happened and how their measures were deployed. We have also engaged extensively with notified VSPs to incentivise compliance with our information requests and opened one formal [investigation](#).
- **Regulatory cooperation:** We have engaged extensively with international counterparts, particularly on age assurance. Ultimately, we intend to coordinate our approach on age assurance as far as possible and establish global alignment where we can – in light of future broader online regulation. We have also deepened our collaboration with other UK online regulators via the Digital Regulation Cooperation Forum (DRCF).
- **Research:** To build our evidence base around the appropriateness of certain protection measures, we have conducted or commissioned research on user experience, including the [VSP Tracker](#), [VSP Parental Guidance research](#), [Adult Users' Attitudes to Age Verification on Adult Sites research](#), and behavioural insight research into [the impact of VSP design on user behaviour](#). Additionally, as part of our statutory duty to promote and research media literacy, we have conducted relevant research to further develop our [understanding of online harms](#), including risks to children.

### **This coming year we will focus on how platforms set, enforce, and test their approach to user safety**

In Year 1 we established a baseline for the user protection measures platforms have in place. Our Year 2 priorities will take a broader look at the way platforms set, enforce, and test their approach to user safety, with a particular focus on robust age assurance. In summary, we will seek to:

- **Ensure VSPs have sufficient processes in place for setting and revising comprehensive terms and conditions (generally known as Community Guidelines)** that cover all relevant harms,
- **Check that VSPs apply and enforce their Community Guidelines) consistently and effectively** to ensure harmful content is tackled in practice,
- **Review the tools VSPs provide to allow users to control their experience** and promote greater engagement with these measures, and;
- **Drive forward the implementation of robust age assurance to protect children** from the most harmful online content (including pornography).

Our priorities for Year 2 will support more detailed scrutiny of platforms' systems and processes. We will continue to take forward learnings from our work in holding VSPs to account as we prepare to take on the much broader online regulatory role set out in the Online Safety Bill.

## The structure of this report

- 1.1 Under the VSP legislation, Ofcom has the power to publish reports about the measures taken by platforms for the purposes of protecting users from videos containing harmful material and the ways in which such measures are implemented by each VSP provider.<sup>1</sup>
- 1.2 In summer 2022 we requested information from each of the notified platforms through our formal information gathering powers set out in VSP legislation. The information we received from the platforms forms the basis of this first VSP Report.
- 1.3 In **Part A: Year 1 Report** we provide **an introduction to VSP regulation** and set out our **Key findings in Year 1**<sup>2</sup> across our five strategic priorities and relating to our overall experience regulating the sector. We have used the information provided by the platforms to produce **Our strategic priorities in Year 2**.
- 1.4 In **Part B: Platform Reports**, we present in detail some of the information provided to us by the VSPs about their measures. These detailed platform reports are designed to increase transparency of platform processes and raise public awareness of the measures VSPs have in place to protect users from harmful content.
- 1.5 In addition to information provided by the VSPs, we include throughout this report information and data from Ofcom research and analysis. This information is important because it provides the context in which platforms are making the decisions about their protection measures.
- 1.6 We are simultaneously publishing separate reports including our report into [the VSP Landscape](#), the [VSP Tracker \(March 2022\)](#), the [VSP Parental Guidance research](#), and the [Adults Attitudes to Age-Verification on adult sites research](#) which supplement this document. Information from these reports is included as appropriate throughout the platform reports.

---

<sup>1</sup> Please refer to section 368Z11(1)(b) of the Communications Act 2003.

<sup>2</sup> Only 'An introduction to VSP regulation', 'The key findings in Year 1' and 'Part B' constitute a report for the purpose section 368Z11 of the Communications Act 2003.

## Part A: Year 1 report

## 2. An introduction to VSP regulation

### The VSP Regime in brief

#### Ofcom is the regulator for video-sharing platforms (VSPs) established in the UK

- 2.1 The VSP Regime is set out in Part 4B of the Communications Act 2003 (The Act) and stems from the revised European Audiovisual Media Services Directive (AVMSD) 2018.<sup>3</sup> The requirements for platforms came into effect in November 2020. Ofcom was appointed to regulate the sector from this point on and began developing and implementing the regulatory framework. In this report we refer to the regulatory framework set out in Part 4B of the Act as 'the VSP Framework' or 'the VSP Regime'.
- 2.2 VSP providers in UK jurisdiction are legally obliged to notify their platform to Ofcom. Providers must make their own assessment of whether their platform meets the legal criteria for notification. We have published guidance to assist providers, [Video-sharing platforms: who needs to notify?](#) Currently, 19 platforms have notified Ofcom as meeting the relevant criteria. Many VSPs are not in scope in the UK because they do not meet the jurisdictional criteria, but might be regulated by EU Member States.
- 2.3 We consulted with a broad range of external stakeholders to produce Ofcom's [VSP Guidance](#). This was published in October 2021 alongside [our plan and approach for VSP regulation](#) for the following year (Year 1).

#### VSPs must protect all users from relevant harmful material, with additional protections for under-18s

- 2.4 Under the VSP Regime, VSP providers must protect all users from video content likely to incite violence or hatred against protected groups, and content which would be considered a criminal offence under laws relating to terrorism; child sexual abuse material; and racism and xenophobia – which is referred to as 'relevant harmful material'. They must also protect under-18s from videos containing R18 or unclassified material, and other material that might impair their physical, mental, or moral development ('restricted material').<sup>4</sup> Ofcom refers to 'relevant harmful material' and 'restricted material', collectively, as 'harmful material'. Platforms must also uphold standards around advertising where the advertising is marketed, sold, or arranged by the VSP provider.<sup>5</sup>
- 2.5 Unlike in our broadcasting regulatory duties, Ofcom's role under the VSP Regime is not focused on making decisions about individual videos suspected of containing harmful

---

<sup>3</sup> The revised AudioVisual Media Services Directive was transposed into UK law under regulations made by the Secretary of State, which introduced Part 4B of the Communications Act 2003

<sup>4</sup> For more information on the definition of Restricted Material, see pages 16 – 18 of the [VSP Guidance \(ofcom.org.uk\)](#).

<sup>5</sup> See [Guidance for providers on control of advertising \(ofcom.org.uk\)](#). We will report on the requirements on platforms around advertisements in a later VSP Report.

material.<sup>6</sup> Rather, our role is to ensure platforms have relevant systems and processes in place that provide effective protection to their users from videos containing harmful material. As presence of harmful content alone will not always indicate a systems and processes failure, we must draw on a wide range of evidence in assessing the measures platforms take and in considering the need for any safety improvements, including the nature of the platform and type of content available to users.<sup>7</sup>

2.6 The VSP legislation lists some measures (Schedule 15A measures<sup>8</sup>) that VSP providers must take, if appropriate, to fulfil their duties to protect users from harmful material, which include:

- Having, and effectively implementing, terms and conditions for harmful material
- Having, and effectively implementing, flagging, reporting or rating mechanisms
- Applying appropriate access control measures to protect under 18s like age assurance and/or parental control measures
- Establishing easy-to-use complaints processes
- Providing media literacy tools and information

2.7 VSP providers are required to determine which of the Schedule 15A measures are appropriate for their platform. The VSP Framework sets out that a measure is appropriate for a certain provider if it is practicable and proportionate for that provider to implement it, considering factors including the size and nature of its platform; the type of material on the platform and the harm it might cause; the characteristics of users to be protected, the rights and legitimate interests of users, and any other non-Schedule 15A measures already implemented on the platform.<sup>9</sup>

2.8 You can read more about the regulatory requirements, the harmful material in scope of the VSP Regime, and the measures VSPs must consider taking in the [VSP Guidance](#).<sup>10</sup> The publication of this guidance in October 2021 marked the end of the 'implementation period' of the VSP Regime. We therefore refer to the period October 2021 to the publication of this report as the first full year of regulation, or Year 1. We refer to the twelve-month period following this report as Year 2.

## The VSP Regime will be superseded by the future Online Safety regime

2.9 The Online Safety regime that will be introduced by the Online Safety Bill will eventually supersede the VSP Regime. The timing for the repeal of VSP legislation and details of any

---

<sup>6</sup> The framework for advertising on VSPs places specific responsibilities on VSPs around the content of adverts. See [Statement: The regulation of advertising on video-sharing platforms \(ofcom.org.uk\)](#) for more details.

<sup>7</sup> The practicable and proportionate criteria that platforms must have regard to when determining which measures are appropriate, are also relevant considerations for Ofcom. You can read more about these criteria in Section 6 of the [VSP Guidance \(ofcom.org.uk\)](#).

<sup>8</sup> Throughout the report we refer to Schedule 15a measures as user protection measures, safety measures or simply measures.

<sup>9</sup> Please refer to section 368Z1(1) and (4) of the Communications Act 2003.

<sup>10</sup> VSPs also have duties with regard to advertising on top of their duties to protect users from videos containing harmful material. These are set out in Ofcom's [Guidance for providers on advertising harms and measures \(ofcom.org.uk\)](#). We will report on the requirements on platforms around advertisements in a later VSP Report.



transitional period for in-scope platforms will be set out by Government. We will continue to engage with Government on transitional arrangements and communicate these to platforms as soon as we can.

## 3. Our approach to VSP regulation in Year 1

### Our strategic priorities for Year 1

3.1 We set out Ofcom's approach and plan of work for Year 1 in our [Plan and Approach document](#). We identified five strategic priorities for Year 1 of VSP regulation. Across these, the overall goal was to ensure that VSPs are taking appropriate measures to protect their users from harmful videos. Our priorities were:

- **reducing the risk of child sexual abuse material (CSAM)** by improving registration and moderation processes on adult VSPs,
- **tackling hate and terror** by ensuring platforms' terms and conditions effectively protect users,
- working with VSPs popular with children to ensure they are providing an **age-appropriate experience**,
- laying important foundations for **age verification** on adult VSPs, and
- ensuring that VSPs' **flagging and reporting processes** are effective and VSPs increase user engagement.

### Establishing supervisory relationships with VSPs

3.2 During Year 1, we established constructive supervisory relationships with each of the notified platforms to deliver across our five strategic priorities. We engaged with a range of individuals and teams within the VSPs – at the smaller platforms these tended to be CEOs, while at larger platforms we generally met with Policy or Trust & Safety teams – to learn more about the measures in place ahead of sending formal information requests.

3.3 We saw some platforms make positive changes to their systems and processes in light of new VSP regulatory requirements. For example:

- a) **OnlyFans has adopted age verification tools for all new UK subscribers.** The primary product is provided by Yoti, a third-party software company that uses facial age estimation technology, along with liveness, anti-spoofing and document authenticity checks to verify the age of potential users. Potential subscribers who fail Yoti's check are directed to Ondato, another third-party age assurance software provider.
- b) **TikTok established an Online Safety Oversight Committee modelled on existing data protection governance processes.** The OSOC provides "executive oversight of content and safety compliance"<sup>11</sup> specifically within the UK and EU region. The OSOC includes the Global Head of Trust & Safety, the General Manager for the UK and EU, and the Director of Legal and Corporate affairs.

---

<sup>11</sup> Quote taken from TikTok's response to the formal information request issued in the summer of 2022.

- c) **Vimeo now allows only material rated 'all audiences' to be visible to users without an account.** Content rated 'mature' or 'unrated' is now automatically put behind the login screen.
- d) **Platforms have introduced a range of other safety measures, including Snap's Family Center.** Snapchat recently launched its parental control feature, Family Center, which allows parents and guardians to view a list of their child's conversations without seeing the content of those message. Parents and guardians can also view who their child is friends with on the service as they are given access to a list of their most recently-added friends.

3.4 We have also worked with VSPs to understand which platforms fall in scope of the VSP Regime, including supporting platforms through the process. Two small adult VSPs decided to close the video sharing sections of their sites after their provider concluded that they would not be able to comply with Ofcom's expectations around age verification. Another VSP, Brand New Tube, notified in October 2022 following a period of close engagement.

## Gathering information from VSPs

- 3.5 In the summer of 2022, we sent VSPs requests for information using our formal regulatory powers set out in VSP legislation. Ofcom has the power to use information requested from platforms for the purposes of preparing a VSP Report.<sup>12</sup> Platforms are legally required to respond to all information requests in a full, accurate, and timely manner. If they do not, we have the power to take enforcement action.
- 3.6 The information we requested from platforms corresponded to our five Year 1 strategic priorities. However, as the platforms we regulate vary greatly in size and nature we did not ask for the same amount of information from each platform.
- 3.7 Over the course of the information gathering process we engaged with several VSPs about potentially incomplete and delayed information responses. Overall, this largely resulted in positive engagement and the platforms submitting the necessary information.
- 3.8 However, on 29 September 2022 we opened an investigation into whether Tapnet Ltd, trading as RevealMe, failed to comply with its duties to comply with a formal information request.<sup>13</sup> While Tapnet Ltd provided its response after the investigation was opened, this has impacted on our ability to comment on RevealMe's protection measures in this report.

---

<sup>12</sup> Please refer to section 368Z10(3)(h) of the Communications Act 2003.

<sup>13</sup> [Investigation into Tapnet's compliance with a statutory information request \(ofcom.org.uk\)](https://www.ofcom.gov.uk/consult/condocs/tapnet/tapnet_220929.pdf)

## 4. Key findings in Year 1

### Summarising our learnings from Year 1

**All platforms have safety measures in place, including rules on what kinds of video material are allowed.** Some platforms made changes to their measures in direct response to being regulated under the VSP Regime.

**VSPs generally provided limited evidence on how well their safety measures are operating to protect users.** As a result, this creates difficulty for Ofcom in determining with any certainty whether VSPs' safety measures are working consistently and effectively.

**More robust measures are needed to prevent children accessing pornography.** We have continued to build our evidence base on the efficacy of age assurance measures. Based on our analysis so far, we think some adult VSPs' access control measures might not be sufficiently robust in stopping children accessing pornography – and we expect improvements.

**Some VSPs could be better equipped for regulation.** Some platforms, across all sizes, are not prepared or sufficiently resourced for regulation. Ofcom is working with platforms to improve their regulatory preparedness and will be standing up a dedicated supervision team to support this work. Going forward, we will be looking for platforms to improve and provide more robust and timely responses to Ofcom's information requests, and connect us to the appropriate expert teams where necessary.

**Platforms are not prioritising risk assessment processes,** which Ofcom believes are fundamental to proactively identifying and mitigating risks to user safety. Risk assessments will be a requirement on all regulated services under the Online Safety regime.

### Progress against our five strategic priorities

4.1 As explained in paragraph 3.1 above, in October 2021 we [set out](#) our priority areas of focus for the year. These priorities have shaped our work, through research, engagement, and how we structured our information requests to the platforms. We set out progress made against our priorities below.

#### Reducing the risk of child sexual abuse material (CSAM)<sup>14</sup>

4.2 Self-generated content is an increasingly significant driver of child abuse images and videos. The Internet Watch Foundation (IWF) reported a 163% increase across the internet in the amount of self-generated CSAM in 2021 as compared with 2020;<sup>15</sup> and a 360%

---

<sup>14</sup> CSAM includes any material which shows the sexual abuse of a child. Under the VSP Framework the definition of CSAM covers the depiction of any person appearing to be a child, as well as realistic images of CSAM (such as computer-generated content) and simulated activity.

<sup>15</sup> [The Annual Report 2021 \(iwf.org.uk\)](#). The webpages assessed by the IWF are a result of reports and proactive searching; the data may not be entirely reflective of material found on the adult VSPs discussed in this report as the material is behind a paywall.

increase in the amount of self-generated CSAM depicting seven to 10-year-old children in the first half of 2022<sup>16</sup> the same period in 2021.

- 4.3 Our [VSP guidance](#) sets out a clear requirement that having and effectively implementing terms and conditions to prohibit CSAM is critical for protecting users and complying with the VSP Regime. Adult VSPs carry a heightened risk regarding the uploading of this material, and so our starting point in tackling CSAM was to focus on the protection measures that adult VSPs have in place, including registration and onboarding processes and expedited process for user reports relating to CSAM.
- 4.4 To build on our requirement that platforms need to have and effectively implement measures to prevent CSAM, we have undertaken a programme of engagement – including with NGOs and civil society – to understand the key drivers of CSAM-related harm, and what the most effective mitigations are in the online space. We have engaged with adult VSPs to understand the action they are taking to protect their users from CSAM.

#### **The adult VSPs that responded all have some CSAM prevention measures in place**

- 4.5 Ofcom requested information about CSAM from six of the notified adult VSPs: OnlyFans, AdmireMe, FanzWorld, PocketStars, RevealMe, and Xpanded.
- 4.6 Although further assessment will be carried out to verify the information we received, we have found that all the notified adult VSPs who responded to us have:
- user rules in place to prohibit uploading illegal material, including CSAM. These rules contain clear sanctions for users who breach them;
  - creator registration and onboarding processes in place that seek to prevent the uploading of CSAM, which appear to Ofcom to be adequate, though no assessment of effectiveness has been conducted;
  - some reporting and moderation mechanisms in place, including prioritising user reports relating to CSAM. Notable here is the fact that all responding adult VSPs use some form of proactive moderation<sup>17</sup> on their platforms; and
  - more work to do in the way they assess and evidence the effectiveness of their measures.
- 4.7 More detail on these measures and how they appear to reduce the risk of CSAM can be found in the OnlyFans and smaller adult VSPs sections in **Part B** of our report.

---

<sup>16</sup> [20,000 reports of coerced 'self-generated' sexual abuse imagery seen in first half of 2022 show 7- to 10-year-olds \(IWF,org,uk\)](#)

<sup>17</sup> 'Proactive moderation' refers to the measures taken by platforms to monitor content online, including through scanning technologies, in order to detect protection violations of their terms and conditions. These measures are generally implemented alongside inputs from user reporting tools. Proactive moderation is not required under the VSP Regime and is therefore not part of this work but is nevertheless of interest to Ofcom.

## Tackling hate and terror

- 4.8 Our 2022 research found that 24% of VSP users said they had come across videos they perceived to be violent, abusive, or inappropriate videos in the last three months.<sup>18</sup> Our [VSP Guidance](#) sets a clear expectation that having and effectively implementing terms and conditions that prohibit the uploading of relevant harmful material is central to compliance. This includes both videos constituting a criminal offence under laws relating to terrorism, racism, and xenophobia, as well as material likely to incite violence or hatred.
- 4.9 Our aims under this priority were to ensure platforms' terms and conditions align with the legal requirements, and to understand how they are communicated to users.
- 4.10 During Year 1, we engaged with VSPs – and other relevant stakeholders – to understand potential areas for improvement as well as notable good practices.
- 4.11 On 12 October 2022, we published a [report](#) on the May 2022 shooting in Buffalo, New York, which was initially livestreamed on Twitch, a notified VSP and then shared on other platforms. The report set out our findings on the role that platforms play in facilitating and preventing the rapid spread of terror footage following such attacks.

### Terms and conditions are generally adequate, though there is some room for improvement

- 4.12 The notified VSPs generally have adequate terms and conditions (sometimes referred to as Community Guidelines) that prohibit material that would come within the scope of laws relating to terrorism, racism, and xenophobia, as well as material likely to incite violence or hatred (content we refer to collectively as 'hate and terror content'). We found platforms largely also have a range of adequate processes for updating these as required.
- 4.13 However, there is room for improvement for some platforms regarding what is prohibited in their terms and conditions and how they communicate these rules to users (e.g. communicating T&Cs during the sign-up process, during the upload process, after a violation etc.). We found that some VSPs' terms and conditions were dense and hard to follow and, in some cases, users do not even need to open the terms and conditions in order to watch content.
- 4.14 VSPs generally apply different types of sanctions, such as removal of content or suspension of a user account, when these T&Cs are breached. Explanations of sanctions and appeals processes is also an area where platforms can improve. For example, we think some VSPs should be quicker at reviewing reported content, while for others the appeals process is unnecessarily opaque.
- 4.15 Finally, as in other areas, we need to deepen our understanding of whether terms and conditions on hate and terror content are consistently and effectively enforced. We will explore this further in Year 2.

---

<sup>18</sup>[Ofcom VSP tracker Wave 2 \(ofcom.org.uk\)](#) (March 2022), Q3a

## Protections for under-18s

- 4.16 Last year, our VSP research found that 79% of 13-17-year-olds who use VSPs said that they had experienced harm on a VSP in the last three months.<sup>19</sup> Users in this age group were also significantly more likely than adult VSP users to say they had been exposed to content they perceived to be harmful online, including negative body image and eating disorder content, content glamourising unhealthy or abusive lifestyles, and the promotion of self-harm.<sup>20</sup> There is a lot more to do to ensure that under-18s are suitably protected online.
- 4.17 In Year 1 we made clear that platforms need to tailor their safety efforts to ensure an age-appropriate experience for all their users. This includes ensuring users of all ages can understand and engage with safety measures and promoting media literacy. Our key aim for this priority was to be able to report in detail on the measures VSPs have in place to protect under-18s and ensure content is age appropriate – improving transparency across platforms for users, parents, and carers.
- 4.18 Over the past year we have focused on understanding the safety measures available to children and their families. This has included research with parents and carers to understand how they engage with guidance and child safety measures.<sup>21</sup> We have also engaged with civil society groups, including hosting a roundtable discussion to better understand children's experiences of VSPs and relevant safety measures.

### More work needs to be done on providing age-appropriate experiences to under-18s

- 4.19 VSPs are at an early stage in providing age-appropriate experiences and do not tend to distinguish between age groups. Most platforms only consider whether a user is the minimum required age (or 18+). Currently, only two of the VSPs (TikTok and The Sponsor Hub) distinguishes between different age groups to adjust the nature of the content and functionalities younger users can access.
- 4.20 Some platforms have introduced new parental controls that empower parents to manage their children's online experiences. For example, TikTok's Family Pairing allows parents to limit the visibility of content with a warning notice through Restricted Mode. Snapchat's Family Center provides parents with increased transparency in relation to their child's conversations and friends.
- 4.21 Platforms also have various ways of informing users that content may be inappropriate for young users. Twitch enables streamers to tag their channels as mature if their content is not intended for younger audiences. TikTok categorises content that may be unsuitable for

---

<sup>19</sup> Ofcom, [Video-sharing platform usage & experience of harms survey 2021 \(ofcom.org.uk\)](https://www.ofcom.gov.uk/consult/condocs/vsp/vsp2021/vsp2021.pdf) These perceived harms do not necessarily map onto the definition of relevant harmful material under the VSP Regime. Even when these harms could potentially fall under material likely to cause harm to the physical, mental, or moral development of under-18s, this may be difficult to establish without assessing individual pieces of content.

<sup>20</sup> Ofcom, [Video-sharing platform usage & experience of harms survey 2021 \(ofcom.org.uk\)](https://www.ofcom.gov.uk/consult/condocs/vsp/vsp2021/vsp2021.pdf).

<sup>21</sup> [VSP Parental Guidance research \(ofcom.org.uk\)](https://www.ofcom.gov.uk/consult/condocs/vsp/vsp2021/vsp2021.pdf)

younger users, notifying under-18s that this content cannot be viewed as it is 'age-protected'.

- 4.22 Some providers have made changes to the design of their platform over the course of their engagement with Ofcom. For example, Vimeo now allows only material rated 'all audiences' to be visible to users without an account. Users must log in to an account to see non-rated and mature content. Users with an account can also opt-in to filter out non-rated or mature content from their search results.
- 4.23 In Year 2, we will drive further change in this important area. We will have a particular focus on the types of content that may be particularly harmful to children, including self-harm and suicide material. We will work with platforms to consider how effectively these tools protect under-18s and look for substantial improvements.

## Age verification on adult VSPs

### We said we would continue laying important foundations for age verification on adult VSPs

- 4.24 We consider VSPs specialising in pornographic material to be 'adult VSPs'. In our [VSP Guidance](#), we list a number of indicators we might consider when assessing whether a platform needs to implement measures to protect under-18s from pornographic material. These include how much pornography is on the platform, the significance of pornography to the platform, and the way the platform is positioned on the market.
- 4.25 Laying foundations for age verification to prevent under-18s from accessing pornographic material on adult VSPs was one of our priorities for Year 1.
- 4.26 As part of laying these foundations, Ofcom commissioned research into user attitudes to age verification. The key findings have been published alongside this report and are summarised in the [Introduction to Adult Platforms](#) section in Part B. The research found that, while participants were broadly supportive of age verification being used to protect under-18s from harm, the risk of harm from pornography felt less tangible than from other activities, such as online gambling. Participants felt more willing to verify their age to access pornography when creating an account or subscribing to content.<sup>22</sup>
- 4.27 We have engaged with providers of age assurance solutions to build our understanding of what technology is currently able to achieve, and future trends in this area. We plan to continue this engagement and deepen our understanding of this market, including the development of technical standards for age assurance. We recognise the need for overall progress in this area.
- 4.28 We have also engaged regularly with our international counterparts through the International Working Group on Age Verification, as well as the ICO to ensure that age verification solutions protect the privacy of the user.

---

<sup>22</sup> [Ofcom's Adult Users' Attitudes to Age Verification on Adult Sites Research, 2022](#)



### OnlyFans has adopted an age verification solution, but smaller adult VSPs have further to go

- 4.29 OnlyFans, the largest of the adult VSPs, has implemented Yoti, an age assurance solution, for all new UK users following the VSP regulatory requirements. Ofcom is pleased to see a VSP provider taking steps to implement age verification processes that appear more robust than the traditional and much more prevalent self-declaration of date of birth.
- 4.30 In contrast, the information received from the smaller adult VSPs did not provide us with confidence that they have robust access control measures in place to prevent children accessing pornography. We will continue to engage with these platforms during Year 2 to drive forward the implementation of age verification.

### Reporting and flagging

- 4.31 We outlined in our [VSP Guidance](#) that we consider reporting and flagging mechanisms fundamental to the protection of users and that these mechanisms should be easy to use. In addition, actions taken in response to reports or flags should be clear, transparent, appropriately timely and proportionate to the size, nature, and risk profile of the platform.
- 4.32 Reporting and flagging is a tool for users and third-party organisations to alert platforms of content that potentially violates a platform's terms and conditions. It can be both a tool for users to take control of their online experiences, and a way for platforms to detect and potentially take action on a piece of content or account. As such, reporting and flagging has a strong relationship with various other protection measures.
- 4.33 Reporting and flagging tools can often function as a last line of detection for harmful or inappropriate content that is not detected by other means, such as automated classifiers. Even as algorithmic tools improve, user reporting can remain an important protection measure for some platforms. User reporting can at times be more precise when it comes to detecting content that requires additional context or contains nuance (e.g. humour, satire, or implicit claims).
- 4.34 As reporting and flagging mechanisms both allow platforms to detect harmful content as well as empower individual users, our Year 1 priorities have been focused on understanding which platforms have implemented these tools, how they have been implemented, and how they can be improved. This included research, engaging with platforms, taking in the views and experiences of civil society organisations, and running online behavioral insights experiments on reporting features. In order to raise awareness for the importance of user reporting and how to do it, we ran [a campaign on reporting and flagging](#) on several social media platforms.<sup>23</sup> The aim was to highlight directly to younger users that if they see video-content that appears to be harmful, they should report it to platforms, as well as to provide information on how they can do this.

---

<sup>23</sup> The campaign ran on [TikTok](#) and [Instagram](#).

## Reporting and flagging is in place on all notified VSPs, but its use and application varies

- 4.35 We found that reporting and flagging is an industry-wide standard practice; all of the currently notified platforms have reporting and flagging mechanisms in place. However, there is significant variation in how they are implemented. There is also variation in how integral they are to platforms' detection and enforcement processes, with some platforms relying much more heavily on automated tools for proactive removal or detection of content breaching their terms and conditions.
- 4.36 This means that Ofcom has more to explore when it comes to understanding how effectively harmful content is identified and acted upon via user reporting and flagging systems. We have carried out [research](#) into the ways in which the design of reporting and flagging mechanisms impacts user behaviour. Various user interface design changes were made to a generic online video viewing portal. The results provide evidence for the use of more visually-prominent reporting tools to increase the likelihood of VSP users reporting potentially harmful videos.
- 4.37 Additionally, most large VSPs have some form of external 'trusted flagger' programme where platforms partner with organisations (such as civil society, government agencies, or other relevant groups) who have a specific expertise in online harm. These 'trusted flaggers' can directly flag potentially violating content to the platforms, which then prioritise reviewing those reports. However, there is no agreed industry best practice for trusted flagger programmes. Going forward, we will work with platforms to better understand how these programmes work, who is eligible to be a trusted flagger, and how reports are prioritised.

## Additional learnings from VSP regulation in Year 1

### Platforms need to better prepare themselves to engage with regulation

- 4.38 Throughout the last year we have seen a large variation in platforms' readiness to engage with Ofcom. This variation had no correlation with the size of userbase or revenue of the platforms, and was particularly evident from the way in which different providers responded to our information requests. Some platforms were fully responsive in the information they provided, while the initial responses from others were less satisfactory. We are conscious of the need to be proportionate and not impose undue burden on regulated companies. However, we have made it clear that we expect better engagement from regulated platforms – of all sizes – and that we will not hesitate to take action where we have concerns that a provider is not meeting its duty to cooperate with the regulator.
- 4.39 VSPs need to properly resource themselves to engage with regulation, whether that be by ensuring they are including the correct staff in meetings with Ofcom or responding to information requests fully and by the correct deadline. In some cases, platforms will need to seek appropriate legal guidance when engaging with regulatory processes. Open, proactive communication between platform and regulator leads to more constructive engagement. We expect to see the platforms we regulate under the VSP regime mature in

their relations with us in the coming year – tailoring this expectation in a way that is proportionate to the size and riskiness of each platform.

- 4.40 We also found that most VSPs generally did not understand what risk assessments are and how to perform them. This suggests that these platforms – and wider industry – need to increase their regulatory preparedness at pace, as risk assessments are expected to become mandatory for all regulated services under the Online Safety Act.

## **International collaboration is central to our approach**

- 4.41 Central to Ofcom's future ambitions under the VSP Regime is our international collaboration on global issues. We recognise the value of working together with other regulators to highlight best practice, identify common risks, and support platforms in developing their approach to complying with different regulatory regimes.
- 4.42 Since we started regulating VSPs, we have attended regular meetings with our counterparts in France, Germany and Cyprus to share best practice relating to our regulatory approaches to age verification. Over the coming year we will continue to work on our collective understanding of assessing proportionality, developing our understanding of consumer attitudes towards age assurance technologies, and creating consensus internationally.
- 4.43 One practical example of where our international work can foster positive change relates to age verification. Some UK-based adult VSPs are concerned that if they were to implement age verification, they would lose customers to platforms based outside of the UK's jurisdiction. As a result, we are seeking to proactively level-up age verification standards applicable to the full range of adult VSPs accessed by UK users, through collaboration with other National Regulatory Authorities.<sup>24</sup> International collaboration and information sharing will be a central priority over the coming year as we seek to implement our Year 2 priorities.

## **Significant events can teach us about the systems and processes platforms have in place**

- 4.44 In our role as VSP regulator, it is not for us to decide if a particular piece of content should or should not be allowed or whether it complies with a platform's terms and conditions. Rather, our role is to ensure that platforms are taking effective measures to protect users from harmful content. Accordingly, the way in which platforms respond to significant events, such as the rapid online spread of footage from the May 2022 shooting in Buffalo, New York, can provide Ofcom with an insight into the strength of a platform's systems and processes. Our report on [The Buffalo Attack: Implications for online safety](#), published on 12 October 2022, touches on our findings relating to three notified VSPs in this report.

---

<sup>24</sup> National Regulatory Authorities' is the term Ofcom uses to refer to authorities in other jurisdictions and international organisations who regulate the same sectors that we do.

- 4.45 We want to see evidence of consistently-applied, effective action to deal with violating content and accounts. The action taken by platforms helps Ofcom understand how their internal detection systems operate and how terms and conditions are enforced. We have found it useful to see in practice how these high-profile events highlight the tradeoffs platforms may face when deciding how to balance keeping users safe from harms and enabling the free and frank exchange of ideas and opinions.

## 5. Our strategic priorities in Year 2

### Our Year 2 priorities build on our Year 1 findings

- 5.1 In Year 1, we engaged with platforms on the measures they use to address the most harmful forms of video content – namely CSAM, hate and terror content, and content that is likely to harm under-18s. As set out above, we found some cross-cutting issues.
- 5.2 All platforms have adequate terms and conditions that prohibit materials that would be in scope of the legislative requirements. However, there is still room for improvement, including in how these terms and conditions are updated. Furthermore, none of the platforms supplied sufficient information for us to assess how effectively and/or consistently their user policies are applied and enforced.
- 5.3 Platforms also have a range of systems to identify and prevent the dissemination of harmful video content, including a mixture of both proactive and reactive content moderation. However, we were also unable to determine whether and how they internally assess the effectiveness of these systems.
- 5.4 We will therefore take a broader look at these cross-cutting issues in Year 2 and focus on **what VSPs are doing to set, enforce, and test their approach to user safety**. We will also deepen our technical knowledge about platforms' systems and technologies, identifying and sharing good practice where we can.
- 5.5 In doing so, we will continue to focus on the most harmful types of content, including CSAM, hate and terror content, and video content that may be harmful to under-18s.

### Our priorities for the year ahead

- 5.6 We have identified four priority areas of focus for Year 2 of VSP regulation. Our overall goal across these priority areas continues to be ensuring that VSPs are taking appropriate measures to protect their users from harmful material.
- 5.7 Our strategic priorities for Year 2 of the VSP Regime are for Ofcom to seek to:
- **ensure VSPs have sufficient processes in place for setting and revising comprehensive user policies** (generally known as Community Guidelines) that cover all relevant harms,
  - **check that VSPs apply and enforce their Community Guidelines consistently and effectively** to make sure harmful content is tackled in practice,
  - **review the tools VSPs provide to allow users to control their experience** and promote greater engagement with these measures, and
  - **drive forward the implementation of robust age assurance**, to protect children from the most harmful online content (including pornography).
- 5.8 These are the areas we will dedicate most of our attention to, but we know that other issues might arise that need prioritising due to the severity of harm or risk to users. We will continue to learn and adapt as we go, as we did in Year 1. We discuss each in turn below.

## Setting Community Guidelines that cover all relevant harms

**We will seek to ensure that VSPs have sufficient processes in place for setting and revising comprehensive Community Guidelines that cover all relevant harms**

- 5.9 In our VSP Guidance published in October 2021, we said that having and effectively implementing terms and conditions for harmful material is central to compliance with the VSP Regime. We have established that notified VSPs all broadly have terms and conditions that prohibit the sharing of forms of relevant harmful material (including CSAM and hate and terror content) and set rules for what content is appropriate for under-18 users. However, we have identified that there is still some room for improvement.
- 5.10 We will focus on ensuring VSPs' terms and conditions cover the full breadth of harmful material described in the VSP Regime. We also want to see VSPs continually reviewing their terms and conditions to make sure they are updated as needed and that these are communicated effectively to users. Ensuring VSPs have effective processes for updating their terms and conditions will be a core foundation for all other measures VSPs have in place.
- 5.11 By the time Ofcom publishes its Year 2 report, we expect all VSPs to have terms and conditions that cover the full breadth of harmful material under the VSP Regime and enable platforms to take effective action to remove or restrict content as appropriate. We are also interested in how VSPs balance user safety with other considerations where competing rights and interests are at stake, although we recognise that VSPs do not have any statutory obligations regarding freedom of expression.
- 5.12 We also expect to have built a detailed understanding of VSPs' internal practices for reviewing and updating terms and conditions and policies on user sanctions, pushing for improved internal processes where needed. We will want platforms to demonstrate the steps they take to assess the effectiveness of these processes.

## Applying and enforcing Community Guidelines to tackle harmful material

**We will seek to check that VSPs apply and enforce their Community Guidelines consistently and effectively to ensure harmful content is tackled in practice**

- 5.13 The VSP Regime requires that where a measure is taken, it must be implemented in a way that achieves the purpose for which it is designed. We said in our guidance that, for terms and conditions (generally referred to as Community Guidelines) which prohibit illegal content, they must be enforced effectively. Sometimes effective enforcement can involve processes or technologies that fall outside the measures listed in the VSP legislation (in Schedule 15A).
- 5.14 Ofcom has received a range of information about platforms' moderation, training, and sanctions processes. However, we found that there was insufficient information to indicate how effectively and consistently user policies are being enforced via content moderation processes (see paragraphs 5.8 - 5.15). In this context, 'content moderation' refers to the process by which a video is reviewed after it has been detected by a platform (e.g., via a

user report) and what happens next (e.g. from doing nothing to taking it down). The phrase 'content moderation' can sometimes refer to proactive measures taken by platforms to monitor content online, including through scanning technologies. Such proactive content moderation is not required under the VSP Regime and is therefore not part of this work but is nevertheless of interest to Ofcom.

- 5.15 Ofcom therefore wants to take a deeper look at how cross-cutting content moderation review processes work to tackle harmful content. For example, we want to explore how moderators are trained, how VSPs measure and improve accuracy, and how problematic content is identified and prioritised from reports received. We want to make sure VSPs' content moderation review processes tackle harmful content and result in a clear and consistent approach to enforcing the platforms' Community Guidelines.
- 5.16 By the end of Year 2, we want to be confident that VSPs pick up and quickly remove or restrict harmful content – once it has been reported – to ensure that users are better protected from encountering such content. We also want to continue building our understanding on the extent to which the efficacy of these measures can be evaluated, sharing our findings where possible.

## Empowering users to control their VSP experience

### We will seek to review the tools VSPs provide to allow users to control their experience and promote greater engagement with these measures

- 5.17 One of our Year 1 priorities was ensuring that VSPs' flagging and reporting processes were effective and that VSPs increase user engagement with them. We found that there is more work to do to understand how effective these measures are on different platforms. Our research also suggests that engagement with these tools is low – and we expect platforms to educate users about available measures as part of promoting media literacy.<sup>25</sup>
- 5.18 We are expanding our focus in Year 2 to find out more about the availability and effectiveness of a broader range of user empowerment tools. This includes reporting and flagging, as well as tagging, rating, parental controls, complaints, and appeals. We want to understand how these tools operate to protect users and, help maintain free and frank conversations where appropriate.
- 5.19 Parents and carers should feel empowered to know how to control their online experiences or those of their children so they can feel safer online. Our [VSP Parental Guidance research](#), published today, shows that parents would like to be made more aware of the safety information and tools provided by VSPs. Ofcom will be exploring how we can promote greater awareness of, and engagement with, such measures. We are also

---

<sup>25</sup>[Just one in six young people flag harmful content online \(ofcom.org.uk\)](#)

particularly interested in how tagging can and does operate to protect children from restricted material.<sup>26</sup>

- 5.20 At the end of Year 2, we expect to have a more detailed evidence base into how user empowerment tools operate and what works best – sharing learnings with VSPs and the public where we can. This should support VSPs in developing even more effective and engaging empowerment tools for users and parents. And ultimately if VSP users, parents, and carers know how to control their experience online, they will feel safer and more confident when using platforms.

## Protecting children from the most harmful video material

### We will seek to drive forward the implementation of robust age assurance, to protect children from the most harmful online content (including pornography)

- 5.21 As we noted in our [VSP Guidance](#), we think adult VSPs should make robust age verification a priority to ensure viewers are 18 or over. We laid some important foundations for age verification on adult VSPs during Year 1, engaging with other domestic and international regulators and building our evidence base on what measures are in place on VSPs. As noted above, we have concerns that some of the adult VSPs do not have robust age assurance measures.
- 5.22 This focus will continue during Year 2. We will continue our programme of research and engagement, building up the picture of what solutions work and are proportionate to expect from different VSPs. We will also continue to engage with adult VSPs – including OnlyFans on its age verification technology – and share learnings with the wider industry where we can.
- 5.23 As the VSP Regime is only applicable to UK-established VSPs, many adult VSPs are beyond Ofcom's jurisdiction. Therefore, protecting children from adult material ultimately demands a global response. Ofcom will continue to work closely with international regulators to understand each other's approaches to assessing the proportionality and efficacy of age assurance technologies and to seek to level-up protection for UK users across the full range of adult VSPs, irrespective of where those VSPs are established.
- 5.24 Ofcom continues to work closely with the ICO to share learnings, ensure clarity of roles and responsibilities, and make sure our respective approaches are coherent and aligned (where relevant). We have published joint [research on age assurance measures](#) and we plan to publish a joint statement on our high-level approach to collaborative working later this year.
- 5.25 By the end of Year 2, we want notified adult VSPs to have in place a clear roadmap to implementing robust age assurance measures – and having them in place where that is proportionate. In the longer term, we want to ensure children are unable to readily access

---

<sup>26</sup> This is relevant to the Schedule 15A measure to include terms and conditions to the effect that if a person uploads to the service a video that contains any restricted material, that person must bring it to the attention of the person who is providing the service.



adult VSPs, protecting them from content that could cause them harm. It is essential we lay the basis for this ahead of the Online Safety regime, where Ofcom's jurisdiction will not be limited in the same way and many more services will be required to protect children from adult content.

## Looking ahead to Online Safety

- 5.26 One of Ofcom's strategic aims for the VSP Regime is to prepare industry, and ourselves, for the upcoming Online Safety regime. We have made progress on this in Year 1, including by recruiting and developing skills and expertise that will be vital to the successful implementation of the Online Safety regime.
- 5.27 While every harm is different, we will expect platforms to take tailored and proportionate approaches to protecting users, ensuring that their cross-cutting measures are working well. Some harms require collective action across the whole sector – otherwise harms and bad actors will simply migrate from services with more effective protections to those with weaker systems. As previously noted, this is particularly relevant to the implementation of robust age verification on adult platforms – age verification will only be effective if it is adopted across the entire industry. We will want to take a comprehensive approach to Online Safety, including working with other regulators in the UK and overseas where appropriate.
- 5.28 Conducting product risk assessments, both before deploying new measures and in reviewing the effectiveness of existing ones, is critical to effective risk management by platforms. Ahead of the introduction of the Online Safety Bill, we encourage services to review their risk assessment processes, consider how effective they are, and explore opportunities for improvement, particularly with respect to the illegal priority harms in the current version of the Bill. As set out in our July 2022 [Online Safety Roadmap](#), we will expect all firms to consider how they prioritise user protection and incorporate safety considerations into product and engineering design decisions, once the online safety regime comes into force. We expect companies to have robust internal governance processes in place to discuss these issues including with senior decision makers, where appropriate.
- 5.29 Our findings on the lack of regulatory preparedness of some VSPs (see paragraphs [5.38-5.40](#)) demonstrate that there is likely to be more work to do for many platforms before the arrival of the Online Safety Regime. Notified VSPs will be in a strong position to build on their existing experiences engaging with Ofcom, but we are keen that these learnings are shared with the wider set of platforms likely to be in scope of the future regime.
- 5.30 We anticipate issuing a number of formal information requests when Parliament passes the Online Safety Bill. As under the VSP regime, relevant platforms will likely have an opportunity to comment on these information requests before they are formally issued. While we welcome suggested changes that help ensure we receive the most relevant information, we will expect services to engage constructively and promptly throughout this process. Transparency, which sits at the heart of both VSP and Online Safety regulation,

relies on our ability to extract meaningful information from platforms on the different ways they protect their users from harm, highlighting areas where they fail to do so effectively.

## Part B: Platform reports

## Guide to the platform reports

This part contains the following platform reports:

**TikTok**

**Snapchat**

**Twitch**

**Vimeo**

**BitChute**

**Smaller VSPs** (Thomas Cook, The Sponsor Hub, Fruitlab, Qurio, Recast Sport)

**OnlyFans**

**Smaller adult VSPs** (Admire Me, Fanzworld, Pocket Stars, RevealMe, Xpanded)

## Structure of the reports

### About the platforms

Each section begins with information about the platforms in order to provide the reader with the context in which their measures are being implemented.

### Governance and risk management

We explain the internal governance processes within each regulated service, describing any systems in place for online safety risk management

### User journey

We set out the protection measures in place along the user journey for each platform or group of platforms.

**Note:** We received information from some VSPs that is not included in these platform reports. Information has been redacted where we understand that the publication of the information might seriously and prejudicially affect the interests of the VSP. We have also redacted information that we have concerns may help users evade the protection measures on a VSP.

## 6. TikTok

### Introduction

TikTok is a video-sharing app owned by the Chinese company ByteDance. It is highly popular in the UK, particularly with younger audiences. TikTok uses an algorithmic recommender system to deliver content to users. Its personalised 'For You' page displays a continuous feed of videos tailored to each user,<sup>27</sup> taking into account individual preferences as expressed through interactions with the app. Users can also search for content via the 'Discover' page. TikTok is known for its short form videos,<sup>28</sup> the power of its algorithmic feed, and the speed with which user-uploaded videos can go 'viral'. Though the platform is often described as hosting dance videos, its content ranges widely from lighthearted satire and makeup tutorials to news footage and cultural commentary. TikTok presents videos paid for by advertisers to users alongside user generated content.

Figure 6.1: Key information on TikTok



Source: © Ipsos, Ipsos iris Online Audience Measurement Service, 1 - 30 April 2022, age 18+ for reach and time spent data, age 15+ for audience share by gender data, UK.

### Key findings

- 6.1 **TikTok has a broad range of protection measures in place.** This is reflective of its large audience size, particularly among younger demographics. It is therefore appropriate that it invests in a variety of protection measures.
- 6.2 **TikTok relies predominantly on proactive detection of harmful video content,** rather than relying on reactive user reporting. User reporting represents a very small part of harmful content detection on TikTok with user reports leading to just 4.6% of videos removed.

<sup>27</sup> TikTok's recommender system takes into account user preferences as expressed through interactions with the app, like posting a comment or following an account. Source: TikTok, [How TikTok recommends videos #ForYou](#), 18 June 2020.

<sup>28</sup> In 2021 TikTok increased the maximum video length time from one minute to three minutes. Source: TikTok, [More Tok on the Clock: Introducing longer videos on TikTok](#), 1 July 2021.

- 6.3 **TikTok has a governance structure in place to oversee decision making around user safety.** TikTok's recently-established Online Safety Oversight Committee has "executive oversight of content and safety compliance"<sup>29</sup> specifically within the UK and EU region.

## Engagement with Ofcom

- 6.4 TikTok's engagement with Ofcom has been largely positive; both sides have worked to establish an effective regulatory relationship. TikTok has provided briefings and attended meetings to discuss different areas of the platform, bringing along relevant staff to explain how measures work and sharing relevant or as-yet-unpublished information with Ofcom related to emerging online safety issues. TikTok provided full and complete responses to the formal information request.

## Governance and risk management

### Decision-making structure

**TikTok employees from a wide range of departments work together to assess the appropriateness and implementation of the protection measures they employ**

- 6.5 TikTok's Trust & Safety team, the European Safety Public Policy team and the Product team working cross-functionally, supported by the EMEA Product and Regulatory Legal team, are responsible for decisions regarding Schedule 15a measures. These teams assess the "appropriateness" of the platform's Schedule 15a measures and "oversee and monitor their implementation."
- 6.6 TikTok's Trust & Safety work regarding users in the UK is led from the regional EMEA Trust & Safety Hub in Dublin, Ireland. This team includes TikTok's Global Head of Trust and Safety, who is based in Dublin.
- 6.7 TikTok listed the following job titles as being "involved in decision-making":
- a) Global Head of Trust & Safety
  - b) Global Head of Trust & Safety Product Policy
  - c) Head of Product & Process (Europe)
  - d) Head of Global Safety Support
  - e) Global Head of Law Enforcement Response Team.
- 6.8 TikTok recently established an Online Safety Oversight Committee (OSOC) to "provide executive oversight of content and safety compliance" specifically within the UK and EU region. The OSOC includes the Global Head of Trust & Safety, the General Manager for the UK and EU, and the Director of Legal and Corporate affairs. TikTok modelled the OSOC on

---

<sup>29</sup> Quote taken from TikTok's response to the formal information request issued in the summer of 2022. Subsequent quotes in the TikTok section are also from TikTok's response to this information request.

the regional governance framework it has in place for data protection matters. Ofcom considers this a relevant point because we would like to see platforms taking online safety into account in the way they currently do with data protection – through proactive risk assessments at the point at which they are developing new or updated products.

- 6.9 TikTok also established a “cross-functional risk assessment review group” responsible for “validating” internally documented risk assessments. This group reports to the OSOC.

## Risk assessment

### TikTok conducts risk assessments relating to the safety of its users

- 6.10 When TikTok launches new features on the platform or makes changes to existing features, it requires cross functional input from its Trust & Safety teams (including the Minor Safety teams), Legal teams, and the Safety Public Policy team.
- 6.11 Some of the “key factors” TikTok say are relevant to its risk assessment processes include:
- a) Legal changes and regulatory guidance;
  - b) Internal and external research;
  - c) Internal data on risks regarding restricted material;
  - d) Feedback from external partners and experts;
  - e) Classification rules and standards in specific countries.
- 6.12 Individuals from across TikTok’s Trust & Safety teams, Safety Public Policy team, and Legal teams carried out a risk assessment prior to the VSP regime coming into effect to determine whether the measures already in place on its platform would meet the requirements of the VSP legislation. It simultaneously carried out an assessment to determine whether it would be appropriate to implement additional measures listed in Schedule 15a.
- 6.13 TikTok “continuously” assesses the effectiveness of its Schedule 15a measures to ensure that they are “appropriate and remain up to date and fit for purpose”. TikTok stated that “this process has involved the implementation of both minor adjustments and also more substantial changes on an ongoing basis.”
- 6.14 TikTok’s risk assessments include a dedicated section assessing the platform’s approach to protection measures. Under the VSP Framework, VSP providers are required to determine whether it is appropriate to implement a particular measure to protect users from harmful material according to whether it is practicable and proportionate to do so. TikTok’s risk assessments do this by taking into account aspects such as the platform’s size, the volume of content on its platform, the nature of the material available to users, and the platform’s functionality.
- 6.15 “Relevant decision makers” within TikTok have also considered various characteristics of its community of users when determining which of the Schedule 15a measures are

appropriate to use on the platform. These include age, physical characteristics and wellbeing, and protected characteristics.

## Assessing effectiveness

### TikTok told us it uses a range of metrics to assess the effectiveness of Schedule 15A measures



6.16 TikTok records the number of videos and accounts removed from the platform, a breakdown of the reasons by reference to the various policy categories under the Community Guidelines, and the number of accounts suspected to be underage removed from the platform per quarter. This information is captured in TikTok's [Community Guidelines Enforcement Reports](#).

### TikTok also told us it collects a variety of internal metrics to assess the effectiveness of its measures

- 6.17 TikTok measures the number of visits to the Community Guidelines webpage and Transparency Centre webpage;
- a) The volume of user reports from UK users by policy category/appeal rate/volume and rate of user reports leading to content removal;
  - b) The number of visits to TikTok's UK AVMS information and complaints webpage;
  - c) The volume of complaints through the AVMS information page, including a breakdown of complaints by categories and level of user follow-ups after initial feedback; and
  - d) Engagement with its media literacy tools.
- 6.18 TikTok conducts its own research on protection measures and how it can improve the effectiveness of these measures. In 2021 the platform commissioned a report exploring things that can be done to prevent harm to children from dangerous online challenges and hoaxes.<sup>30</sup>






## User journey

Figure 6.2: TikTok User Journey

<p><b>User opens website or app</b></p>		<p>Users can view a limited amount of content without an account (and only on the web). TikTok applies a range of measures that limit the potentially sensitive content users can see without logging in</p>
<p><b>User sign-up and log-in</b></p>		<p>Users are asked to enter their date of birth when they are creating an account. TikTok detects and removes suspected underage accounts through a combination of automated detection measures.</p>

<sup>30</sup> [Exploring effective prevention education responses to dangerous online challenges \(internetmatters.org\)](#) (November 2021)



<p><b>VSP recommends or user searches for content</b></p>		<p>The Family Pairing parental controls can place limits on what a child can search for; see and how they might interact with other users.</p>
<p><b>User watches and engages with content</b></p>		<p>Terms of service restrict what users can see, but TikTok does not currently give users that upload content tools to give that content a rating (i.e. whether content might contain material unsuitable for u18s).</p>
<p><b>User encounters harmful material and flags or reports</b></p>		<p>TikTok places emphasis on its own proactive moderation to review and detect uploaded content in breach of its policies Users also have tools to report and flag, but users cannot apply ratings to potentially harmful material. Trusted third-party flaggers monitor and flag harmful content.</p>
<p><b>User informed of outcomes</b></p>		<p>Within the TikTok app, users get updates on the outcome of TikTok's review of content that users have reported, including whether the content was in violation of the platform's terms and conditions.</p>
<p><b>User access to other tools</b> <i>(e.g. Media Literacy)</i></p>		<p>TikTok offers a range of tools and information to improve users' media literacy, including: TikTok-funded videos by creators; on-site safety centre and portals; offline initiatives with other organisations.</p>

## Signing on

### Users can view a limited amount of content without an account

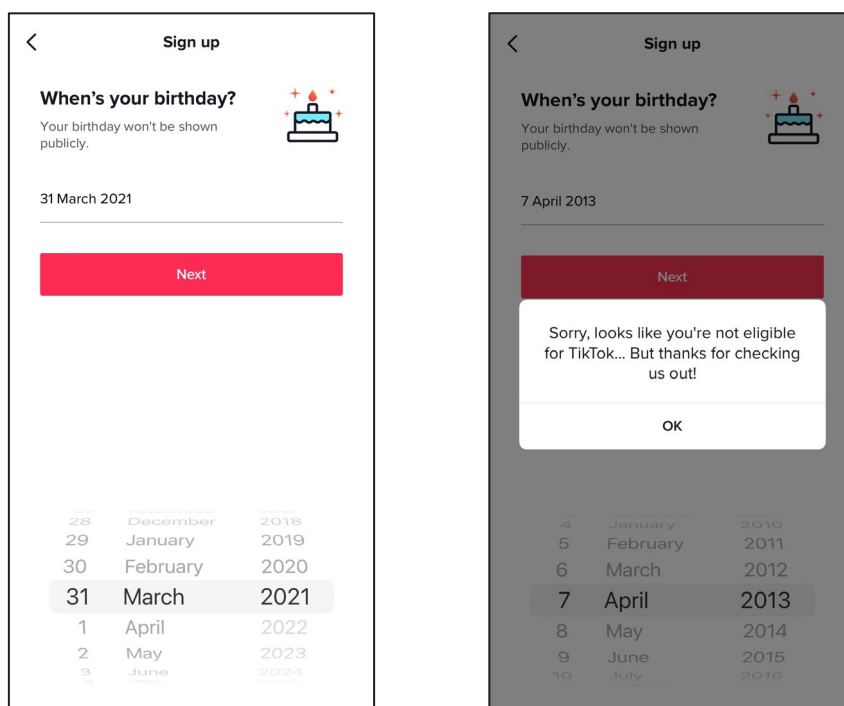
- 6.19 TikTok has said that, when using the mobile app, UK-based users are required to register an account and log-in to view video content on the platform. For visitors to the TikTok website who have not logged in, the videos that are available to those without an account are subject to a number of restrictions:
- They have been through several rounds of human moderation;
  - Any content that has a warning notice will not be eligible to appear; and
  - Videos with captions or hashtags that hit "sensitive word lists" will not be eligible to appear.
- 6.20 Users must create an account and sign in to upload any videos to TikTok.

### Users are asked their age when they are creating an account

- 6.21 TikTok asks its users to self-declare their age when registering for an account. The platform uses a neutral age gate to avoid nudging users to select the minimum required age.
- 6.22 [Yonder research](#) commissioned by Ofcom in 2022 showed that one third of under-18s lie about their age when creating a social media account. The research also showed that some younger respondents, aged 8-12, have done this when registering for accounts on platforms with an age requirement of 13+. TikTok was the platform where children 8-12

were most likely to say they had lied about their age (65%).<sup>31</sup> Age assurance is one of **Our strategic priorities in Year 2**, and we will continue to engage with TikTok on this important topic.

**Figure 6.3: The alert shown to users under the age of 13 if they attempt to sign-up for a TikTok account**



Source: Screenshot provided by TikTok in response to Ofcom's information request, June 2022

- 6.23 TikTok detects suspected underage accounts through a combination of user reporting, third-party reporting, and automated detection measures based on textual indicators. Suspected underage accounts are subject to moderator review and are removed if the user is considered to be under the age of 13.
- 6.24 TikTok is the only VSP that currently reports publicly on the number of suspected underage accounts it removes. According to its own Transparency Reports, TikTok removed nearly 60 million accounts globally on the basis of suspected underage use from April 2021 to April 2022.
- 6.25 Suspended users are temporarily blocked from creating a new account from any device and can appeal this ban by confirming their age through submission of either ID, a selfie with a trusted adult (under 18), or temporary credit card authentication (over 18).

<sup>31</sup> [Children's Online User Ages Quantitative Research Study \(ofcom.org.uk\)](https://www.ofcom.gov.uk/consult/condocs/childrens/childrens-ages-quantitative-research-study/)

## Agreeing to terms and conditions

### TikTok's terms and conditions set the rules for the content that users are allowed to upload

- 6.26 TikTok's Terms of Service and Community Guidelines outline what is not allowed on the platform. We will refer to these collectively as 'terms and conditions'.
- 6.27 Users are presented with the Terms of Service during sign up. This includes a section on 'Using the Platform' and 'What you can't do on the Platform'. In this, TikTok states 'in any event, you must not use the Platform to do anything illegal'.
- 6.28 The Terms of Service and Community Guidelines are linked at the bottom of the 'sign in' or 'register' page and users must agree to the terms before they can upload content to the platform. Users are also reminded of the Terms of Service through an in-app notification if they violate TikTok's Community Guidelines.
- 6.29 The Terms of Service includes a link to the TikTok [Community Guidelines](#). These contain policies around categories of content which are prohibited on the site. The policies detail what is not allowed and tells users 'do not post, upload, stream or share' a list of categories of harmful content.
- 6.30 TikTok's Community Guidelines expressly prohibit users posting, live streaming or distributing a comprehensive list of harmful content including content that is not suitable for under 18s or would cause 'physical, mental or moral detriment to minors' and details behaviour which is not allowed on the site.

### TikTok's terms and conditions prohibit terrorist and hateful content

- 6.31 TikTok's Terms of Service states that users cannot 'do anything illegal' or behave in a way that 'negatively impact the enjoyment of other users'. This is expanded upon in TikTok's Community Guidelines that reiterates its 'firm stance against enabling violence on or off TikTok' where it does 'not allow people to use [its] platform to threaten or incite violence, or to promote violent extremist organizations, individuals or acts'.
- 6.32 Where TikTok believes there could be a threat to public safety, or an account could be used to promote or glorify off-platform violence, it will ban the account. It may also consider off-platform behaviour to identify violent extremist organisations and individuals (TikTok was one of three platforms that told us it does this).
- 6.33 In September 2022 TikTok became a member of Tech Against Terrorism.<sup>32</sup> This membership commits it to exploring new technical solutions and working with civil society on combating violent extremism.
- 6.34 The Community Guidelines outline and provide examples of content that is unacceptable for users to post, upload, stream, and share; this includes but is not limited to statements

---

<sup>32</sup> [Partnering to prevent violent extremism \(tiktok.com\)](#)

to inflict harm, encouraging others to commit violence, and calls to bring weapons to a location with the intent to intimidate or threaten an individual or group with violence.

- 6.35 The use of slurs is also highlighted within TikTok's Community Guidelines as 'derogatory terms... intended to disparage groups or individuals' and states that it removes all slurs, unless the terms are "reappropriated, use self-referentially (i.e., by members of the protected group), or used in way that does not disparage (e.g., educational context)".<sup>33</sup> This is a unique approach and has not been highlighted by any other platform.
- 6.36 The Community Guidelines also expand on the consequences of sharing this content on the platform, stating that users breaching these rules can have their account suspended or terminated. Accounts or users that engage in multiple hate speech violations are banned.

### **TikTok's terms of service are dense, but the Community Guidelines are user friendly**

- 6.37 TikTok's terms are very dense, but its Community Guidelines are more user-friendly, with a contents list with hyperlinks, large headings and sub-headings, and bullet-point lists.

### **When necessary, TikTok update their Community Guidelines**

- 6.38 TikTok told us it aims to maintain consistency on its platform by not updating its terms too frequently, but also wishes to be flexible enough to respond to emerging developments and risks. When TikTok updates its terms and conditions, it follows a review, consultation, and approval process.
- 6.39 An internal team continuously reviews the Community Guidelines to assess whether any changes are needed. This work is led by the Product Issue Policy team that sit within the Trust and Safety team. This team is made up of specialists in particular categories of harm who conduct literature reviews, engage with internal colleagues to obtain insights, collaborate and consult with relevant internal and external stakeholders, and ensure the terms are up to date.
- 6.40 Consultation takes place through outreach teams such as TikTok's Outreach Partnership Management (OPM) team. This team builds relationships with external experts and groups, such as Stonewall UK, to support policy improvement. TikTok also consults internally on changes across teams such as the global security team, legal teams, public policy and communications teams.
- 6.41 For example, the changes made to the Hateful Behaviour Policy in February 2022 were informed by consultations with civil society groups and enhanced the clarity on hateful ideologies. TikTok made changes to its Violent Extremism policy to "focus more holistically on the issue of violent extremism", after work with Tech Against Terrorism.<sup>34</sup>
- 6.42 The conclusions of these processes must then be approved by the Trust and Safety senior leadership team.

---

<sup>33</sup> [Community Guidelines: Hateful Behaviour \(tiktok.com\)](#), page 9.

<sup>34</sup> TikTok Response to Ofcom Information Request – Cover Letter and Main Response – 25 July.

### TikTok has notified users of updates to their terms of service via an in-app notification

- 6.43 TikTok sent users an in-app pop up in May 2022 to notify users of a change of its Terms of Service that would come into effect the following month. Users were given three options ('Agree', 'Remind me later', and 'Learn more') and could not continue onto the app without selecting one of these options. The 'Agree' option was in bold, inviting users to select this. Only by selecting 'Learn more' was a user actually presented with the details of the changes to the terms.
- 6.44 TikTok also publicises changes through 'newsroom posts' on their website that explain the rationale behind the updates.

## Uploading and watching content

### TikTok categorises its users by age, and this determines the features they have access to and the type of content they can see

- 6.45 The VSP Regime requires TikTok to protect under-18s from restricted material<sup>35</sup>. TikTok groups users under the age of 18 into two categories: users aged 13-15 and users aged 16-17. TikTok was one of two platforms to tell us it does this.
- 6.46 TikTok told us that its general approach is to implement various measures for under-18s, and then to further consider the two different age groups (13-15 and 16-17) for certain measures. TikTok will then shape the user experience according to the age category that the user belongs to. This results in younger users having increased privacy and less access to potentially risky features.
- 6.47 The age category will also influence the nature of the content eligible for a user's For You feed. Some types of content are considered 'Ineligible for the For You feed', and therefore not in the For You feed of under-18s. It is not clear from TikTok's response whether under-18s can still view this content if they search for it.
- 6.48 The types of content that are not prohibited from the platform entirely but may be unsuitable for under-18s include content featuring dangerous stunts and sports, regulated goods such as alcohol or tobacco, violent and graphic content, and overtly sexualised content, and misleading content. Any content uploaded by users under the age of 16 is also not eligible for recommendation.
- 6.49 However, this grouping is based on the date of birth entered by users when they register, which could be falsified. Yonder research commissioned by Ofcom showed that a third of social media users aged 8-17 with their own profile have a user age of at least 18. 32% of children aged 13-15 with their own profile said they used an older date of birth when creating their profile on TikTok.<sup>36</sup>

---

<sup>35</sup> Restricted material refers to videos which have or would be likely to have an R18 certificate, or which have been or would likely be refused a certificate.<sup>11</sup> It also means other material that might impair the physical, mental or moral development of persons under the age of 18.

<sup>36</sup> [Children's Online User Ages Quantitative Research Study \(ofcom.org.uk\)](https://www.ofcom.gov.uk/consult/condocs/childrensresearch/childrensresearch.pdf)

- 6.50 TikTok relies on its own moderation systems and processes to identify content that may be unsuitable for under-18s, rather than allowing uploaders or viewers to apply these ratings. For example, TikTok uses a profanity indicator to filter out mature content that features on under-18s’ For You Feed. It also uses a variety of other indicators to classify content into three age categories, including content appropriate for users aged 13-15, 16-17, and users aged 18+.
- 6.51 In July 2022, TikTok began to implement the first phase of its content classification system, which sorts videos into content levels based on thematic comfort levels. The initial rollout encompasses content that does not violate Community Guidelines but is considered to be suitable for users aged 18+. Such content is not recommended to under-18s and they will not be able to search or view the content, even if it is shared with them. If under-18s attempt to view this content, they will be informed that the content is unavailable as it is age protected.

**The Family Pairing Feature provides the option to apply parental controls**

- 6.52 TikTok launched its Family Pairing feature in April 2020. Once enabled, this feature allows parents and guardians to link their account with their child and take certain decisions regarding their child's use of the platform, though TikTok does not verify that the ‘parent’ does in fact have parental responsibility for the child. Under-18 users can disable Family Pairing at any time, at which point the parent would be notified of this.
- 6.53 TikTok did not provide evidence to support the effectiveness of its parental control systems. Yonder research commissioned by Ofcom showed that 41% of parents have never used parental controls on TikTok, despite being aware of them.<sup>37</sup>

**Figure 6.4: TikTok Family Pairing features**

Feature	Description
Screen Time Management	Enables a guardian to set a limit (between 40 and 120 minutes) on how long their teen is able to spend on the platform each day. If a teen wishes to use the platform outside of the allotted time, the guardian is required to enter a passcode.
Restricted Mode	Enables a guardian to turn on Restricted Mode. Restricted Mode limits the visibility of any content that contains a warning notice in their child’s For You Feed, which means the child can only view videos that have been subject to several rounds of content moderation.
Search	Enables a guardian to turn on or off a teen’s ability to use the search function, where they would otherwise be able to search for content, users, hashtags, and sounds.

<sup>37</sup> [VSP Parental Guidance Research 2022 \(ofcom.org.uk\)](https://www.ofcom.org.uk/research-and-data/parental-guidance-research-2022)

<b>Direct messages</b>	Users become eligible for using the direct messages feature at 16. Through Family Pairing, a guardian can turn on or off the ability of a teen over-16 to direct message friends.
<b>Privacy settings</b>	Enables a guardian to choose the following privacy related settings: <ul style="list-style-type: none"> <li>• To choose if their teen's account is public or private</li> <li>• To choose if other users can see their teen's 'liked videos'</li> <li>• To choose if their teen's account could be suggested to other users</li> <li>• To limit who is allowed to comment on their teen's videos.</li> </ul>
<b>Dashboard</b>	Where enabled, a teen is given access to a dashboard where they can see the choices made by their guardian. They cannot change the individual settings selected by their guardian but can disable Family Pairing at any time.

### TikTok cited several other sources it has considered when seeking to design an age-appropriate experience

- 6.54 TikTok did not tell us the types of profanity that would be permitted under each age classification. However, it told us it considered the most recent Ofcom commissioned [report into offensive language on TV and Radio](#) when designing the classification system.
- 6.55 TikTok said it considers the [ICO's guidance on Age and Developmental Stages](#) and has produced its own internal guidance on age-appropriate design. Its internal guidance is intended to be followed by Product Feature teams to support the integration of child safety into the design process.

### TikTok uses video content, off-platform resources, and offline initiatives to improve the media literacy skills of its users

- 6.56 TikTok told us it uses videos to improve its users' understanding of potential harms that may be encountered and the safety measures available to protect users from those harms. For example, the 'TikTok Tips' page promotes a range of short videos on topics such as critically engaging with content by questioning the source, graphics, and users' own bias, the difference between fact and opinion, and encouraging users to reflect on when it is appropriate to share or report content.
- 6.57 In addition to its own video campaigns, TikTok has collaborated charities and prominent content creators to produce and promote videos on a broad range of online safety issues. These resources aim to raise user awareness about reporting and comment filtering tools, inform users what happens once content is reported and how the feed works, inform users around setting privacy levels and how to create a more secure safe password, and help users recognise when they are spending too much time online.
- 6.58 TikTok reported on user engagement with three of the media literacy campaigns run in 2021. The data provided indicate that:

- a) The #SaferTogether video campaign on online safety received views from approximately 9.4m viewers in the UK, with a further 1.07 million user engagements. It is reported to have resulted in an 8% rise in user familiarity with TikTok Safety measures and facilitated increased traffic directed towards safety measures such as the Digital Wellbeing section (19% increase), Screen Time Management (53% increase), Comment Filters (5% increase), and switching on the Private Account setting (5% increase);
  - b) The #FactCheckYourFeed campaign received 15 million viewers, including over 550 thousand visits to a dedicated #FactCheckYourFeed media literacy hub where users were reported to have spent four times as much time engaging;
  - c) The #SwipeOutHate campaign received 9.4 million views and over 200 thousand engagements in the UK on a single day.
- 6.59 In addition to its video campaigns, TikTok reported several off-platform media literacy tools hosted on its safety center, help center, youth portal, and shared in newsroom posts. These resources range from guides for parents/caregivers on privacy controls and information on how to report harmful material to safety information for under-eighteens on interacting with users and managing suspicious follow requests.
- 6.60 The platform is also involved with offline media initiatives which entail collaboration, information-sharing and discussing best practice with external stakeholders and partners.
- a) TikTok partnered with Internet Matters to develop Parental Controls and Privacy Setting Guides, running focus groups with parents/caregivers and teenagers to understand how best to address conversations about online safety, and developing a guide to TikTok for teachers and educators.
  - b) It also told us it partnered with the South West Grid for Learning (SWGfL) to develop a TikTok Guide with Media Smart to inform under-18s, their parents, guardians, and educators on how advertising works on TikTok.
  - c) TikTok has joined DCMS's Media Literacy Steering Group and is participating in Ofcom's Making Sense of Media (MSOM) Network to share insights and engage on best practices for media literacy.

### **TikTok takes steps to promote users' awareness of media literacy tools and information**

- 6.61 TikTok encourages users to report or flag content on the platform, for instance by providing information in the Help Center, Safety Center, and various video series with instructions on how to report harmful content. It also uses real-time in-app feedback after a user reports content.
- 6.62 It also promotes content to raise under-18 users' awareness of media literacy tools and information on the platform. These users are presented with a Privacy Highlights for Teens video series as part of the registration process which are displayed when under-18 users subsequently open the mobile application.



- 6.63 TikTok informs users of changes to their terms and conditions. Users receive in-app notifications to make them aware of these changes and supplements the updates with video and blog posts explaining the purposes of those changes.

## Detecting harmful content

### TikTok relies mostly on proactive content moderation, rather than user reporting

- 6.64 Rather than rely on user reports, TikTok aims to proactively detect and remove harmful material before it is reported to them by users or third parties.
- 6.65 Only 4.6% of content found to be violative is flagged via a user report. This aligns with separate TikTok submissions to other authorities where it claimed 95.1% of such content was removed before a user reported it.<sup>38</sup> TikTok provided global data on video removals and user reporting rather than UK specific data.
- 6.66 When a user uploads content to TikTok it goes through a technology-based automated review process designed to flag content that may violate its Community Guidelines. While a piece of content is being reviewed it is visible only to the uploading user. This review process involves several tools including:
- a) use of computer vision models designed to detect and classify objects within video content
  - b) comparison of videos to match hashed content previously found to violate the Community Guidelines
  - c) review of text and audio content to detect potentially harmful material.
- 6.67 If these reviews flag that content may be harmful, this is then reviewed by a person in TikTok's content moderation team or it is removed from the platform by automated systems.

### Users can report harmful content for review by human moderators

- 6.68 Users do not need to be logged in to flag content. There are two options to report in-app, either after clicking share and selecting the flag icon, or selecting the 'report' option after long holding on a video. Users can report content on the website as well.
- 6.69 Users are able to report against the following categories: minor safety dangerous acts or challenges; suicide, self-harm, or disordered eating; adult nudity or sexual behaviour; bullying and harassment; hateful behaviour; Violent extremism; spam and fake engagement; hateful misinformation; illegal activities and regulated goods; violent and graphic content; intellectual property infringement; and if unable to categorise they can report it as 'Other'.

---

<sup>38</sup>[TikTok Written Submission to the Consultation on the Regulation on the Digital Content Platforms \(gov.uk\)](#), April 2022

- 6.70 TikTok claim in their 'Learn How Reporting Works'<sup>39</sup> video that every report that is sent to TikTok is 'checked'. However, it is not clear from the information shared that this is checked by a human moderator or the average timescales for this process.

#### **The moderation process differs depending on the type of harmful content reported**

- 6.71 If suspected CSAM is detected by TikTok, this is escalated to its Child Safety Team (CST), which forms part of the Law Enforcement Response Team. The CST team is trained to handle such content and determine the appropriate course of action to be taken, such as reporting to relevant authorities and banning of associated accounts.”
- 6.72 If a violent threat is credible (with a detailed, achievable plan), TikTok will report it to its Emergency Response Team.

#### **TikTok makes users aware of reporting tools in a variety of ways**

- 6.73 TikTok presented a range of options and avenues for users to understand and be made aware of user reporting:
- a) Instructions and guidance in the Terms of Service, Community Guidelines, and Help Centre;
  - b) Via their Safety Centre and Online Youth Portal.
- 6.74 TikTok also undertook a proactive awareness raising campaign with journalist Sophia Smith Galer and podcaster Boni Odoemene. This included two videos under the hashtag #SaferTogether designed to educate people about reporting harmful content.

#### **Trusted flaggers, reviews of popular content and targeted sweeps also form part of the manual reviewing process**

- 6.75 In addition to user-reporting TikTok has a Community Partner Channel in which non-governmental organisations (NGOs) perform the role of trusted flaggers (entities who can submit reports of suspected harmful content via dedicated email channel monitored by TikTok's Risk and Response team).
- 6.76 TikTok's content moderators also manually review content when it reaches a certain level of popularity in terms of video views.
- 6.77 TikTok's Trust and Safety team will also regularly undertake targeted searches of the platform for specific risks.

---

<sup>39</sup>[Let's be #safertogether \(tiktok.com\)](#), November 2021

## Enforcing against harmful content

### Sanctions are in place to deal with content that breaches the terms and conditions

- 6.78 In its terms and conditions, TikTok sets out that it can remove or restrict content which breaches its rules and can also temporarily or permanently suspend a user where it has, or are about to, seriously break their terms and conditions.
- 6.79 Users who breach TikTok's Terms of Service or Community Guidelines could be sanctioned in a number of ways depending on the nature of the breach as well as other factors:
- a) Content removal;
  - b) Account ban;
  - c) Do not proactively promote;
  - d) Feature ban (restrictions on certain features);
  - e) Device ban.
- 6.80 TikTok also can geo-block certain content, so it is not available in certain jurisdictions. According to the training materials shared, potentially harmful content is blocked from regions where it could trigger a potentially violent response.

### Moderation teams apply sanctions to hate and terror content in a variety of ways

- 6.81 In relation to hate and terror content, the sanctions applied are generally the same. However, for content that violates violent extremism and hateful behaviour policy categories in the [Community Guidelines](#), such content is generally removed from the platform. "Hard to find" (i.e. do not recommend) and "Do not proactively promote" (i.e. not for feed) are only used for two limited exceptions within the hateful behaviour category. These exceptions relate to less severe content.
- 6.82 Where content contains hateful or illegal content, TikTok has moderation processes which will either make the content harder to find or remove it completely, as appropriate. Depending on the seriousness of the suspected breach, TikTok can suspend access to some platform features and remove or restrict content while an investigation is underway. In addition, users or third parties may report users to the platform for uploading potentially hateful or illegal content.

### Moderators are trained to ensure consistency when apply sanctions

- 6.83 The Trust and Safety team are trained on the relevant systems, sanctions and the policy categories outlined in the Community Guidelines. The aim is to ensure that the team understands the core issues, terms and can differentiate between the subcategories of harm, to properly identify and appropriately apply sanctions. The documents also detail clear examples and exceptions to moderation such as satirical humor or criticism of a hateful ideology.

- 6.84 The internal training documents that TikTok shared with us separates the guidance on video policies for hateful behaviour and violent extremism:
- a) Hateful behaviour is then separated between hate; hate speech, slurs in degrading contexts, promotion of hateful ideology and inciting religious conflict; and stereotypes;
  - b) Violent extremism focuses on terrorist and terrorist organisations on international and regional levels, TikTok will report users to local legal authorities according to the laws and regulations.
- 6.85 Stereotyping is only considered harmful behaviour when geared towards members of a protected group and in a negative context, this includes imitating the accent of a minority group, or harmful generalisations based on nationality. If content is negatively stereotyping without attacking a group with protected attributes, TikTok moderators will make this content harder to find.
- 6.86 Moderated content is then subject to Quality Assurance (QA) through regular sampling.
- 6.87 TikTok told us it continues to develop specialist moderation queues for issues, such as violent extremism, and these are also used to help with the development of focused expertise, such as terrorism. TikTok also uses automated systems to detect harmful content which are subject to extensive human oversight, testing and adjustments, including manual reviews from moderated samples by their QA team.
- 6.88 TikTok also monitors user appeals and complaint rates with processes in place to investigate significant fluctuations.

#### **Users can be informed when sanctions are applied, but not always**

- 6.89 TikTok provides in-app notifications to users in real time to inform them of content removal, account bans and feature bans - this generally happens within several seconds or up to a few minutes at most. If the user is offline at the time of the ban, then they are informed of the sanction when they next attempt to open the app.
- 6.90 However, there are specific instances that users are not directly informed, for example, if content has been made "hard to find", TikTok have told us that users are made aware of this feature and sanction in the general terms of the Community Guidelines. Additionally, for device bans, users are not informed in specific instances but will have been aware of specific account bans that result in a device ban.

#### **Users can appeal against sanctions**

- 6.91 There are varying approaches to appeal; for content removal and account bans, users can simply click the "submit appeal" button, the process beyond this is not specified. However, there are some sanctions that users are not able to appeal:
- a) For content that is made "hard to find" or where content is "not actively promoted", users are unable to appeal. This is considered as proportionate relative to the limited nature of the sanction;

- b) Users cannot appeal “feature bans” and “device bans” (which arise as a result of an accumulation of content-level violations), but they can appeal the underlying content-level violations and/or account bans.

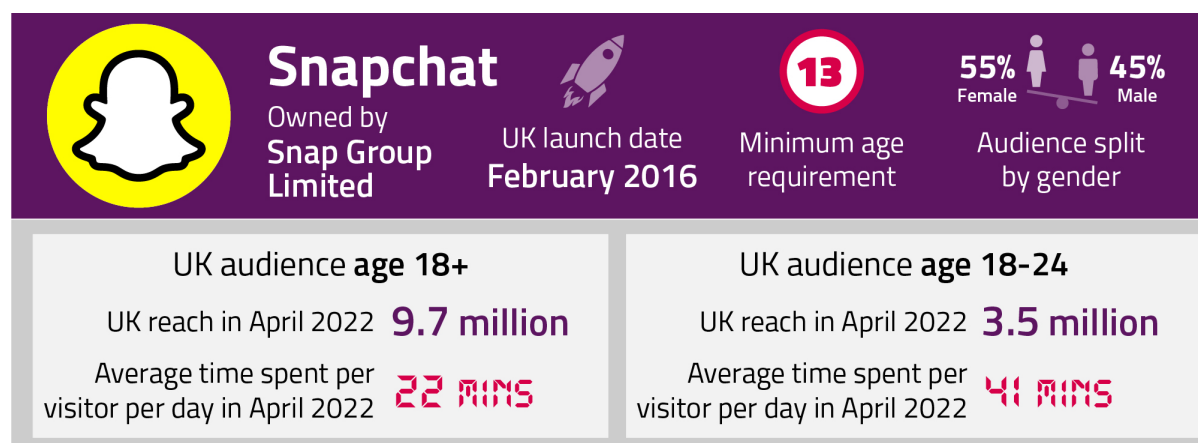
6.92 In all appropriate cases, users will be notified and may be given time to download their data.

## 7. Snapchat

### Introduction

Snapchat is most well known as a messaging app that lets users exchange pictures and videos which typically are only available for a brief period. Since September 2022 it is also now available to access through a website. Only two areas of Snapchat are currently notified to Ofcom as being in scope of the VSP Regime: **Spotlight** and **Discover**. Spotlight, a personalised feed for users showing popular user-generated content, goes through human moderation before content can reach more than 25 viewers. Users can add hashtags to their content to increase its exposure on the platform.<sup>40</sup> Discover is a feed featuring curated content from verified media outlets and vetted content creators.<sup>41</sup> Discover also recommends content to users in a personalised feed, and users can subscribe to specific channels. Snap generates revenue through advertising; users are shown video advertisements alongside videos made by content creators.

Figure 7.1: Key information on Snapchat



Source: © Ipsos, Ipsos iris Online Audience Measurement Service, 1 - 30 April 2022, age 18+ for reach and time spent data, age 15+ for audience share by gender data, UK. Note: UK launch date refers to date that Snapchat services under scope of VSP regulation (Spotlight and Discover) launched.

### Key findings

- 7.1 **Snap moderates popular user-generated content before it is shared widely to users on Discover or Spotlight**, making it less likely that harmful content will be found here. We do not know if videos shared to other public pages, such as Maps, are subject to the same moderation process.<sup>42</sup>
- 7.2 **From Snap's responses to the information request, it is not clear that it has a formalized online safety governance structure in place** to ensure it is implementing appropriate

<sup>40</sup>[Spotlight 101 \(snap.com\)](https://www.snap.com/spotlight-101)

<sup>41</sup>[What is Snapchat Discover: Fresh Content at Your Fingertips \(g2.com\)](https://www.g2.com/articles/what-is-snapchat-discover-fresh-content-at-your-fingertips)

<sup>42</sup> The Maps part of the platform, which allows users to click on a map of the world and watch videos uploaded by users located in that area, is not currently regulated by Ofcom.

measures to protect users from harmful content on Spotlight and Discover. However, this is not a formal requirement on platforms under VSP legislation.

- 7.3 **Snap is taking some innovative approaches to providing media literacy tools.** It has integrated a 'Here for You' tool into its search tools that shows users resources from localized partners if users search for topics related to mental health, anxiety, depression, stress, suicidal thoughts, grief, and bullying.

## Engagement with Ofcom

- 7.4 Snapchat is the only platform we regulate in which only part of the service is notified. While our engagement with Snap focused on the two notified areas, some of the measures taken may be implemented beyond the notified areas. We gathered information about all these measures using our formal powers, and we will look to deepen our engagement with Snap in the coming year to better understand how it is working to protect its users in line with the VSP legislation.

## Governance and risk management

### The decision-making structure that Snap follows to implement appropriate safety measures is not clear

- 7.5 Ofcom asked platforms to provide a description of the decision-making structure in place for considering whether it is practicable and proportionate to implement certain safety measures. We hoped to gain a better understanding of the process through which these platforms assess the nature and scale of risks on their platforms against the cost of mitigations and how they make sure these processes have appropriate checks and balances built in throughout.
- 7.6 Snap's response did not clearly show a well-defined and identifiable process for doing this. It also did not show that assessments of the practicability and proportionality of measures were being carried out.

### Snap told us it collects several metrics to assess the effectiveness of their protection measures

- 7.7 In assessing the effectiveness of its protection measures, Snap told us it collects a number of metrics, such as the timeliness and accuracy of moderation decisions and Snap's effectiveness in enforcing against its Transparency Reporting categories. It also referred us to its transparency report site for the UK. This report indicates that in the UK from July 2021 to December 2021 there were over 750,000 user reports made to Snap about either a piece of content or an account. To see more of Snap's transparency reporting statistics please visit [Snap's Transparency Report site](#) for the UK.

7.8 The metrics in the Transparency report site also include the breakdown of the total number of pieces of content reported by reason, such as sexually explicit content, drugs, harassment and bullying, or spam.

## User journey

Figure 7.2: Snapchat user journey

<p><b>User opens website or app</b></p>		<p>Most content is only accessible with an account on the Snapchat app, but its Discover and Spotlight feeds are viewable on its website without an account. In this case account-holder protection measures will not work.</p>
<p><b>User sign-up and log-in</b></p>		<p>Users must provide a birthdate when signing up (anyone putting a birthdate under 13 will not be allowed an account). Users must agree to Snapchat's terms of service when signing up for an account.</p>
<p><b>VSP recommends or user searches for content</b></p>		<p>Content is categorised based on suitability for under or over-18s, under-18s are restricted from viewing mature content where it has been labelled as such by publisher partners. Snapchat also refers to measures to monitor a user's age (i.e detecting underage users), which feeds into content recommendations and account monitoring.</p>
<p><b>User watches and engages with content</b></p>		<p>Videos on its Discover and Spotlight feeds are pre-moderated or rely on content from vetted third-party creators. Snapchat's Family Centre gives parents some oversight of who their child is interacting with. Snap is yet to implement specific parental controls over the content their children can see.</p>
<p><b>User encounters harmful material and flags or reports</b></p>		<p>Snapchat provides a reporting function to users. It also operates a Trusted Flagger programme, using third parties to identify content that violates policies.</p>
<p><b>User informed of outcomes</b></p>		<p>Snapchat does not provide feedback to users on the outcome of its review of reported content.</p>
<p><b>User access to other tools (e.g. Media Literacy)</b></p>		<p>Snapchat provides a range of media literacy tools and information for users both on and off platform to raise awareness, including informative videos from content creators, notifications and prompts to users. Snap collaborates with external partners to develop these tools and user-friendly guides for parents.</p>



## Signing on

### Users without an account are unable to upload content, but can view content on the Spotlight and Discover feeds via Snap's website

- 7.9 Users need to create an account on Snapchat to access many of its features. Users must create an account and log in to upload video content. User-generated images and videos can either be shared privately with other users as a "Snap" or posted as a "Story", which can be shared either with groups of friends or publicly. Snap content is ephemeral, meaning it disappears after a certain amount of time.
- 7.10 Some features are available without an account via Snap's website (i.e. not on the mobile app), including Spotlight and Discover.

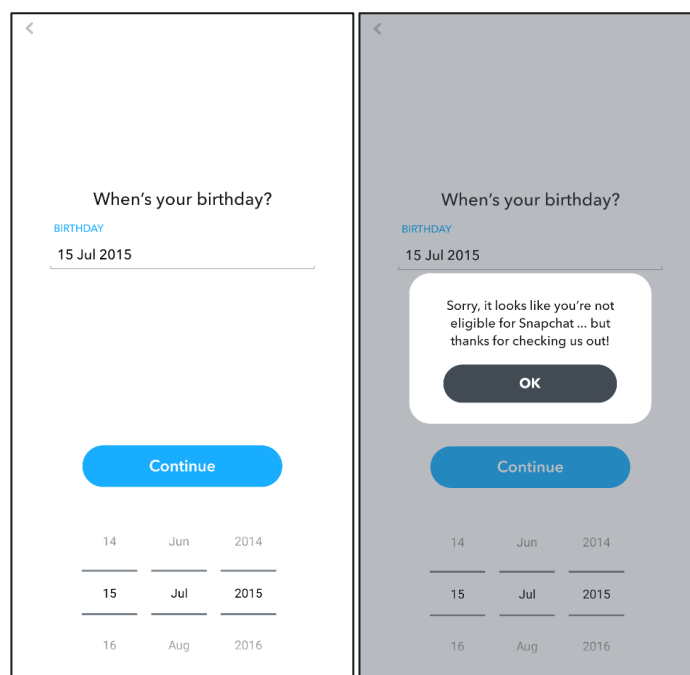
### Users are asked to enter their date of birth when signing up for an account on Snapchat

- 7.11 If the date entered indicates that they are below the age of 13, the user is not permitted to create an account. In this case, the user is not told that they are underage, just that they are not eligible. If users under the age of 18 do create an account, they are unable to change their date of birth. Yonder research commissioned by Ofcom showed that 59% of children aged 8-12 that used social media and had their own profile said they had used a false date of birth to appear older when registering for an account on Snapchat.<sup>43</sup> Age assurance is one of **Our strategic priorities in Year 2** and we will continue to engage with Snap on this.
- 7.12 Snap uses tools to estimate the age of users already on the platform to identify where users may have misrepresented their age at sign-up. This is based on a number of factors and is used to stop users suspected to be under 18 from seeing content that is inappropriate for them, such as advertisements of regulated goods. However, Snap did not disclose whether this information is used to validate information about the age of its users or whether it would prompt further investigation of the user.
- 7.13 Snap investigates user reports of underage accounts and provided examples of training material followed by moderators who handle these reports.

---

<sup>43</sup> [Children's Online User Ages Quantitative Research Study \(ofcom.org.uk\)](https://www.ofcom.gov.uk/consult/condocs/childrens-online-user-ages-quantitative-research-study/childrens-online-user-ages-quantitative-research-study)

**Figure 7.3: The alert shown to users under the age of 13 if they attempt to sign-up for a Snap account**



Source: Screenshot provided by Snap in response to Ofcom's information request, June 2022.

## Agreeing to terms and conditions

### Users are presented with Snapchat's terms and conditions when registering to make an account

- 7.14 Terms and conditions appear when registering and signing in. Users are not required to review the terms prior to sign-up, but by signing up they accept the terms of service.
- 7.15 Creators are required to agree to Snapchat's terms and conditions prior to being able to upload content to Snapchat.
- 7.16 Snap told us that its cross-functional team frequently discusses and reviews its terms and conditions, including its Community Guidelines, and related policies; and these discussions inform whether changes to these policies are warranted. Changes may also be made in light of new data or events.

### Snapchat's terms and conditions prohibit terrorist and hateful content

- 7.17 Snapchat's [terms and conditions](#) state that terrorist, extremist and hate groups are prohibited from using the platform, and that Snapchat has no tolerance for content that advocates or advances violent extremism or terrorism.
- 7.18 The terms and conditions also note that Snapchat may remove the offending content, terminate or limit the visibility of an account and/or notify law enforcement. They state that Snapchat engages with global law enforcement when activity poses a threat to human life. In addition to this, Snap is a member of the EU Internet Forum (EUIF) which brings

together governments, civil society, and the tech industry to reduce accessibility to terrorist content online.<sup>44</sup>

- 7.19 In its terms and conditions Snap reserves the right to remove users that the platform has reason to believe pose a clear and present danger to others, on or off of Snapchat. These include leaders of hate groups and terrorist organisations, and individuals with a reputation for inciting violence or behaviour that the platform believes pose a threat to human life.
- 7.20 Snapchat's terms and conditions also prohibit "hate speech or content that demeans, defames or promotes discrimination or violence on the basis of race, colour, caste, ethnicity, national origin, religion, sexual orientation, gender identity, disability or veteran status, immigration status, socio-economic status, age, weight or pregnancy status".

### Ofcom has assessed the clarity of communication of Snapchat's terms and conditions

- 7.21 We found that the terms and conditions were clearly presented, making use of sub-headings, different font sizes and bold text. However, there was no table of contents to make navigation easier. Snapchat's use of bullet points helps readability, but some paragraphs are also very long and harder to digest.
- 7.22 For a user seeking the terms and conditions and community guidelines, we found it fairly straightforward to navigate to and between the relevant webpages.

## Uploading and watching content

### Content on Snapchat is categorised based on whether it is considered suitable for under or over-18s

- 7.23 Snapchat is popular among children, with 42% of children in the UK aged 3-17 using the platform.<sup>45</sup> Under-18 users have a different experience of the platform to those over 18. Snapchat does not distinguish between different age groups for users under the age of 18.
- 7.24 Snap restricts under-18s from viewing advertisements of regulated products such as alcohol, and content flagged as being inappropriate for under-18s by its publisher partners. In addition, in instances where Snap discovers that an advertiser or a publisher partner has not set appropriate restrictions on the users who can view their content, the content is removed.
- 7.25 All content on the Discover channel is pre-moderated before being promoted for wider visibility on the platform.
- 7.26 When deciding what measures to put in place to protect under-18s, Snap considers the regulatory landscape, the platform's Privacy and Safety by Design Principle's, Community Guidelines, terms and conditions and any user reports. Each user report is reviewed by

---

<sup>44</sup> [EUIF \(ec.europa.eu\)](https://ec.europa.eu/euif/)

<sup>45</sup> [Children and parents: media use and attitudes report 2022 \(ofcom.org.uk\)](https://www.ofcom.gov.uk/consult/condocs/childrenandparents/childrenandparents2022/)

Snap's Trust and Safety team and assessed against the Community Guidelines. In the case of potentially disruptive global events, the team is briefed for potential increases in content violating its terms and conditions and Community Guidelines. Snap is alerted by trusted sources of information and third parties on any false information. As a result, the Trust and Safety team then decide to approve or delete the reported content or decide to delete the account entirely.

- 7.27 Snap also provided us with details of the training Trust and Safety staff are given during onboarding as well as information on how staff are kept informed of emerging safety trends.

### **Snapchat provides tools and information for users with the aim of improving their media literacy**

- 7.28 Snapchat provides several tools and resources to assist under-18 and adult users navigate the platform in a safe and secure way. These range from on-platform information integrated into product design features to resources for users and other stakeholders hosted on the off-platform Safety and Transparency Centres.

### **As users access and post content, they are offered videos and tools to help them stay safe**

- 7.29 Snapchat offers the 'Here For You' feature, an innovative feature that prompts users towards support if they search for terms that might indicate they're at risk of harm. This includes search terms related to suicide and self-harm. The prompts take potentially at-risk users to resources from local partners on mental health support. Snap also sends users suicide prevention resources when it finds users who post content relevant to self-harm, in addition to removing the content.
- 7.30 On the Discovery feature Snapchat promotes videos designed to reduce the risk of potential harms to users. Examples of the content promoted on this channel include its 'Heads up' campaign which informs users about the dangers of drugs and other substances, and an official account called 'Privacy Snapshot' which is used to share safety and privacy tips and tricks with users.
- 7.31 Snapchat uses notifications and prompts across the platform to raise users' awareness of causing certain harms. For instance, it sends warnings to users posting content that violates its community guidelines, reminders to users regarding prohibited content in its terms and conditions after a video has been uploaded, and on-platform prompts to notify users of changes to its terms and conditions or other policies.
- 7.32 Snapchat also offers resources to users to access outside the platform, on their website. The Snapchat Safety Centre and Privacy Centre contain resources to help users, parents, and the wider public understand and navigate the platform. It has published a Parent's Guide to Snapchat, which aims to provide caregivers with information about the platform's operation and support its safe use, alongside other articles such as one on Tips for Staying Safe Online. Snap report that its Privacy and Safety Centres are visited by approximately 80 thousand UK users every month.

### Snapchat recently launched parental controls

- 7.33 In August 2022, Snap launched its parental control feature, Family Center, which allows parents and guardians to view a list of their child's conversations without seeing the message's content. Parents and guardians can also view who their child is friends with on the service, as they are given access to a list of their most recently-added friends.
- 7.34 To set up Family Centre, the child must accept an invitation sent by parents and guardians through Snapchat's direct chat. As with regular conversations on Snapchat, the message will delete either immediately or 24 hours after the child has viewed it, depending on their settings. For this reason, parents are encouraged to remind their children to accept the invite.
- 7.35 Snap plans to introduce content controls for parents and guardians, alongside additional features for the Family Centre, which will allow parents or guardians to receive notifications when their child has reported an account or content.
- 7.36 As Snapchat is currently launching its parental controls, we recognise the challenges of providing information on effectiveness and take-up given that these tools are yet to be rolled out to users. Yonder research commissioned by Ofcom showed that 47% of parents have never used parental controls on Snapchat, despite being aware of them. Among those using parental controls on Snapchat, 91% found them easy to set up.<sup>46</sup> We will be considering the effectiveness of these measures in our next report.

## Detecting harmful content

### Users can report harmful content using a tool on the platform

- 7.37 In-app reporting consists of holding down on content or tapping the "report snap" option which is visually highlighted to the user in red text (with the other options in black).
- 7.38 Users can report against the following categories: Bullying & harassment; Nudity & sexual content; Threats, Violence & dangerous behaviour; Hate speech, terrorism & violent extremism; Drugs & weapons; Suicide & self-harm; False information; Intellectual property. Further sub-categories are available under each category including a "it involves a child" sub-category under the "Nudity & sexual content" category.
- 7.39 In addition to in-app reporting, individuals can report content via the safety centre and the support website. Users do not need to be logged into Snapchat to report content.
- 7.40 Snap has confirmed that it operates a Trusted Flaggers programme to detect harmful content.

### User reports are reviewed either manually, automatically, or a combination of both

- 7.41 Snap told us that every user report is reviewed by its Trust & Safety team by humans and through automated processes.

---

<sup>46</sup> [VSP Parental Guidance research 2022 \(ofcom.org.uk\)](https://www.ofcom.gov.uk/consult/condocs/vsp/vsp2022/vsp2022.pdf)

- 7.42 Snap anticipates that some world events may cause an increase in user reports, and its teams are briefed on this before they take place.
- 7.43 As an example of automated processes, to combat misinformation Snap leverages technology that automatically detects and rejects keywords commonly found in false information content.

### **Snap moderates content that is likely to reach a large audience**

- 7.44 Content that is submitted by users to be included in the Spotlight feed is moderated before it is shared more widely.<sup>47</sup>
- 7.45 Snapchat promotes some popular user-generated content on its Discover channel alongside content from media companies. Snap moderates this content prior to being promoted to ensure that harmful content is not widely shared.

## **Enforcing against harmful content**

### **Sanctions are in place to deal with content that breaches the Terms of Service**

- 7.46 Snap uses a variety of Trust & Safety moderation sanctions, including:
- Deleting a user account
  - Locking a user out of their account
  - Deleting videos content (Snap or Story)
  - Suppressing video content from being shared widely in the Spotlight and Discover feeds
- 7.47 Content moderators on average take two hours to review for potential violations. For violations that are not sufficiently severe to warrant account-level sanctions (i.e. account locking or deletion) based on a single offence, the user's record of previous offenses and warnings will be referenced to determine whether account-level action is nonetheless warranted based upon a pattern of harmful behavior. Otherwise, the platform will remove the piece of content and warn the user of the violation.
- 7.48 Snapchat also reserves the right to revoke verification of Snap Star status. Snap Star status is a form of verification granted by Snap to users with large follower counts who produce high quality and authentic content.<sup>48</sup> This status can be removed for reasons including violations of their terms of service or community guidelines, or dangerous behaviour that puts others at risk of harm.

### **Hate and terror content is prioritized over other kinds of potential violations**

- 7.49 Snap asks third-party content moderators to prioritise content in their queue depending on the severity of the violation type. Hate and terror content are considered high priority, as well as extremism, self-harm, threats/violence, weapons, and CSAM. Snapchat will also

---

<sup>47</sup> [Are Snap Submissions moderated before being posted to Spotlight? \(snapchat.com\)](#)

<sup>48</sup> [How do I become a Snap Star? \(snapchat.com\)](#)

deactivate an account if it appears to be dedicated to hate or terror content, with the Snaps taken on that account also deleted.

### **Moderators are trained to ensure consistency when applying sanctions**

- 7.50 Agents are also provided with a non-exhaustive table of prohibited slurs that are actionable under Snap's policies against hate speech, which is subject to frequent updating.
- 7.51 Snap also conducts frequent and in-depth collaboration sessions with its teams that review more nuanced cases and ensure consistency of policy application.

### **Users are informed when sanctions are applied and can appeal against sanctions**

- 7.52 Creators are informed of violations, most typically by receiving a warning or notice of violating content upon signing into Snapchat for the first time after the imposition of a sanction. If a user wishes to contest the sanction, Snap provides a form<sup>49</sup> for users to appeal content and account sanctions.

---

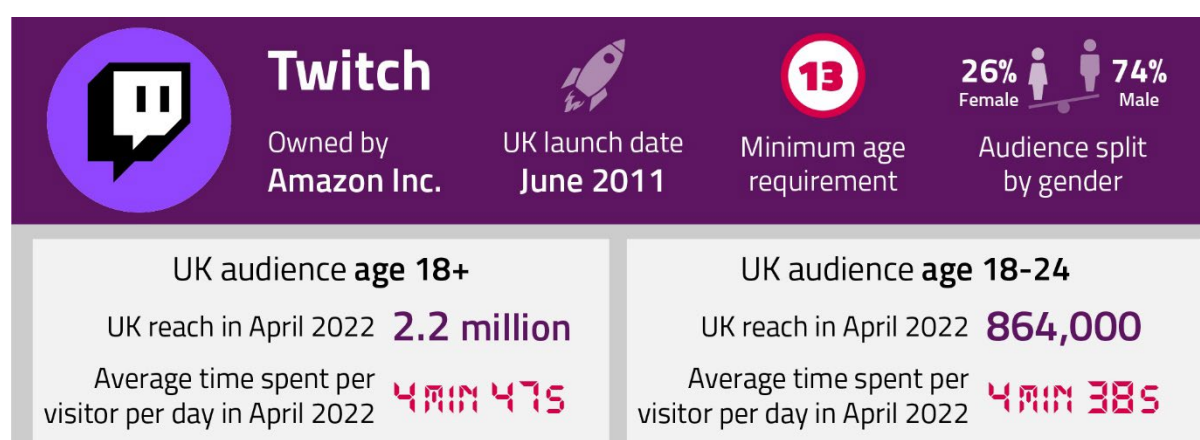
<sup>49</sup> [Contact Us \(snapchat.com\)](https://www.snapchat.com/contact-us)

## 8. Twitch

### Introduction

Twitch is a video-sharing platform predominantly known for its livestreamed content. The platform heavily features video game livestreams; users can watch other people playing video games or stream themselves doing so. Twitch also broadcasts e-sports competitions, as well as music and other entertainment broadcasts. Viewers can interact with each other in real time in a chat bar alongside live video content. Twitch has diverse revenue streams consisting of subscriptions, donations, and advertising.

Figure 8.1.: Key information on Twitch



Source: © Ipsos, Ipsos iris Online Audience Measurement Service, 1 - 30 April 2022, age 18+ for reach and time spent data, age 15+ for audience share by gender data, UK.

### Key findings

- 8.1 Viewers do not need to create an account on Twitch to watch video content. This means that many viewers do not have to declare their age or read or agree to any terms and conditions. However, only users with an account can report videos containing harmful content within the app – users without an account have to send an email.
- 8.2 **Viewers of any age, with or without an account, can easily watch videos that are marked by creators as 'mature'.<sup>50</sup>** Even when viewers do declare their age as under-18, they are able to watch mature content.
- 8.3 **Twitch is unique in its approach to enforcing on-platform sanctions, such as account bans, for severe off-platform conduct,** such as terrorist activities or sexual exploitation of children.

<sup>50</sup> Mature content is defined by Twitch simply as content that 'may be inappropriate for younger audiences'. Twitch specifies to users who select the mature content tag that they 'may never broadcast sexual activity, nudity, threats or extreme violence.'



8.4 All reported videos are reviewed by human moderation teams.

## Engagement with Ofcom

8.5 Twitch engaged positively with Ofcom during the first full year of the VSP Regime. For example, Twitch voluntarily offered teach-ins to Ofcom employees on several subjects. Twitch provided detailed information in its response to our formal requests, while at the same time engaging on issues related to the Buffalo shooting.

## Governance and risk management

### Decision-making structure

#### Twitch includes teams across the organisation in decisions about its protection measures

8.6 Twitch told us that decisions regarding protection measures on the platform are made across the Legal, Trust & Safety, Community Health Products, Advertising Product and Advertising Policy teams. The table below shows which teams have primary responsibility for decisions about each measure, and the teams that are involved in those decisions.

Figure 8.2: The teams within Twitch with responsibility for decisions about measures

Measures	Primary responsibility	Key stakeholders
Community Guidelines User Age Restrictions Prohibited Games	Trust & Safety (Policy)	Legal, Trust & Safety, Community Health Products
User Report Tool Account Enforcements Appeals Portal Safety Center	Trust & Safety (Operations) Trust & Safety (Law Enforcement Response)	Legal, Community Health Products, Trust & Safety
Proactive Detection	Community Health Products	Community Health Products, Legal, Trust & Safety
Terms of Service	Legal	Legal, Trust and Safety,
Mature Content Label	Advertising Product	Advertising Product, Trust & Safety, Legal

8.7 Within the Trust and Safety team, Twitch has a Safety Operations team that enforces the company's Community Guidelines and Terms of Service. Within the Safety Operations team there is a team that evaluates and acts on user reports of violations of Twitch's rules.

## Risk assessment

**Twitch analyses data relating to risks on its platform twice a year and publishes some of this in transparency reports.**

- 8.8 It uses this process to look at what is being reported to it by users and to analyse the statistics to reactively assess what risks exist on the platform and what can be done to mitigate them.
- 8.9 Twitch also said it considers the percentage users that are under 18 when deciding what safety measures it should implement.

## Assessing effectiveness

**Twitch uses the metrics published in its transparency reports to assess the effectiveness of its safety measures**

- 8.10 Twitch did not provide a full list of the metrics it collects to assess the effectiveness of its measures, instead referring us to the published metrics in its transparency reports. These reports can be found on the [Twitch Safety Center](#).


**Twitch assessed its measures against the framework of the VSP regulations**

- 8.11 Twitch reviewed its processes against the measures set out in Schedule 15A. It assessed two of the measures in Schedule 15A against the practicability and proportionality criteria and decided not to implement them:
- Parental controls:** Twitch deemed it would not proportionate or practicable to implement this measure, as under-18 protections could be achieved through other measures it had in place.
  - Systems for viewers to rate harmful material:** Twitch deemed that these tools would not be practicable and proportionate to implement given as it is predominantly a live-streaming platform. It decided that it makes most sense to give viewers the opportunity to mark content which should not be permitted on Twitch via viewer reporting tools and to rely on creators to mark content which may be permitted but rated as mature. Twitch considered that creators will have a better idea of what their live content may contain and are better placed to rate it as mature, but viewers can still signal to Twitch when content may need to be removed.


## User journey

Figure 8.3: Twitch user journey

---

<p>User opens website or app</p>		<p>Viewers can watch livestreams on the web with or without an account. Creators live streaming must have an account. Creators can use the Twitch mobile app or use Twitch's own streaming software if on desktop.</p>
----------------------------------	-------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

---

<p><b>User sign-up and log-in</b></p>		<p>Users input their birthdate and must be over 13 in order to create an account (T&amp;Cs require under-18s to have parental permission).</p> <p>Twitch reviews user reports to monitor any accounts that may be held by under-13s.</p> <p>Twitch does not have any parental controls.</p> <p>It allows parents to request account suspension if their child is found to have an account.</p> <p>Viewers must agree to Twitch's terms of service when creating an account. This includes a requirement that under-18s only use the platform when supervised by an adult.</p>
<p><b>VSP recommends or user searches for content</b></p>		<p>Viewers receive warnings about content labelled as mature. All viewers can still see this content, including those with under-18 accounts.</p>
<p><b>User watches and engages with content</b></p>		<p>Twitch's terms of service prohibit a range of illegal and harmful content</p> <p>Twitch highlights that the majority of content is livestreamed, with only a few pre-vetted partners and creators able to share pre-recorded videos.</p> <p>Community Moderators (typically volunteer users) engage and monitor content interactions.</p> <p>Creators can mark their content as mature.</p>
<p><b>User encounters harmful material and flags or reports</b></p>		<p>Users must create an account and sign-in to report videos.</p> <p>All reported videos are reviewed and acted upon by moderation teams.</p> <p>Users and community moderators can report any prohibited content.</p> <p>Twitch also uses AI to flag potentially harmful material which is then reviewed.</p>
<p><b>User informed of outcomes</b></p>		<p>Where enforcement action is taken (and for the reason the user reported) the user will be contacted via email.</p>
<p><b>User access to other tools</b> <i>(e.g. Media Literacy)</i></p>		<p>Twitch's terms of service prohibit a range of illegal and harmful content</p> <p>Community Moderators (typically volunteer users) engage and monitor content interactions.</p> <p>Creators can mark their content as mature.</p>

## Signing on

### Viewers can watch videos without creating an account, meaning they can watch mature content without first declaring their age

- 8.12 Viewers on Twitch can create an account, but do not need to do so to view the majority of the video content available on the platform. Creators must create accounts and be logged in to live stream or upload video content to Twitch.
- 8.13 If viewers without an account search directly for mature content streams or a mature category, they will see a message that tells them the stream or category is mature and asks

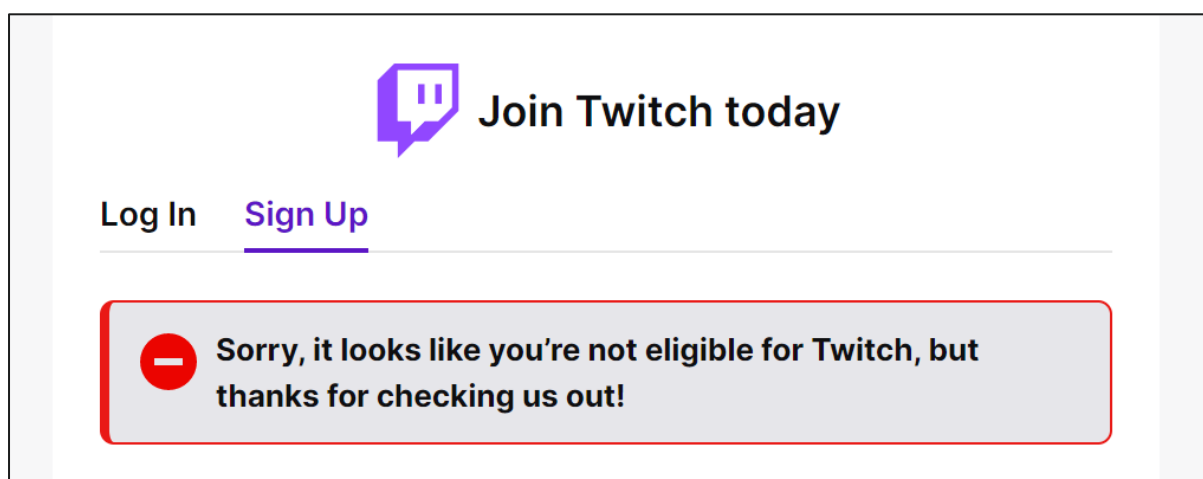
if they want to continue. If viewers click yes, they are free to watch mature content without any proof, or even declaration, of their age.

- 8.14 Mature content is defined by Twitch simply as content that 'may be inappropriate for younger audiences'. Twitch specifies to users who select the mature content tag that they 'may never broadcast sexual activity, nudity, threats or extreme violence'.

#### When users create an account, they are asked to enter their date of birth

- 8.15 Prospective users are asked to self-declare their date of birth when registering for an account on Twitch. If the date entered indicates that the user is under 13, they are unable to create an account for a certain period of time.

**Figure 8.4: The alert shown to users under the age of 13 if they attempt to sign-up for a Twitch account**



Source: Screenshot provided by Twitch in response to Ofcom's information request, June 2022

- 8.16 Twitch uses a combination of user reports and machine learning to identify underage accounts. Suspected underage users are asked to provide ID to confirm their age to avoid account deletion.
- 8.17 When considering protection measures for different age-groups, Twitch only considers two age categories: 13-18 and over-18.
- 8.18 Account holders who have declared themselves to be 13-18 year-olds must agree in the terms of service that they will not use Twitch unsupervised.

#### Twitch does not have any parental control measures that allow guardians to control the content an under-18 can access

- 8.19 Twitch does not have parental controls available to users. It considered the practicability and proportionality of parental controls and decided that its terms and conditions provide adequate protections for under-18s.
- 8.20 Twitch's [Terms of Service](#) state that under-18s may only use "Twitch Services under the supervision of a parent or legal guardian who agrees to be bound by these Terms of

Service". However, viewers can search and view mature content without having set up an account on the platform and therefore without having accepted, or even read, the Terms of Service.

- 8.21 If Twitch is made aware that a user is violating this term, through a guardian, another user or by the admission of the underage user themselves, Twitch will suspend the account pending receipt of parental permission. Twitch does not publish figures for the number of accounts that are suspended for this reason. It is therefore unclear how often this happens in practice.
- 8.22 Parents and guardians can request for their child's account to be deleted. Upon receiving the request, a customer service agent will share a certification form which includes instructions on how to close the account within 48 hours.

## Agreeing to terms and conditions

### Users are presented with Twitch's terms and conditions when registering to make an account

- 8.23 These terms and conditions include Terms of Service, a privacy notice and Community Guidelines.
- 8.24 When signing up to Twitch and making an account, users must confirm they have read and agree to Twitch's Terms of Service and Privacy Notice. The Terms of Service contain Community Guidelines which outline categories of material which are not allowed on the platform.
- 8.25 Users who wish to review the Terms of Service outside of the context of sign-up may do so from the Twitch homepage on desktop and via the settings in the mobile app.

### To upload content to Twitch, creators must create an account and sign in

- 8.26 Creators uploading content must make an account and sign in, and therefore will have been presented with Twitch's terms and conditions.
- 8.27 When beginning a stream from the mobile app, creators are prompted to comply with the [Community Guidelines](#). Creators on desktop cannot stream directly from the webpage, and instead need to use special software, this means that a reminder of the Community Guidelines is not provided to those streaming from desktop.
- 8.28 To upload pre-recorded video content (i.e. not livestreamed) on Twitch, creators are required to obtain 'Affiliate or Partner status'. To obtain this status, a creator must have a proven record of behaviour and streaming content that complies with the Terms of Service and Community Guidelines. When signing up to the Affiliate and Partnership programme creators must once again read and agree to the Terms of Service, and face having the status removed if they do not adhere to them.

### **Twitch continuously updates its terms and conditions and communicate these changes to users in a variety of ways**

- 8.29 The terms are continuously reviewed and updated through the Policy Authorization Process (PATH), with changes prioritised based on the severity and scale of the associated harm. Twitch's response did not provide further detail and specifics on the updating process.
- 8.30 Twitch updated its Community Guidelines following the Buffalo shooting. It observed an increase in conversation on 'The Great Replacement Theory', a white supremacist theory claiming that immigration patterns will result in the extinction of 'the white race'. Twitch worked with external misinformation consultant Global Disinfo Index (GDI) and concluded that this theory is closely associated with hateful ideologies. The result was an internal review and update of the Hateful Conduct and misinformation policies to specify that supporting/espousing the Great Replacement Theory should be considered a violation under the policy. It also provided training materials on this topic to those enforcing the policy.
- 8.31 Changes and updates to Twitch's terms and conditions are communicated to users differently depending on the 'significance' of the change. Twitch informs creators directly of changes through the creator dashboard, as well as broad and targeted emails. Outside of these communication channels, Twitch also shares details of changes through blog posts, Twitch-produced shows broadcast on in its platform, and Twitter posts. For particularly extensive or notable changes, Twitch engages with interested press outlets to "amplify the changes".

### **Twitch's terms and conditions prohibit terrorist and hateful content**

- 8.32 Twitch's Community Guidelines are incorporated into its Terms of Service, which prohibit hate and terror content and any associated material, as well as violence and threats on the platform. Twitch does not allow content that "depicts, glorifies, encourages, or supports terrorism, or violent extremist actors or acts... [including] threatening to or encouraging others to commit harmful or illegal acts." Users also cannot display or link terrorist or extremist propaganda, including pictures and videos of associated violence, even for the purposes of denouncing it.
- 8.33 The Community Guidelines state that Twitch will indefinitely suspend accounts associated with hate and terror and in "exceptional circumstances" Twitch can pre-emptively suspend accounts where it believes there is a high likelihood of inciting violence offline.
- 8.34 Amazon, the owner of Twitch, became a member of the Global Internet Forum to Counter Terrorism (GIFCT) in 2019. Ofcom understands that to be granted membership to GIFCT a company must, among other criteria, demonstrate 'terms of service, community guidelines, or other publicly available policies that explicitly prohibit terrorist and/or violent extremist activity'.<sup>51</sup>

---

<sup>51</sup> [Membership \(gifct.org\)](https://www.gifct.org/)

- 8.35 On hateful content, Twitch states that it has “zero tolerance” approach for hateful conduct and “act[s] on every valid reported instance of hateful conduct”. The Community Guidelines prohibit the uploading or streaming of hateful conduct and harassment. Hateful conduct is defined in the Community Guidelines as “any content or activity that promotes or encourages discrimination, denigration, harassment, or violence” this based on those with specific protected characteristics, including race, ethnicity, religion and gender identity.
- 8.36 The Community Guidelines acknowledge that there are many manifestations of harassment including stalking, personal attacks, promotion of physical harm, hostile raids<sup>52</sup>, and malicious false report brigading.<sup>53</sup> Sexual harassment is also specifically highlighted, outlining this can be formed of unwelcome sexual advances and solicitations, sexual objectification, or degrading attacks relating to a person’s perceived sexual practices.

### Ofcom has assessed the clarity of communication of Twitch’s terms and conditions

- 8.37 We found that Twitch’s Terms of Service were not particularly user-friendly, especially considering Twitch’s core demographics, with lots of long and very detailed paragraphs. Twitch’s Terms of Service contain a list of each section with hyperlinks to allow users to navigate to different sections.
- 8.38 Twitch’s Community Guidelines are more user-friendly than its Terms of Service, containing a list of all subheadings and hyperlinks to each in a box on the side of the webpage. Each section is also shorter and clearer than in the Terms of Service.
- 8.39 Twitch’s Terms of Service and Community Guidelines are not immediately visible from its homepage – users have to click three dots at the top of the page, and scroll down to see the options ‘Community Guidelines’ and ‘Terms’.

## Uploading and watching content

### Twitch allows creators to tag their content as mature

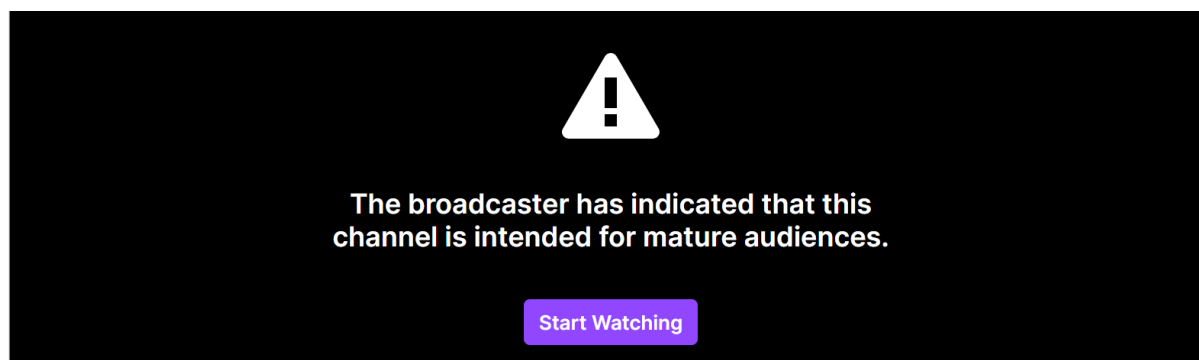
- 8.40 Creators on Twitch can tag their channels as mature if they plan to stream content that is not suitable for younger audiences.
- 8.41 When a viewer encounters a stream labelled mature for the first time, they are shown the warning in figure 8.5.

---

<sup>52</sup> ‘Raids’ are when a user on Twitch sends their viewers to another streamer, usually as a show of support. Hostile raids are when the user instead sends their viewers to post hate on another streamer’s chat.

<sup>53</sup> ‘Brigading’ is a term that originated on Reddit for a coordinated attack by a group of users of an antagonistic subreddit (forum dedicated to a particular topic). For more details, see [Social Media Futures: What is Brigading?](#).

Figure 8.5: The Twitch mature content warning



Source: Screenshot provided by Twitch in response to Ofcom's information request, June 2022

Once viewers have seen this warning, they may begin watching the content by selecting the start watching button, they are not then shown this warning for any subsequent content they wish to watch that is rated as mature.

- 8.42 This process is the same for all viewers on the site, even those who have created an account and identified themselves as under-18.
- 8.43 When a creator sets a mature content flag, they are given a reminder of the Terms of Service to prevent them from uploading content that breaks Twitch's terms and conditions. They are shown the following message: "Enable this setting if your stream contains content that may be inappropriate for younger audiences. You may never broadcast sexual activity, nudity, threats, or extreme violence. Doing so will result in immediate, irrevocable termination of your account. Please make sure your content will comply with the Terms of Service before broadcasting".

#### **Twitch has measures planned to stop under-18s from stumbling upon mature content**

- 8.44 Twitch has shared with us its plans to make changes to its mature content labelling and access systems. However, as we understand the proposal, Ofcom has concerns that these plans may not fully address this problem and will continue to engage with Twitch on this.

#### **Most of Twitch's media literacy tools are provided off-platform**

- 8.45 Twitch provides tools and information to advance its users' media literacy, these are almost all provided via off-platform webpages. Resources on its Help Page and Safety Center include plain language materials explaining safety tools, reporting process, moderation practices, and guidance for account management.
- 8.46 Twitch's Safety Center hosts information targeted at both under-18 and adult users to raise awareness of on-platform safety tools including viewer controls and managing harassment.
- 8.47 A series of articles and user guides are provided off-platform. For example, Twitch has published a Guide for Parents and Educators with resources relevant for parents/caregivers and educators' role protecting under-18 users. The guide explains Twitch's age restrictions, reporting mechanisms as well as how to set chat filters and moderate live streams. It also aims to educate parents/caregivers and educators on topics such as avoiding online harms



including scams and risks of impersonation. Twitch has additionally published a guide for all users on [Crisis Prevention](#) and [Preventing Doxxing, Swatting and other IRL harm](#).

- 8.48 The off-platform Help Center gives users access to searchable resources and information on a wide range of other topics, some of which are related to safety, such as its dedicated page on Moderation and Safety.

## Detecting harmful content

### Logged-in users can report videos containing harmful content to Twitch using an on-site reporting tool

- 8.49 Users can report videos containing harmful content to Twitch using the on-site reporting tool.
- 8.50 To use the on-site reporting tool, viewers must create an account and be logged-in to Twitch. Given that being logged-in is not a requirement for watching content, many viewers may be watching content and not have the ability to quickly report videos they consider contain harmful content.
- 8.51 Twitch has an email address that users without an account can contact to flag content or voice complaints. However, the ability to flag content is not embedded into the design of the platform interface for users without an account – rather than clicking on a flag icon near the video, for example, users must find the email address and draft a separate email. This creates extra friction and may result in slower, or less reporting.
- 8.52 When reporting via the on-site tool, users can select to give further details on the harmful content they have seen. Viewers can select where they found the harmful activity or content, e.g. livestream, clip, chat message. They can also select the kind of harmful content they are reporting e.g. Hateful content, Terrorism, Nudity or Sexually Explicit.
- 8.53 When a user submits a report they receive confirmation in the form of a dialogue box and an email. When Twitch takes action on content as a result of a user report, that user is notified via email.

### Content moderation professionals review all user reports

- 8.54 Twitch uses content moderation professionals to review and respond to potentially offending content and has teams dedicated to reviewing certain classes of high-priority content such as hate and terror and other potentially illegal content.
- 8.55 All reports submitted to Twitch by users are reviewed by these trained specialists. These specialists also review content that is signalled by Twitch's automated tools. When these tools detect potentially harmful content, these are always reviewed by a human before action is taken.
- 8.56 Twitch prioritises reports within its moderation system to ensure that reports of the most harmful behaviour receive prompt attention by moderators. Twitch's training materials for moderators set out a number of factors, including the severity of harms, so that moderators know which harms to prioritise.

- 8.57 Twitch has a dedicated Law Enforcement Response team of experienced investigators to investigate the most egregious reports related to hate and terror. This team works with the Met Police and NGO's such as the Diana Trust.
- 8.58 Where users select that a report is about CSAM content, these reports are transferred to a specialist internal team. This team will submit its findings to the National Center for Missing and Exploited Children (NCMEC).

### **Twitch takes a 'layered approach' to safety that combines user reporting, machine detection, and moderators**

- 8.59 Twitch attached its quarterly Transparency Reports to its response, which outlines its "layered approach" to safety infrastructure. This layered approach combines its Community Guidelines, Service Level Safety, Channel-Level Safety and Viewer-Level safety.
- 8.60 The approach relies on a combination of user reporting, machine detection, and moderators. The Global Safety Operations team works in tandem with auto-moderation systems that work 24 hours a day, 365 days a year to respond swiftly to user reports, and to help further mitigate harm.<sup>54</sup>

## **Enforcing against harmful content**

### **Twitch applies a variety of sanctions depending on the nature and circumstances of the content**

- 8.61 Twitch maintains an internal set of guidelines that set forth sanctions for violative content. The sanction applied depends on the nature and circumstances of the violative content, creators face increasingly severe sanctions ranging from:
- Warnings
  - Loss of user privileges
  - Temporary suspensions
  - Indefinite suspensions
- 8.62 Twitch issues warnings to users for a small number of less severe violations, building to temporary suspensions for repeated, less severe violations, which can last between 24 hours and 30 days, depending on the severity of the violation. Another sanction is the loss of user privileges such as being featured on the home page or in marketing campaigns or participation in programmes and events. Indefinite suspensions are reserved for the most severe violations and for serious repeat offenders.
- 8.63 Twitch will investigate and issue sanctions to users if it receives verifiable evidence of certain conduct that takes place off platform. [Twitch's Off-Service Conduct Policy provides](#) details of off-platform conduct liable to generate sanctions.

---

<sup>54</sup> [H2 2021 Transparency Report \(twitch.tv\)](#)

### How sanctions are applied for hate and terror content

- 8.64 Twitch moderators consider numerous factors including the context of the content, the intent, the potential for, or actual harm, to the community and legal obligations, as well as consideration of prior offenses. The response explained that the sanction is then decided based on the conduct category.
- 8.65 Moderators use a strike guide to determine the appropriate action that includes such information as enforcement criteria, examples of violative content and penalties associated with such actions.

### Users are informed when sanctions are applied

- 8.66 Twitch users receive an email notification once the sanction has been imposed, to inform them of the type of sanction applied, the relevant duration and detailed reasoning for the sanction. Twitch also provides details to the user of their rights and means to appeal the decision.

### Users can appeal against sanctions

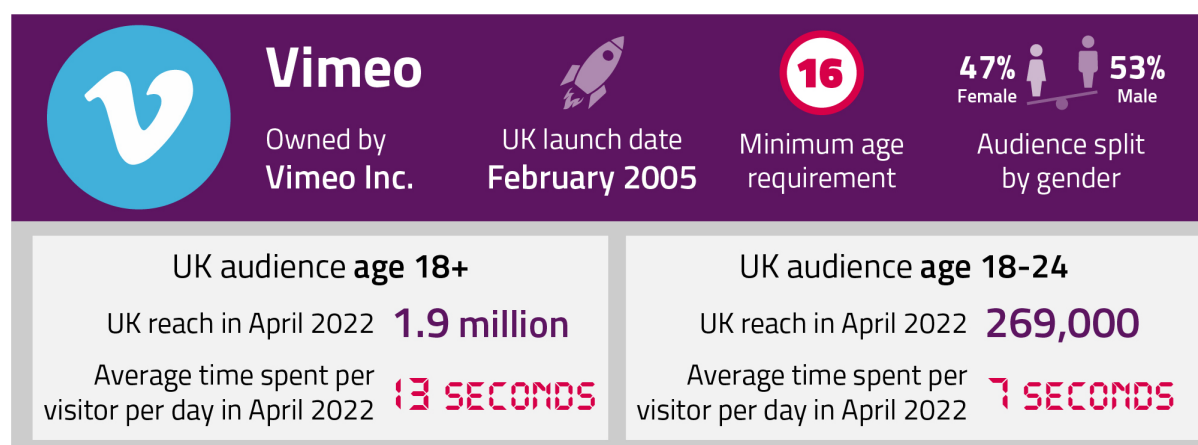
- 8.67 Twitch operates an 'Appeals Portal' which allows users to appeal an enforcement sanction that the user believes to be incorrect, unwarranted, or unfair. Users can only appeal against enforcement sanctions issued within the last 60 days, and in relation to indefinite suspensions.

## 9. Vimeo

### Introduction

Vimeo is a video-sharing platform that allows users to share high-quality videos. Vimeo makes money by selling platform subscriptions to businesses. Seventy-one percent of Fortune 500 companies have a paid subscription with Vimeo for their video sharing activity, with another 18% having at least one free subscription.<sup>55</sup> Vimeo offers customisable video templates and access to stock photos to allow businesses to create professional video content. Users can also password-protect videos in order to allow users to create and share videos with a limited group of people. Vimeo does not consider under-18s a target audience for its platform.

Figure 9.1: Key information on Vimeo



Source: © Ipsos, Ipsos iris Online Audience Measurement Service, 1 - 30 April 2022, age 18+ for reach and time spent data, age 15+ for audience share by gender data, UK.

### Key findings

- 9.1 **Vimeo does not allow users without an account to view videos that are rated mature or left unrated**, reducing the likelihood that under-18s will stumble across unsuitable videos. This is a change it has made in response to VSP regulation. However, desk research indicates that **Vimeo does not ask users to declare their date of birth when creating an account**.
- 9.2 Vimeo carried out a risk assessment on its service in response to VSP regulation.

### Engagement with Ofcom

- 9.3 During the first full year of the VSP Regime, Vimeo has demonstrated effective regulatory engagement with Ofcom. Vimeo met with us throughout the year and were open about

<sup>55</sup> [Vimeo 2022 Annual Shareholder Letter \(vimeo.com\)](https://www.vimeo.com/press/2022-02-02-annual-shareholder-letter), February 2022.

the protection measures it has in place and its plans for improvement. Vimeo's response to our information request was notably comprehensive and detailed.

## Governance and risk management

### Decision-making structure

#### Vimeo includes Policy, Trust and Safety, Product and Legal professionals in decisions about suitable protection measures

- 9.4 Vimeo's Director of Policy is in charge of:
- a) reviewing Vimeo's requirements under the VSP Regime;
  - b) carrying out assessments to determine Vimeo's risk level;
  - c) determining which measures are relevant to Vimeo;
  - d) working with other teams within Vimeo to develop and implement solutions to comply with relevant measures; and
  - e) working with other teams to monitor and review the effectiveness of the measures implemented; make recommendations where needed; and implement changes.
- 9.5 The General Counsel, Director of Trust and Safety Operations and those in the Product team are also key stakeholders in decisions relating to protection measures. Any decisions made by the Director of Policy will be approved by Vimeo's General Counsel. Vimeo's Director of Trust and Safety Operations is responsible for identifying areas of risk and developing solutions. The Vimeo Product Team provides support in developing solutions to comply with VSP regulations.

### Risk assessment

#### Vimeo considers the nature of its platform and its audience when assessing risk

- 9.6 Vimeo considers itself to be a platform for businesses to host video content, as opposed to a social network site. As such, Vimeo does not consider its platform at a high risk of being used as a content viewing platform for the viral spread of videos containing harmful content.

### Assessing effectiveness

#### Vimeo told us that it assesses the effectiveness of its 'mature content'<sup>56</sup> tag using a variety of metrics

- 9.7 These metrics are:

---

<sup>56</sup> Where content is permissible based on Vimeo's terms of services but may not be suitable for all audiences, users are asked to rate content as 'mature'.

- a) The weekly count of unique users who rate their content as mature,
- b) The count of users who mark their content as including adverts, and
- c) The weekly count of unique 'mature or unrated' clips that are blocked for public users (i.e. those users who are not logged in).

## User journey

Figure 9.2: Vimeo user journey

<b>User opens website or app</b>		Some content is available to view without logging in to an account, but Vimeo does not allow content labelled as mature or left unrated to be publicly available.
<b>User sign-up and log-in</b>		Vimeo does not apply specific age verification tools. Viewers must agree to Vimeo's terms and conditions when creating an account. Vimeo terms and conditions requires users to be over 16 to create an account or to use a service. However, users are not asked their age when creating an account.
<b>VSP recommends or user searches for content</b>		Vimeo allows users to set filters for the type of content they can see. Vimeo does not pre-moderate any videos uploaded. Vimeo does not apply parental controls, stating that under-16s are not allowed under its terms and conditions.
<b>User watches and engages with content</b>		Vimeo's terms and conditions do not allow users to upload illegal, harmful, or restricted material. Where content is permissible based on Vimeo's terms and conditions but may not be suitable for all audiences, users are asked to rate content as "mature".
<b>User encounters harmful material and flags or reports</b>		Users can report content and behaviour that violates Vimeo's policies. Vimeo endeavours to review specific material that is flagged by users, third-parties or its software-based systems.
<b>User informed of outcomes</b>		Users are sometimes informed when sanctions are applied depending on the nature of the content uploaded. If moderators determine content is not in violation, it will be 'whitelisted' so that the user can no longer report the same content again.
<b>User access to other tools (e.g. Media Literacy)</b>		Vimeo notes it has limited media literacy tools as its focus is not as a social media platform, but is used by small businesses, larger enterprises, marketers, agencies and creative professionals. Vimeo has some guidance documents and it also provided support to the UN's initiatives around Covid misinformation (Validate).

## Signing on

### Users need to create an account to watch videos rated as mature or left unrated

- 9.8 Users can create an account on Vimeo and need to do so to upload video content.
- 9.9 Users without an account can view some video content, but not content that is rated mature or left unrated. This measure is intended to stop under-18s from being exposed to videos containing unsuitable content. For the same reason, users who are not logged in are not able to hover over the thumbnail of a mature or unrated clip to auto-play it, they are shown a static thumbnail for these clips.
- 9.10 When a user that is not logged into their account tries to view content rated mature or unrated, they are prompted with a message explaining that the content may not be suitable for all audiences, and they will be required to log in to view.

### Users must be over 16 to use Vimeo, but there are no age assurance methods in place nor any parental controls

- 9.11 Vimeo's terms and conditions require users to be at least 16 years old or the applicable age of majority in their jurisdiction (whichever is greater), to use the platform.
- 9.12 Users must agree to these terms and conditions, but desk research indicates that Vimeo does not ask users to declare their date of birth when creating an account.
- 9.13 Furthermore, Vimeo does not have a parental control system on its platform. Vimeo has said that this is because it does not see under-16s as a target audience.

## Agreeing to terms and conditions

### Users are presented with Vimeo's terms and conditions when registering to make an account, and are informed of any changes to this

- 9.14 Users are required to agree to Vimeo's terms and conditions ('Terms of Service') when joining Vimeo and must have an account to upload content.
- 9.15 Vimeo gathers information from its Legal, Trust & Safety, Security, Product, and other teams regarding the services and user behaviour on its platform. The platforms says that it makes changes to its terms and conditions, including the Acceptable Use Policy, if and when needed.
- 9.16 Vimeo publishes changes to its terms and conditions once a month, on the 15th. It also conducts a review of all legal documents, including its Terms of Service, and policies once a year.
- 9.17 Users are informed of changes and updates by a notification banner on the main landing page and an email is sent to users' personal addresses. Users are suggested to read through these changes, but the response did not expand on whether users must agree to these changes to continue using the platform.

### Vimeo's Terms of Service prohibit hateful and terrorist content

- 9.18 Vimeo's Terms of Service state that users may not submit any content that "promotes or supports terror or hate groups". The [Vimeo Acceptable Use Community Guidelines](#) builds on this stating that content from hate or terror groups that aims "to spread propaganda designed to radicalise and recruit people or aid and abet attacks" is prohibited.
- 9.19 Vimeo's Community Guidelines also has a section that defines hate groups, hateful ideologies and terror groups. This section states that restricted users, such as terrorist organisations, and groups considered "banned" by governments and NGO designations of hate groups are also prohibited from the platform.
- 9.20 In 2021 Vimeo became a member of [Tech Against Terrorism](#), giving it access to resources and tools to support it in implementing policies to prevent terrorist use of its platform.<sup>57</sup>
- 9.21 The Acceptable Use Community Guidelines states that Vimeo "[does] not allow hateful and discriminatory speech", which is then defined as any expression directed to members based upon personal characteristics including race, gender identity, religion, disability, and age, and "conveys a message of inferiority or contempt". The Community Guidelines states that a negative expression is something that "would be considered extremely offensive to a reasonable person".
- 9.22 Examples are given of content generally considered categorical hate speech and this includes advocating or celebrating violence against an individual group based upon personal characteristics, spreading racial superiority theories or views, pickup-artist (PUA) content,<sup>58</sup> and conversion therapy relating to sexual orientation.

### Vimeo's terms and conditions are clearly communicated to users

- 9.23 Vimeo's terms and conditions are fairly user-friendly. They include a table of contents, sub-headings and bold text. Additionally, Vimeo's Acceptable Use Community Guidelines makes good use of expanding coloured boxes to assist user understanding.
- 9.24 Furthermore, it is easy to navigate to and between the terms and conditions and Community Guidelines.

## Watching content

### Users can rate content as mature when they upload it to Vimeo

- 9.25 Content can be rated as mature on Vimeo by content creators if it contains content that is allowed within the Terms of Service but not suitable for under-18s. This includes permissible nudity, violence, profanity, or illegal substances. This must comply with the

---

<sup>57</sup> [Announcing Tech Against Terrorism's Newest Members \(techagainstterrorism.org\)](#)

<sup>58</sup> Vimeo defines pick-up artist content in the [Vimeo Acceptable Use Community Guidelines](#) as content that markets, sells, or constitutes classes or tutorials that seek to teach seduction techniques.



terms and conditions and only allowed when the context justifies it. Users can find on Vimeo's Help Centre on the type of content that may qualify.<sup>59</sup>

- 9.26 Users with an account can use a content filtering tool to choose content that is rated as mature to not be shown.

### Vimeo makes some media literacy resources available to users

- 9.27 Vimeo provides some media literacy tools to help users better understand the content they are allowed to submit on the platform. These extend to an Acceptable Use policy, Community Guidelines and articles published on its Help Centre on topics such as determining the difference between pornography and artistic or non-sexual nudity; how Vimeo considers commercial content of an erotic nature; and how it defines hateful, harassing, defamatory, and discriminatory content.
- 9.28 Several features are in place to raise user awareness about the acceptable use of the platform. These include communicating changes to terms and conditions in a global banner displayed on login and sending emails to its users; the operation of a strike system against users who violate terms and conditions; a notification procedure for users found posting violating content; as well as notifying account holders of enforcement decisions by emailing the registered address.

## Detecting harmful content

### Users can flag content using the reporting tool, or by directly contacting the platform

- 9.29 Vimeo suggested that users can either flag content using a reporting function or by directly contacting the platform. Videos have a flag in the bottom right corner of the player section and accounts can be reported in the bottom left corner of the page.
- 9.30 Once a piece of content is flagged, it is placed in a queue for Vimeo's Trust & Safety agents to review whether it has violated the acceptable use policy. If the item is deemed to have violated Vimeo's guidelines it is removed, otherwise the content remains on the website and viewers will not be able to flag it again.
- 9.31 No system is in place to explain to the user flagging the content the outcome of the review.

## Enforcing against harmful content

### Sanctions are in place to deal with content that breaches the Terms of Service

- 9.32 Vimeo can apply the following sanctions:
- a) Removal of content;
  - b) Limiting access to a user's account or removing account privileges;
  - c) Account termination;

---

<sup>59</sup> [Content ratings \(vimeo.zendesk.com\)](https://www.vimeo.com/help/content-ratings)

- d) And in the case of Child Sexual Abuse Material (CSAM), users who upload this content are reported to the National Center for Missing & Exploited Children (NCMEC).
- 9.33 The nature of the content (i.e. whether it objectively violates Vimeo's Acceptable Use Policy) and the user's behaviour on Vimeo are considered when deciding whether to remove a piece of content and/or terminate an account for violations to Vimeo's Acceptable Use Policy.
- 9.34 Vimeo uses a "strike" system which typically gives a user a strike each time they violate Vimeo's Acceptable Use Policy. A user's account may be terminated after three strikes. In certain cases, accounts may be terminated from the first strike. The determination will depend on factors such as the egregiousness of the violation and the type of account in question.
- 9.35 Trust and safety agents have discretion about the nature of sanctions applied to different types of content in some cases. Depending on the egregiousness of the violation, agents may decide to remove an account in its entirety rather than removing individual video(s).
- 9.36 Once a user's account is terminated, they may not create another account. Vimeo also blocks the user's email from re-registering. Vimeo also stated that, in rare cases where it identifies a coordinated effort to upload a specific type of content, Vimeo might block an IP address. This is to prevent further spread of the videos containing harmful content and platform abuse.

#### **Hate and terror content is more likely to result in content removal**

- 9.37 Where a user uploads content that falls into one of the following categories, the account is (in most cases) banned for one strike:
- CSAM
  - Terrorist content
  - Violence or incites violence
  - Abusive, harassing, bullying, intimidation
  - Pornography, selling sex, sextortion, sexual stimulation
- 9.38 Accounts which seem to have been created expressly for the purpose of uploading infringing content may also be terminated from the first strike.
- 9.39 The team aims to review and make determination on 80% of flagged content within 24 hours. Videos flagged by users for containing terrorist content or hate speech may require a longer investigation. Terrorist or extremist content flagged by Active Fence, a third-party safety technology company, is taken down within 24 hours of Vimeo receiving notice.

#### **Moderators are trained to ensure consistency when applying sanctions**

- 9.40 Vimeo's Trust & Safety moderation agents are trained on its Acceptable Use Policy and the onboarding process to train all new moderators is a three-month process. There are several internal guidelines used in this process.

- 9.41 Vimeo also uses a Slack channel to let Trust and Safety agents communicate with each other and with other teams within Vimeo.

**Users are sometimes informed when sanctions are applied, depending on the nature of the content uploaded**

- 9.42 Enforcement decisions are communicated to the flagged account holder via email. Where Vimeo assigns a strike or terminates an account because it violates one of their content restrictions, the user will receive an email explaining the reasons for take down or termination.
- 9.43 However, a notification is not provided for CSAM, terrorist, fraud, spam, sextortion, and illegal content, so as not to alert the creator that their potentially criminal activity has been detected.

**Users can appeal against sanctions**

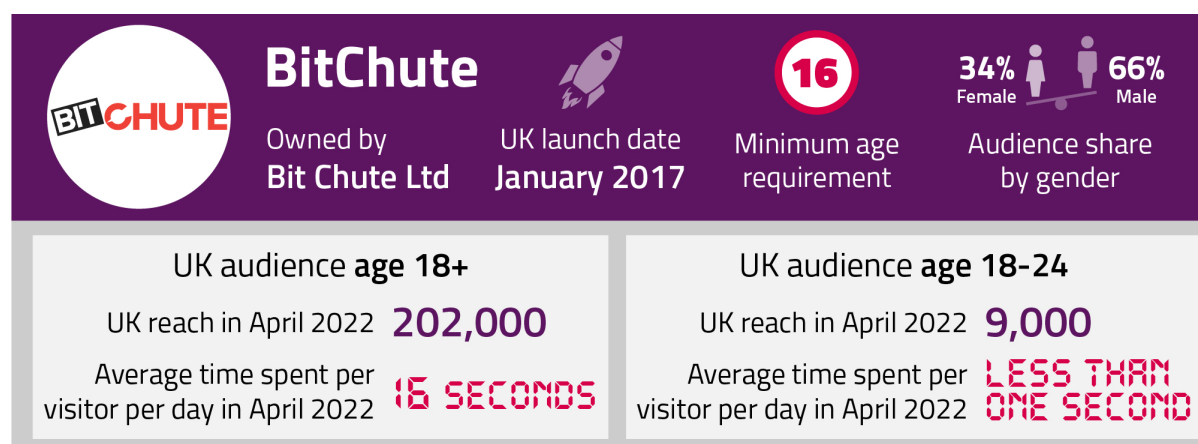
- 9.44 Information regarding appeals is located in the Community Guidelines. Users can fill in a form to submit an appeal request for moderators to reconsider their decisions, they must provide an explanation of why they believe there has been an error and Vimeo can ask for additional supporting information. Vimeo's Trust and Safety team then receives the appeal via a support ticketing system; can request further information from the user if needed, and the video will be re-reviewed by a different moderator to the one who applied the initial decision.
- 9.45 Vimeo can deny appeals made 30 days after the removal of content, and 60 days after the termination of an account. It can also deny appeal in the cases of extreme content such as CSAM but a user posting a video containing hate speech will just be assigned a strike, where the user's account is terminated after three strikes.

## 10. BitChute

### Introduction

BitChute positions itself as a “free speech” video-sharing platform that aims to empower its users to share their thoughts and opinions “without unjust criticism or discrimination”.<sup>60</sup> BitChute is often described as hosting alternative content, including far-right views and conspiracy theories.<sup>61</sup> BitChute’s userbase is skewed towards older adults. 45-55-year-olds were the most likely age demographic to visit BitChute in April 2022, making up roughly 27% of BitChute’s UK audience.

Figure 10.1: Key information on BitChute



Source: © Ipsos, Ipsos iris Online Audience Measurement Service, 1 - 30 April 2022, age 18+ for reach and time spent data, age 15+ for audience share by gender data, UK.

### Key findings

- 10.1 BitChute does not currently use automated detection measures to remove content, relying instead on human moderation and user reports.
- 10.2 **Viewers do not need to create an account on BitChute to watch video content.** This means that many viewers do not have to declare their age or read or agree to any terms and conditions. **An account is however needed to easily report videos containing harmful content.**
- 10.3 **BitChute reviews and removes content violating its terms and conditions when it is reported by third-party organisations like Tech Against Terrorism** in addition to reports from its own users. BitChute [became a member](#) of Tech Against Terrorism in October 2022, following its participation in the Tech Against Terrorism Mentorship Programme.

<sup>60</sup> [Our Commitment \(bitchute.com\)](#)

<sup>61</sup> [New breed of video sites thrives on misinformation and hate \(reuters.com\)](#); [News on Fringe Social Sites Draws Limited but Loyal Fans, Report Finds \(nytimes.com\)](#); [Fringe social media networks sidestep online content rules \(politico.eu\)](#)

## Engagement with Ofcom

- 10.4 Since the start of the VSP regime, BitChute has engaged constructively with Ofcom and has improved the measures in place on its platform in response to the new regulatory regime. This includes updating its terms and conditions and increasing the number of moderators it employs. BitChute also submitted its information request response before the deadline. Nevertheless, the number of users accessing BitChute is growing, and we expect the team will have to grow to meet the increased risk that comes along with a greater number of users.
- 10.5 Under the VSP Regime there is no requirement on platforms to proactively remove or monitor content. Therefore, the mere existence of harmful content on the platform would not mean that BitChute is failing to comply with the VSP Regime. However, we consider that the nature of BitChute's platform raises the risk that users may seek to use it to upload videos containing harmful content. We will therefore continue to engage with BitChute to ensure that it has measures in place to protect its users.
- 10.6 Through BitChute's engagement with third-party organisations such as Tech Against Terrorism, and as a result of our engagement with BitChute following the tragic shooting in Buffalo, NY, we have also been able to observe the platform's response to major events. BitChute has consistently removed content violating its terms and conditions when third-party organisations have reported it to the platform, including footage of the Buffalo attack. We are engaging with BitChute closely to drive improvements where needed.

## Governance and risk management

### Decision-making structure

#### Decisions about protection measures are made by BitChute's CEO

- 10.7 BitChute told us its Chief Executive Officer is responsible for making decisions about the implementation of specific protection measures.

#### BitChute has become a member of Tech Against Terrorism

- 10.8 On 6 October 2022, BitChute announced it had become a formal [member of Tech Against Terrorism](#). Ofcom understands that Tech Against Terrorism membership gives BitChute access to support and expertise to improve its efforts to make the platform safer. BitChute engaged with Tech Against Terrorism prior to becoming a member through the organisation's Mentorship Programme, which also resulted in BitChute [publishing its first annual Transparency Report](#).

## Assessing effectiveness

### BitChute measures effectiveness of a range of Schedule 15A measures on its platform




- 10.9 BitChute told us that it measures the effectiveness of communicating the changes to their policies through various metrics around its terms and conditions page: number of views, average time spent on the page, page likes, and page dislikes.
- 10.10 BitChute also told us it measures the effectiveness of its terms and conditions via the number of complaints received and monthly reports from their reporting and flagging mechanism which includes information on:
- d) Moderation action applied (total and per moderation reason)
  - e) Moderation reasons
  - f) Average resolution time
  - g) Number of tickets raised by each source and from whom.

### BitChute published its first Transparency Report in 2022





- 10.11 Following engagement with Tech Against Terrorism and Ofcom, BitChute published its first [Transparency Report](#) in June 2022, covering the previous 12 months. The report sets out BitChute's content moderation processes and publishes data on the number of user flags ('moderation requests'), enforcements against videos or channels, and user appeals.
- 10.12 BitChute has committed to publishing these reports at least once a year.

## User journey

Figure 10.2: BitChute User Journey

<p><b>User opens website or app</b></p>		<p>Without an account users can view most user generated videos, unless the content creator specifically added a sensitivity label or BitChute moderators deem it appropriate to change the sensitivity label after user-generated reporting and flagging.</p>
<p><b>User sign-up and log-in</b></p>		<p>BitChute does not have any specific age assurance measures or parental controls on its platform.</p> <p>It relies on users' self-declaration of being over 18 when they create an account.</p> <p>When users sign-up they must agree to BitChute's terms and conditions.</p> <p>BitChute updated its Community Guidelines following consultation with Ofcom and users can see online a published log of major changes in policies.</p>
<p><b>VSP recommends or user searches for content</b></p>		<p>Viewers can choose to limit the sensitivity of the videos that appear in their listings and searches.</p> <p>By default, content is rated as 'normal sensitivity', but users uploading content have tools to rate their content, either as not safe for work (NSFW) or not safe for life (NSFL).</p>

---

		BitChute moderators also change ratings to reflect the sensitivity of that content.
<b>User watches and engages with content</b>		BitChute terms of service prohibit illegal content but does not include prohibitions around all types of restricted material given the target age of the platform.
<b>User encounters harmful material and flags or reports</b>		A reporting tool is available for logged in users to report content that violates terms and conditions. Other users can report via email.
<b>User informed of outcomes</b>		BitChute's response did not suggest it keeps users informed about the outcome of moderation following user reports or flagging.
<b>User access to other tools (e.g. Media Literacy)</b>		BitChute publishes a transparency report.

---

## Signing on

### Users can create an account on BitChute

10.13 To create their account users must verify an email address and confirm understanding of, and agreement to, the platform's terms and conditions. Users with an account can view all videos available in their region. They can choose to limit the sensitivity of the videos that appear in their listings and searches.

10.14 Users without an account cannot view videos labeled NSFL,<sup>62</sup> or equivalent to BBFC 18.<sup>63</sup>

### BitChute does not have any age assurance measures or parental controls on their platform

10.15 When users register, they have to confirm they agree to terms and conditions that state they are at least 18 years old. BitChute's terms and conditions require that users without a registered account are above the age of 16, though the platform does not require users to read or accept these terms in order to use the website.

---

<sup>62</sup> Not Safe for Life

<sup>63</sup> The British Board of Film Classification (BBFC) sets guidelines for age ratings for films and other audiovisual content. [A rating of BBFC 18 means the content is only suitable for adults](#), meaning no one under the age of 18 can watch a BBFC 18 film in a cinema.

## Agreeing to terms and conditions

### Following a period of focused engagement with Ofcom, BitChute made changes to its terms and conditions and how it communicated these changes to users

- 10.16 Prior to the VSP regime coming into effect in November 2020, BitChute had limited measures to protect its users from harmful content, specifically from material that was “likely to incite hatred”, in its terms and conditions. The VSP Regime requires platforms to have terms and conditions to prohibit the upload of relevant harmful material and to require users to bring to the platforms’ attention the upload of restricted material.
- 10.17 Following structured engagement with Ofcom between April 2021 and September 2021, BitChute updated its terms and conditions to include ‘Incitement to Hatred’ in its prohibited content. Further changes to improve their policies (including their community guidelines) in relation to hate and terrorism content were made in June 2021.

### Ofcom have assessed the clarity of communication of BitChute’s terms and conditions

- 10.18 BitChute’s terms and conditions and Community Guidelines contain hyperlinks to each section (most detailed content restrictions are only mentioned in the Community Guidelines). BitChute’s Community Guidelines do not separately link different types of prohibited content. BitChute links to a separate incitement to hatred page on its website.
- 10.19 BitChute’s Community Guidelines use lots of sub-headings with brief explanations that are fairly easy to understand, although some links are to outside content which may be hard for users to engage with, such as links to pages from the legislation.gov.uk website.

## Uploading and watching content

### Users can rate the sensitivity of content when they upload it

- 10.20 BitChute’s Community Guidelines outline that it expects users to rate the sensitivity of videos they upload. These ratings range from normal, NSFW<sup>64</sup> and NSFL; the latter holds an 18 in the BBFC standards rating. By default, all content is marked as ‘normal sensitivity’ and is visible to all users.<sup>65</sup>
- 10.21 Users can select the sensitivity level that they are exposed to if they wish to avoid harmful content.

---

<sup>64</sup> Not Safe for Work

<sup>65</sup> [Choosing content sensitivity \(bitcoute.com\)](https://www.bitcoute.com/choosing-content-sensitivity)



## Detecting harmful content

### Logged-in users can flag harmful content for moderation using a flag icon

- 10.22 After clicking the flag icon, logged-in users are presented with a form to complete. This includes selecting a reason for the flag from a list of options that are directly related to the terms and conditions.
- 10.23 Users who are not logged into an account can email a moderation report to BitChute's reporting email address.
- 10.24 Users can also send complaints to the platform through a separate email address that is dealt with by BitChute's management rather than the moderation team.

### BitChute operates a 'priority flagger' programme

- 10.25 While BitChute operates a 'priority flagger' programme, they choose not to refer to this as trusted flagging as they do not immediately remove the flagged content as happens in some other trusted flagging programmes.
- 10.26 Priority flaggers can flag harmful content to BitChute, and these flags will be prioritised for review. The prioritisation is done within each type of violation, so that the most harmful content remains at the front of the queue for moderation. Organisations can contact BitChute to express interest in becoming priority flaggers.

## Enforcing against harmful content

### Sanctions are in place to deal with content that breaches the Terms of Service

- 10.27 BitChute uses a range of potential sanctions that may be applied. These are:
  - a) Blocking (removal of content)
  - b) Geo-blocking
  - c) Removal of Thumbnail
  - d) Modify Metadata
  - e) Modify Category
  - f) Modify Sensitivity
  - g) Account Deactivation.
- 10.28 BitChute shared its Moderation Action Matrix which is given to moderators to ensure the appropriate action is taken for a violation of the terms and conditions.
- 10.29 Geo-blocking is used to restrict content in specific regions and is only applied to content deemed 'incitement to hatred' and 'illegal in country'. A review is also conducted on the parent account to check if the account should also be geo-blocked

- 10.30 If the thumbnail of a video contains a harmful material, BitChute will remove it and replace it with a blank image. This can be applied to material classed as 'harassment' and is an approach not taken by any other notified VSP.
- 10.31 Moderators can modify the sensitivity label of a video. NSFL videos are only accessible to users who are logged-in to an account. Therefore if a moderator changes the sensitivity label of a video to NSFL, it will no longer be available to non-logged in users. If this is changed by a moderator, it cannot be altered by the creator. BitChute can also modify metadata attached to content, such as video titles, descriptions, tags or thumbnails. Unusually, BitChute does not apply any temporary restrictions, stating in its response that all sanctions that are applied are permanent unless successfully appealed. As BitChute only uses human moderators, the response suggests that moderators detail and record their justifications in the event of review or an appeal.

### BitChute adopts different approaches to terror content and hateful content

- 10.32 BitChute blocks content classified as 'terrorism and violent extremism'. BitChute applies geo-blocking to content it classifies as 'incitement to hatred', whereas content classified as 'harassment' may result in either being blocking, removing thumbnails, or modifying the video's metadata.
- 10.33 Geo-blocking prevents the content from being viewed in the UK, EEA, EU and the territories of those countries.

### Moderators are trained to ensure consistency when applying sanctions

- 10.34 BitChute moderators are trained to follow the same process to allow for consistency.
- 10.35 The process is as follows:
- a) **New Report Raised** – This can come from an external or internal flag. External flags can come from users, specialist organisations and state authorities.
  - b) **Priorities Assigned** – Reports are prioritised according to a number of factors. This can be based on who made the report, the reason for the report, and how recently the report was made.
  - c) **Material Reviewed** – A human moderator reviews reports in priority order.
  - d) **Moderation Applied** – If the content violates BitChute's terms and conditions, a sanction will be applied. These sanctions include setting an appropriate sensitivity or category, applying a geo-block, and blocking the material. For serious and repeated violations, accounts may be suspended.
  - e) **Stakeholders Notified** – The user that reported the content will be notified of the outcome of the moderation decision. If action is taken on the content, the user who uploaded it will be able to see this on their page but will not be proactively notified.

### Users can appeal against sanctions

10.36 Creators can appeal any sanctions that have been imposed. They can do this via email or through an online form. The appeals are reviewed by BitChute's moderation team.

# 11. Smaller VSPs

## Introduction

### Our approach to smaller VSPs

- 11.1 As set out in [Our approach to VSP regulation in Year 1](#), in considering our approach to VSP regulation in the first year of this regime, Ofcom have taken into account both the reach of each platform and the potential risks of harm to users.
- 11.2 In undertaking this assessment, Ofcom determined that five VSPs had relatively low reach and risk to users. While we still felt it necessary to seek information from these VSPs about how they protect their users from videos containing harmful material, we felt that a lighter touch analysis would be more appropriate and proportionate for these VSPs, especially in light of their more limited resources.
- 11.3 Ofcom consider those VSPs here, which we collectively refer to as “smaller VSPs”. These platforms are:
- a) Fruitlab
  - b) ReCast Sport
  - c) Qurio
  - d) Thomas Cook
  - e) The Sponsor Hub.
- 11.4 While some of the adult platforms are of a similar reach to the VSPs considered in this section, Ofcom felt that the distinct nature of adult platforms meant that the questions we asked VSPs and the focus of our assessment were materially different between smaller non-adult VSPs and smaller adult VSPs. We consider the latter in [Smaller Adult VSPs](#).
- 11.5 While of similar reach and risk, these five VSPs cover a range of different types of services. We set out a brief description of the nature of each of these VSPs below.

### Background to the smaller VSPs

- 11.6 **Qurio** is an app that allows users to upload videos reviewing products that they have purchased from QVC UK's shopping channels. Viewers are able to comment on these videos. Its userbase is relatively small and is made up of mostly QVC customers.
- 11.7 **Recast Sports** is a streaming platform for sports and entertainment that is still in a testing phase. It streams live and on-demand content and does not require users to purchase a subscription. Currently its creators are solely specially selected sports figures and organisations, although this is set to change in future. Viewers can earn or buy “cast credits” and redeem these in return for content. Creators can upload “Watch and Earn” videos which allow users to earn cast credits; these videos are labelled as advertisements on the platform. The platform earns 15% commission from creators' revenue.

- 11.8 The **Sponsor Hub** is a free platform for athletes to share videos of extreme sports, gain discounts from sports brands, and form partnerships with collaborating brands. It is open to professional and amateur athletes.<sup>66</sup>
- 11.9 **Fruitlab** is a “community platform” for those interested in video games.<sup>67</sup> Users can watch, upload and comment on content, and play games and challenges to earn “PIPs.” These are tokens which can be spent in the FruitLab shop on third-party digital vouchers or charity donations, or sent to other users and creators.<sup>68</sup> The platform receives a share of the revenue from advertisements.<sup>69</sup>
- 11.10 The **Thomas Cook** App allows users to share content related to travel and book holidays. Viewers can swipe through the ‘Explore’ page, or search for specific creators and places. There is also a live chat function for users to speak to customer services about their bookings.<sup>70</sup>

## Key findings

- 11.11 Ofcom found that the measures put in place by smaller VSPs tended to be less extensive and detailed than those of larger platforms.
- 11.12 Largely, however, we found that all the smaller providers had in place key terms and conditions in relation to videos containing harmful content, with some going into greater detail. For example, **Thomas Cook** mentioned mental health factors that might lead to harm as well as physical health factors.
- 11.13 The smaller VSPs all have some degree of reporting and flagging mechanism in place, although providers could improve these processes by ensuring users can choose between options behind the reason for content reporting, to the extent that they are not already doing so.
- 11.14 **Recast Sport** highlighted its practice of having different terms and conditions which apply to different types of users.

## Engagement with Ofcom

- 11.15 Early in 2022, we met with all the smaller VSPs to set out our programme and information gathering process, as well as to learn about them and their concerns.
- 11.16 We requested information from the smaller VSPs in summer 2022. Our findings in this section are largely based on those responses.

---

<sup>66</sup> [About Us \(thesponsorhub.com\)](https://thesponsorhub.com)

<sup>67</sup> [fruitlab.com](https://fruitlab.com)

<sup>68</sup> [PIPs \(fruitlab.com\)](https://fruitlab.com)

<sup>69</sup> [Terms of use \(fruitlab.com\)](https://fruitlab.com)

<sup>70</sup> [Mobile app \(thomascook.com\)](https://thomascook.com)

## Governance and risk management

### Decision-making structure

#### Decisions about protection measures on these platforms tend to be made by small teams

- 11.17 Smaller VSPs tended to have smaller teams involved in decision-making around what protection measures the platform will implement or it consulted with external firms to aid their decisions.
- 11.18 **The Sponsor Hub** told us that it is the co-founders who make a joint decision on the protection measures after consultation with an external legal team. **Qurio** told us that the measures in place were agreed at Executive level.

### Assessing effectiveness

#### Only The Sponsor Hub said that it assesses the effectiveness of its measures

- 11.19 Only one smaller VSP (**The Sponsor Hub**) told us it assesses the effectiveness of their platform's measures. The provider told us it does this through exploring their complaints and harmful material reports data. However, it also noted that it has received no complaints so far.
- 11.20 For other smaller VSPs it is either not clear whether they are measuring effectiveness of their measures, or they have confirmed they currently do not do so. **Recast Sports** did acknowledge that as the platform grows in size and risk of users encountering harmful material, it will need to review its process and have suggested metrics it might consider using to measure effectiveness of their protection measures.

## User journey

### Signing on

#### Users must sign-in to all platforms to fully interact with all content

- 11.21 On all of the VSPs, it is possible to create an account and on all of them an account is needed to create and upload video content. The type of content that can be viewed with or without an account varies from platform to platform.
- 11.22 Four of the platforms (**Fruitlab, Recast Sports, SponsorHub** and **Thomas Cook**) limit the content that users can engage with without an account. The extent to which users without an account can engage with videos varies by platform. For example, Recast Sports only allows viewers to watch short previews of videos without logging in, while Thomas Cook lets viewers who are not signed in watch content but not interact with it.
- 11.23 **Qurio** does not allow viewers to see any video content without an account.

### Fruitlab asks users to declare their age when registering for an account

- 11.24 As **Fruitlab** is a gaming VSP and therefore appeals to younger users, we asked the provider specifically about how it verifies the age of its users.
- 11.25 On **Fruitlab**, users must self-declare their age when registering for an account. Users under the age of 13 are not allowed to register an account. **Fruitlab** told us that its human moderators close the accounts if any users found to be under the minimum required age. Fruitlab did not disclose how it verifies information regarding the age of its users.
- 11.26 **Fruitlab** told us that harmful material is excluded from its curated video feed. Content with excessive violence and nudity is considered inappropriate for under-18s.

### The Sponsor Hub categorises its users by age

- 11.27 These age categories determine the requirements for users when signing up and the type of content they can see. The Sponsor Hub groups users under the age of 18 into three categories: 18-13, 12-10 and 9-6.

## Agreeing to terms and conditions

### Not all smaller VSPs present users with terms and conditions during sign up

- 11.28 Two of the smaller VSPs (**The Sponsor Hub** and **Thomas Cook**) told us that they present users with the sites' terms and conditions during the sign-up process
- 11.29 **Qurio** do not present users with terms and conditions during the sign-up process. On Qurio, these can be accessed by clicking through a user's Profile Settings (found under a three horizontal dot icon) under "Legal Terms & Privacy Settings" and then "Terms and Conditions."
- 11.30 **Recast Sports** told us it has three distinct terms and conditions for different types of users of the site: General Terms which apply to everyone visiting the site, User Terms for anyone with an account, and Verified User Terms for Verified Users which it explains are creators.

## Prohibitions on videos containing harmful content

### The smaller VSPs all prohibit the upload of different types of harmful content in their terms and conditions

- 11.31 Notified VSPs are required, if appropriate for their platform, to have terms and conditions prohibiting relevant harmful material.<sup>71</sup>
- 11.32 All smaller providers told us that they have some prohibitions in relation to relevant harmful material, but the wording of these prohibitions and what explicitly is prohibited

---

<sup>71</sup> Relevant harmful material refers to video content likely to incite violence or hatred against protected groups, and content which would be considered a criminal offence under laws relating to terrorism; child sexual abuse material; and racism and xenophobia.

varies by platform. One shared prohibition across platforms is that of racism and xenophobia.

- 11.33 All platforms also have some terms in place, either prohibiting content that incites violence or content that “promotes” it. Two of the smaller providers (**Thomas Cook** and **Fruitlab**) mention protected groups in this context.
- 11.34 Terrorist content is explicitly prohibited by the terms and conditions on **Thomas Cook**, **The Sponsor Hub**, and **Recast Sport**. Both **Qurio** and **Fruitlab** do not make explicit reference to the prohibition of terrorist content in their terms and conditions.
- 11.35 CSAM is explicitly prohibited by the terms and conditions of **Thomas Cook**, **The Sponsor Hub**, **Recast Sport**, and **Fruitlab**. **Qurio's** terms and conditions do not mention CSAM content, but do refer to nudity, pornography, sexual acts or provocative imagery in their list of prohibited content.
- 11.36 For material that might impair the physical, mental, or moral development of under-18s, all smaller VSPs prohibit sexually explicit or pornographic content in their terms and conditions as well as refer to at least one of the following categories of content: aggression, hate speech, violent material, dangerous behaviour, cyberbullying, online harassment, and cyberstalking.
- 11.37 Four of the smaller VSPs, (**Thomas Cook**, **Qurio**, **Fruitlab**, and **Recast Sports**) have terms and conditions which prohibit self-injurious content which may cause physical harms. Only Thomas Cook's terms and conditions mention mental health factors which might lead to a harm.

## Detecting videos containing harmful content

**All five smaller VSPs have reporting and flagging mechanisms in place.**

- 11.38 Of the five smaller VSPs, three (**Thomas Cook**, **Fruitlab**, and **Recast Sports**) provide categories which reporters can select from, although some of these categories are very widely defined, for example, “inappropriate for Fruitlab”. Recast Sports only provides four categories, one of which is “Other”.
- 11.39 **Qurio** and **The Sponsor Hub** indicated in their responses that their reporting categories were identical to their T&Cs. However, their apps do not appear to provide a list of categories when making a report. Instead, users press a single button to report an item without giving further details.



## 12. An introduction to adult VSPs

This section provides some context for the following two sections, which focus on the platforms that Ofcom collectively refers to as “adult VSPs”.

### Key terms

In the next two sections we use the following terms, defined here for reference:

**Age Assurance (AA):** a broad term that refers to the spectrum of methods that can be used for a platform to be informed about a user's age.

**Age Estimation (AE):** methods of age assurance that can estimate or infer a person's age. This may include, but is not limited to:

Biometric analysis, such as the analysis of facial features, fingerprints, and retinal patterns to estimate age;

Behavioural analysis, i.e. behaviour patterns of the user on the platform and their interaction with it to determine likely age;

Linguistic analysis, i.e. analysis of written language structure to evaluate age;

Profiling, such as using a user's past online activity or browsing history to evaluate certain aspects relating to the user.

**Age Verification (AV):** a form of age assurance where a user's age is established to the greatest degree of certainty practically achievable and can constitute the strictest form of access control. It is likely to rely on data sources that can secure a high level of confidence in the information provided (i.e., hard identifiers such as passport scans, credit details, driving license, etc.).

**Creator:** a platform user who can upload their own content to the platform. This often requires a different type of user account.

**Hashing technology:** the ‘hashing’ of CSAM images streamlines the process of detecting and removing illegal imagery. Hashing algorithms turn images into a unique series of numbers known as a ‘hash value’ - like a digital fingerprint. Platforms can use the *hash value* to speed up the search for duplicated content.

**Subscriber:** a platform user who pays to watch the content uploaded by creators. Some platforms use different terms to describe their subscribers, for example OnlyFans uses the term “Fans”.

### The adult VSPs in this report host sexual content that can be accessed through a subscription model

The platforms discussed in the following two sections are: OnlyFans; Admire Me; Fanzworld; PocketStars; Reveal Me; and Xpanded. We collectively refer to these as the adult VSPs.

These VSPs are all subscription-model services. This is a type of service where users sign up and subscribe to follow the accounts of specific creators, to gain access to exclusive content. These

platforms have two distinct types of account: a subscriber account where people sign up to view content, and a creator account where people sign up to create and upload content.

Creators on all notified adult VSPs can upload material of their choosing, though the platforms covered in this report are likely to be best known for hosting sexual content including pornography. Ofcom therefore refers to these as “adult VSPs”.

## **We expect access control measures to be in place where a VSP hosts pornographic material**

The VSP Framework requires platforms to apply the principle that restricted material that has the most potential to harm the physical, mental, or moral development of under-18s must be subject to the strictest access control measures.

In our [VSP guidance](#), we set out some of the indicators we may take into account when considering whether we believe a VSP should implement robust measures to prevent under-18s from accessing pornographic material. These include:

- How much pornography is on the platform;
- The significance of pornography to the service;
- The way the service is positioned on the market. This could be how it brands its own offering, or how the platform is viewed by users;
- Third-party insights which indicate the service specializes in pornography, or that there is a high risk of under-18s being able to access pornographic material on the platform.

Pornographic material is classed as restricted material for the purposes of the VSP Framework. In our Guidance, we stated that “if a VSP has restricted material on its service that is of a pornographic nature, providers should have a robust access control system that verifies age and prevents under-18s from accessing such material.” Such age verification measures were deemed to be a priority for VSP providers specialising in pornographic material.

We also describe several measures that we would not consider appropriate forms of age verification for this purpose, including:

- Self-declaration of date of birth or a ‘tick box’ system to confirm the user is over the age of 18 (‘self-declaration’);
- General disclaimers asserting that all users should be deemed to be over the age of 18;
- Relying on age verification through online payment methods which may not require a person to be over 18.

Laying foundations for age verification on adult VSPs was one of Ofcom’s priorities for the first full year of the VSP regime. It is therefore important that adult VSPs can demonstrate that they have implemented robust age assurance and verification measures and that they regularly test the effectiveness of the systems and processes they have in place.

We asked each of the notified adult VSPs questions about their age verification measures. Details of their responses to these questions are set out in the following two sections.

## Research commissioned by Ofcom shows that adults are broadly supportive of age verification measures

Ofcom has published the findings of research we commissioned which looks into adult internet users' attitudes towards age verification<sup>72</sup>. Participants:

- told us that they were broadly supportive of age verification measures to prevent under-18s from accessing pornography.
- **accepted** age verification measures where they were **expected**. For example, participants said that they accept the requirement to verify their age whilst purchasing alcohol online or participating in online gambling.
- expressed greater willingness to verify their age to access pornography if they were creating an account or subscribing to a creator to access content. In contexts when they were paying to access pornography, using a credit card was their preferred means of age verification.
- had serious concerns about how their data may be processed and/or stored whilst verifying their age to access pornography. This was reflective of a very **low level of trust** in the practices of adult sites.
- said that these concerns could be addressed by increased transparency about how their data would be used, stored, and deleted; a range of options of methods to verify their age; and potentially independent third-party providers performing the age check, rather than the adult sites themselves.

## There may be a heightened risk of adult VSPs hosting child sexual abuse material

Child sexual abuse material (CSAM)<sup>73</sup> is one of the most serious harms that the VSP Framework protects against, with a requirement to protect **all users** from it. Ensuring that platforms are working to keep this content off their services is a priority area of Ofcom's VSP work.

Research suggests that **youth-produced** or '**self-generated**' sexual material of under-18s (sometimes referred to as 'sexting', 'nudes', or 'nude selfies') is an increasingly significant driver of harm. We refer to this material as 'self-generated CSAM' in this report.

Of the 252,194 webpages that the Internet Watch Foundation (IWF)<sup>74</sup> actioned in 2021, it assessed almost three quarters (182,281 or 72%) as containing self-generated CSAM – this is an increase of 163% compared with the same figure for 2020. For the first half of 2022, the IWF reported a further 360% growth in the amount of self-generated CSAM of 7 to 10-year-old children compared to the same period in 2021. At 19,670 webpages, this is an increase of almost 8,000 pages.

---

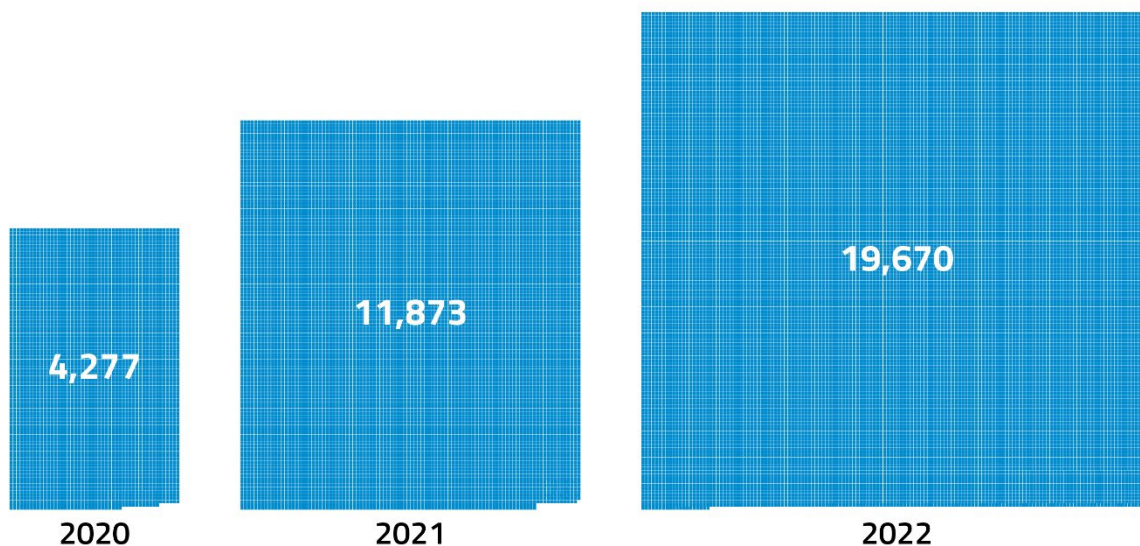
<sup>72</sup> [Adult Users' Attitudes to Age Verification on Adult Sites Research, 2022 \(ofcom.org.uk\)](https://www.ofcom.gov.uk/consult/condocs/adultusers/adultusers2022/adultusers2022.pdf)

<sup>73</sup> CSAM is classed as relevant harmful material and includes any material which shows the sexual abuse of a child. Under the VSP Framework the definition of CSAM covers the depiction of any person appearing to be a child, as well as realistic images of CSAM (such as computer-generated content) and simulated activity

<sup>74</sup>The IWF is a non-profit organisation that aims to remove online CSAM content through a combination of technology tools and human analysts.

The increase in self-generated CSAM within this particular age group might be indicative of online grooming; a harm that does not fall within the scope of the VSP regime, but which may be covered by the Online Safety Act.

**Self-generated CSAM imagery, 1 Jan-30 June 2022, 7–10-year-old children**



Source: [Internet Watch Foundation data release, Aug 2022 \(iwf.org.uk\)](https://www.internetwatchfoundation.org.uk/news-releases/internet-watch-foundation-data-release-aug-2022)

Adult VSPs may carry a heightened risk of under-18s uploading self-generated CSAM either through underage users successfully opening creator accounts despite a platform's policy, or through accounts held by adults posting CSAM they have obtained, or by under-age participants. We have therefore started our work to reduce the risk of CSAM by focusing on the protection measures that adult VSPs have in place.

## 13. OnlyFans

### Introduction

OnlyFans is a subscription-model video-sharing platform where subscribers (referred to by the platform as 'fans') can purchase and view content which has been created by other users (referred to as 'creators'). Creators upload a variety of types of content. Because OnlyFans has a minimum age requirement of 18 and is generally known for hosting sexual content, including pornography, Ofcom considers OnlyFans an adult VSP. OnlyFans also hosts a wide range of content that is not explicit.

Figure 12.1: Key information on OnlyFans



Source: © Ipsos, Ipsos iris Online Audience Measurement Service, 1 - 30 April 2022, age 18+ for reach and time spent data, age 15+ for audience share by gender data, UK.

### OnlyFans generates revenue through subscriptions, transactions and donations

13.1 OnlyFans collects 20% of consumer subscriptions and donation payments before passing on the balance to content creators. OnlyFans is less reliant on advertising than other platforms.

### OnlyFans' UK aged 15+ audience is largely younger and male, with some users exposed to videos containing harmful content

13.2 OnlyFans' userbase in the UK is predominantly younger and male. OnlyFans is particularly popular with young adults aged 18-34. In April 2022, over half (57%) of OnlyFans' total UK audience aged 15+ were aged 18-34, and 89% of visitors were male.<sup>75</sup>

13.3 Ipsos iris data suggests that people aged 15 – 17 have been navigating to OnlyFans' website or app.<sup>76</sup> The data does not provide insight into which parts of OnlyFans the users have been accessing. For example, it could be that users in this age group are only able to

<sup>75</sup> © Ipsos, Ipsos iris Online Audience Measurement Service, 1 - 30 April 2022, age: 15+, UK.

<sup>76</sup> © Ipsos, Ipsos iris Online Audience Measurement Service, 1 - 30 April 2022, age: 15-17, UK.

navigate to the log-in page before navigating away. Ofcom will be exploring this data further and discussing this with OnlyFans.

- 13.4 In our VSP Tracker survey conducted in March 2022, 15% of OnlyFans users said they had come across violent, abusive or inappropriate videos in the previous three months and only 17% of those that had been exposed said that they recall seeing safety measures in place.<sup>77</sup>

## Key findings

- 13.5 OnlyFans has put a broad range of measures in place to protect its users from harm on its platform, including those to detect harmful content and ensure users and creators are over-18. In particular, the platform has implemented age verification for all new UK subscribers, using third-party technological solutions provided by Yoti and Ondato. This comes after engagement with Ofcom on this issue.
- 13.6 Creators who post content on its platform must go through multiple steps to verify their age and identity. These appear to be relatively robust as presented to us by OnlyFans, though Ofcom has not been able to assess the effectiveness of its measures.
- 13.7 OnlyFans also uses several proactive moderation methods to detect harmful content in addition to its user reporting mechanisms. It uses hashing technology and its relationship with The National Center for Missing & Exploited Children (NCMEC)<sup>78</sup> to try to mitigate the risk of any CSAM it has detected from being spread further.
- 13.8 It has also taken steps to improve users' understanding of the risks posed by CSAM, how seriously it takes the challenges in this space, and its approach to tackling the problem.

## Engagement with Ofcom

- 13.9 OnlyFans' engagement with Ofcom during the first full year of regulation and during the information gathering processes was constructive. In the main, the information provided was detailed and supported Ofcom's understanding of how to address harms which manifest on platforms that feature significant amount of pornographic content. We have included information in this report which we think could help other adult VSPs to understand how they might mitigate such harms.

## Governance and risk management

### Decision-making structure

- 13.10 The following personnel make decisions in relation to OnlyFans' protection measures, with input from various team members:

---

<sup>77</sup> [Ofcom, VSP Tracker Wave 2 \(ofcom.org.uk\)](https://www.ofcom.gov.uk/consult/condocs/vsp/vsp-tracker-wave-2/) (March 2022).

<sup>78</sup> [The National Center for Missing & Exploited Children \(missingkids.org\)](https://www.missingkids.org/) is a private, non-profit corporation based in the USA whose mission is to help find missing children, reduce child sexual exploitation, and prevent child victimization.



- a) Chief Executive Officer;
  - b) Chief Strategy & Operations Officer;
  - c) General Counsel; VP & Deputy General Counsel;
  - d) Chief Technology Officer;
  - e) Head of Content Moderation; and
  - f) Head of Customer Service.
- 13.11 Decision makers are sent metrics monthly in order to inform regular assessment of the effectiveness of OnlyFans' measures. These include:
- a) the type of law enforcement requests received and the categories they fall into (e.g. child sexual abuse / exploitation, non-consensual sharing of intimate images, sexual assault, etc.)
  - b) the number of requests for assistance received and the methods by which they were received
  - c) metrics related to the processing time for enquiries to the compliance teams and customer support teams
  - d) a summary of notable enquiries
- 13.12 These metrics are likely to be of interest to Ofcom as we move into our second full year of VSP regulation. In particular, our priority is to assess the effectiveness of cross-cutting and harms-specific measures (see [Our strategic priorities in Year 2](#)).

## Risk assessment

### OnlyFans assesses risk through trend analysis and analysis of de-activated posts

- 13.13 OnlyFans has told Ofcom that it takes a dual approach to assessing risk on the platform. The first level is an internal risk assessment through trend analysis and analysis of de-activated posts to identify key risk areas. The second is a review by external lawyers and cybersecurity analysts who assess the effectiveness of the measures the platform has in place.
- 13.14 The only data that the platform collects on users relevant to a risk assessment is whether they are over the age of 18.
- 13.15 OnlyFans has also engaged a third-party independent monitor to assess and validate the design, implementation, and effectiveness of its safety compliance program, and to provide transparency. The report produced by this independent monitor examined:
- a) the risks posed by the underlying business model;
  - b) the perceptions of external stakeholders on the risks related to the platform;
  - c) key compliance obligations;

- d) the design effectiveness of policies, measures, etc. and whether these are "reasonably tailored" to mitigate the specific risks identified, and;
- e) the presence and effective implementation of corporate governance and compliance infrastructure with regard to mitigation of risks.

## Assessing effectiveness

### OnlyFans assesses the effectiveness of its creator and subscriber age verification and its measures to identify and remove CSAM

- 13.16 Creator age verification: as OnlyFans hosts predominantly adult content, it is important that it regularly reviews the effectiveness of its creator age assurance methods. The platform told Ofcom that these protection measures act to minimise the risk of under-18s uploading self-generated CSAM to the site.
- 13.17 OnlyFans says it uses several quality assurance measures. For example, reviewing any false positives<sup>79</sup> that occur during the age assurance process and testing the accuracy of its human content moderators.
- 13.18 Given the adult content popular amongst a large proportion of OnlyFans' userbase, it is also important to consider the efficacy of access control measures for subscribers. This is to ensure that under-18s are not able to access material which could impair their moral, physical, or mental development.
- 13.19 The VSP Framework accounts for the fact that technology evolves at pace. In our [VSP Guidance](#) Ofcom makes it clear that they should bear this in mind - regularly monitoring the effectiveness of their protection measures and evolving, improving, or changing them to continue to protect their users.
- 13.20 OnlyFans has reported that its senior management team conducts root cause analysis where measures fail, and uses its findings to evaluate whether additional changes, revisions, or enhancements are necessary. Further, OnlyFans says it regularly tests out new technologies to identify where to make improvements to existing processes. OnlyFans' reported that its Chief Technology Officer evaluates the effectiveness of these technologies. We support this level of senior engagement.
- 13.21 To reduce the risk of creators uploading CSAM to the platform, and of it remaining there if uploaded, OnlyFans regularly assesses the quality of its hash technology and hash list. It also uses data collected for its monthly transparency reports to conduct a full analysis of the measures it relies on to identify and remove CSAM from the platform. Additionally, when OnlyFans receives third-party complaints regarding potential CSAM it says it uses them to analyse whether they are symptomatic of 'gaps' in its other CSAM-related protection measures.

---

<sup>79</sup> A false positive refers to an error in which a test result, in this case the verdict of the age assurance mechanism, incorrectly indicates the presence of a condition (for example, the individual passes the age assurance check when they are in fact underage).



13.22 OnlyFans' also collects monthly data on measures such as law enforcement enquiries, the number of requests to remove copyright infringement content and the volume of account and posts deactivated and removed due to breaching their acceptable use policy. Further metrics are also shared with their decision makers to inform conversations on the effectiveness of OnlyFans' controls.

## User journey

### Subscriber user journey

Figure 12.2: OnlyFans Subscriber User Journey

<p><b>User opens website or app</b></p>		<p>Users can access a limited amount of content on the OnlyFans homepage without an account. According to Ofcom desk research, none of these videos are pornographic in nature.</p>
<p><b>User sign-up and log-in</b></p>		<p>OnlyFans accounts (subscriber or creator) are only intended for over 18s. To sign up for a subscriber account, users must provide an email address and an OnlyFans verified payment-method (credit or debit card). Users also need to agree to OnlyFans' terms and conditions.</p>
<p><b>VSP recommends or user searches for content</b></p>		<p>OnlyFans have adopted age verification measures with all new UK users, using technology provided by third-party providers. To sign up for an account as a subscriber, users must verify their age by completing a face scan, using Yoti's age estimation technology. If this fails, the user is re-directed to a backup age verification solution provided by Ondato, which requires a copy of an identity document to be uploaded.</p>
<p><b>User watches and engages with content</b></p>		<p>OnlyFans has very limited platform-wide search. Users must find and subscribe to specific creators to view content. Users can only subscribe to creators and view content once they have verified their age, a payment method is verified and the transaction is complete.</p>
<p><b>User encounters harmful material and flags or reports</b></p>		<p>Subscribers can only share videos with creators they subscribe to. This can be done via messaging, which is subject to moderation. OnlyFans uses technologies to scan for suspicious words and to prioritise for moderation by a human for all video content uploaded.</p>
<p><b>User encounters harmful material and flags or reports</b></p>		<p>OnlyFans allows all registered users to report harmful material and select a reason for reporting. Non-users can also report, e.g. accounts of concern.</p>
<p><b>User informed of outcomes</b></p>		<p>Users receive notifications of decisions only for unlawful/non-consensual content, not for complaints against content breaching acceptable use policies.</p>
<p><b>User access to other tools</b> <i>(e.g. Media Literacy)</i></p>		<p>OnlyFans has a Safety and Transparency centre which contains helpful guides on some of the safety measures the platform has in place. OnlyFans publishes monthly Transparency Reports.</p>

## Signing on as a subscriber

### Users cannot view video content without an account

- 13.23 OnlyFans told us that users cannot view any video content without an account and a subscription to a particular creator, using a payment card.
- 13.24 No pornographic content is available to view without first signing up and registering card details.

### OnlyFans' adoption of age-estimation technology for new UK subscribers

- 13.25 One form of age-assurance is the use of access controls that estimate the age of users, or 'age estimation'. OnlyFans only allows over-18s to create accounts on its platform and has said it has adopted this type of age assurance to prevent under-18s from creating subscriber, or 'Fan' accounts. This is intended to prevent under-18s from subscribing to any content, including pornographic content, on its platform. The technology is provided by Yoti, a third-party software company that uses facial age estimation technology, along with liveness, anti-spoofing and document authenticity checks to verify the age of potential users.<sup>80</sup>
- 13.26 The Yoti age estimation check forms part of the account registration process. A new UK subscriber scans a QR code that appears on their screen using their mobile device. This opens the software's age estimation portal on their phone. The user must first agree to a data handling notice before the platform displays instructions on how to complete the age estimation. The user must position their face within a frame that is displayed on their screen and take a selfie. An "Encrypted 3D FaceMap" is then uploaded, the user's age is estimated, and the selfie image is deleted. Ofcom understands that Yoti do not retain any of the user's data.

---

<sup>80</sup> "Liveness checks" are used to prove that a supposed user is actually present. For example, the user may be asked to move closer to the camera or rotate their head to prove that they are not presenting the software with a photo. "Anti-spoofing" refers to checks which use sophisticated technology to avoid a user circumventing or "spoofing" the age check through the use of masks, make-up etc.

### How accurate is Yoti's age estimation technology?

Yoti supply OnlyFans with age estimation technology to help OnlyFans prevent under-18s from creating accounts on their platform. While Yoti claims its age estimation technology is better than the average human at estimating an individual's age, no age assurance solution can be expected to be completely accurate. Yoti publishes regular White Papers, explaining how its technology works, its accuracy, and other information about the product and business. We have taken the information in this box from the latest White Paper.<sup>81</sup> As such these figures have not been independently verified by Ofcom.

#### Key terms

Two key measures Yoti uses to analyse its solution's accuracy rates are '**Mean Absolute Error**' (**MAE**) and '**True Positive Rate**' (**TPR**).

The **MAE** describes the average discrepancy between a user's estimated age and their real age. For example, if Yoti *overestimates* one user's age by three years, and *underestimates* another user's age by one year, the MAE of those two checks will be 2.

The **TPR** describes the probability that the technology has correctly allowed a user to pass an age threshold (i.e. to be over or under a set age). For example, if the age threshold is set to 23 years and nine out of a sample of ten users whose real age is under 23 are estimated by Yoti to be under 23, the TPR of that sample would be 90%.

#### Why does OnlyFans set its age limit to 23?

Yoti reported that its MAE is 1.52 years for 13- to 19-year-olds. This means that it could effectively estimate a user's age as being 2 years below or above their real age. For example, it may estimate a 16-year-old user as 18 years old (or estimate an 18-year-old as 16 years old). Therefore, the age threshold that users must pass to be able to create an OnlyFans subscriber account should be set higher than 18.

Yoti's latest White Paper reported that the TPR for correctly estimating 13–17-year-olds as being under 23 is 99.65%. If these results are accurate, by setting its Yoti age estimation 'pass' age as 23, OnlyFans is able to ensure that there is only a very small probability (0.35%) that under 18s will be able to pass through the check. However, it is likely that a number of users who are old enough to access the platform, but whose age is estimated to be under 23, will not be able to pass the age check.

- 13.27 If a user fails the Yoti age estimation process, they have the option of verifying their age using third-party Ondato.
- 13.28 Ondato's verification process analyses a potential user's identification document (i.e. passport, ID), then asks the user to record and upload a live video of themselves holding their ID. Ondato compares the image of the potential user's face to the image on the identification document to confirm that the ID belongs to the user. These checks are

---

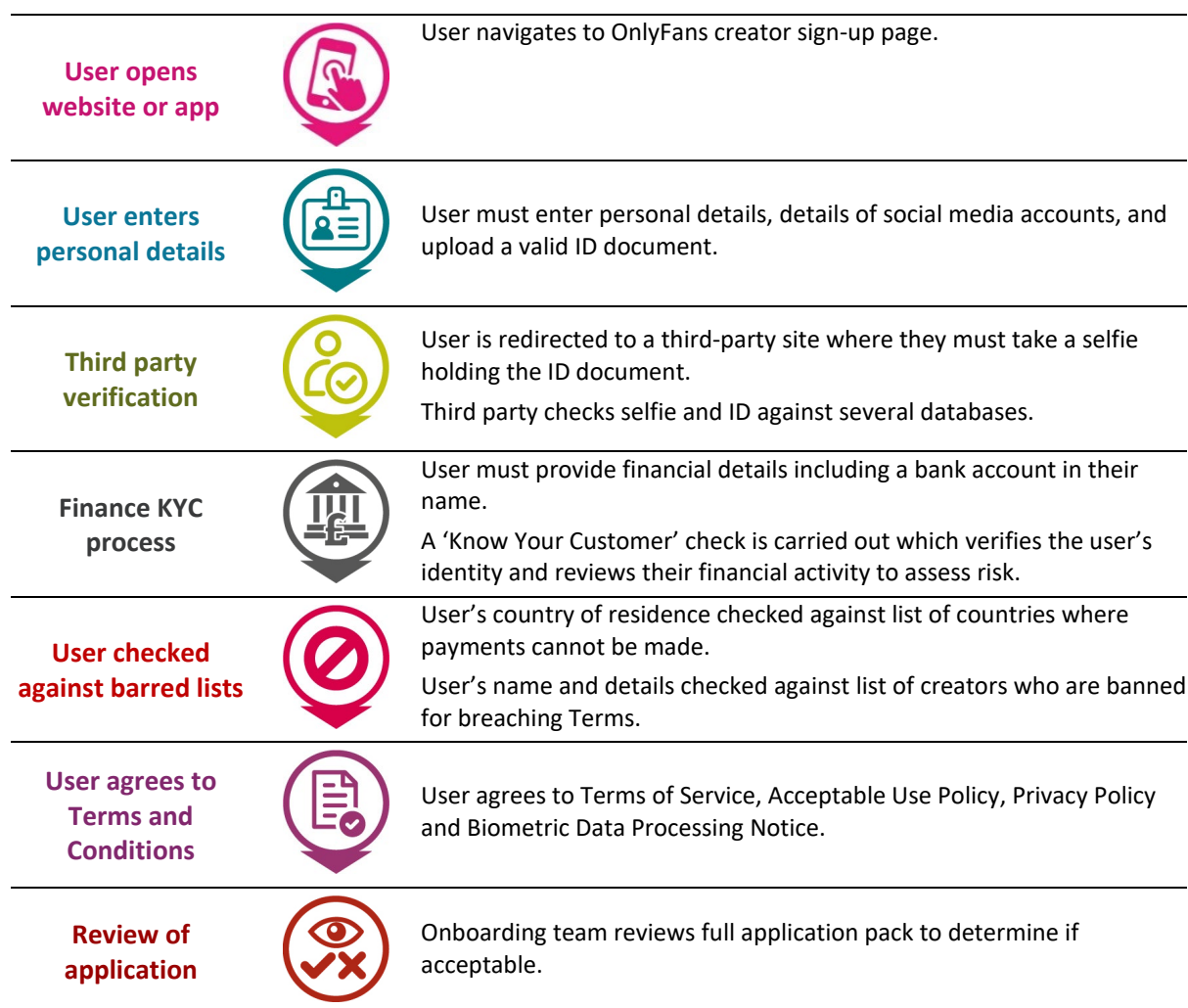
<sup>81</sup> [Yoti Age Estimation White Paper \(yoti.com\)](#), May 2022

followed by checks for sanctions (i.e. banned individuals), politically exposed persons, and adverse media screenings.

- 13.29 If a user refuses or is unable to prove their age using either the Yoti or Ondato solutions, they are unable to create a subscriber account. Where users attempting to create an account fail the Ondato age verification, their details are stored for future reference.

## Signing up as a creator

Figure 12.3: User journey to becoming a creator on OnlyFans



- 13.30 To address the elevated risk to under-18s associated with adult content creator accounts, OnlyFans has put added access measures in place during creator sign-up.

- 13.31 The age verification measures as described appear relatively robust, though no assessment of their effectiveness has been conducted by Ofcom. It is positive to see that OnlyFans has recognised the risk that creator sign-up represents and has spent time implementing a variety of ways to mitigate this.

- 13.32 It is important to note that age verification technology is evolving quickly as companies develop innovative new technologies, so standards are constantly rising. As a result, Ofcom's assessment of what constitutes 'robust age verification' will develop over time. In our guidance to VSP providers we explained that Ofcom expects providers to stay informed of emerging technological developments and solutions for online safety. We also ask providers to consider these developments as part of their ongoing assessment of the measures that are appropriate for their services. We expect OnlyFans and all other VSPs to continue to monitor and innovate in this space.

## Agreeing to terms and conditions

### CSAM Terms and conditions

- 13.33 OnlyFans' terms and conditions ban the upload of CSAM and make it clear that under-18s are not eligible to sign up for an account on the platform. The service also explicitly bans deepfakes<sup>82</sup> depicting under-18s, as well as emphasising the need to verify the age of anyone who appears in videos, not just the uploading party. This is in line with the definition of CSAM used in the VSP Framework and is something Ofcom would encourage all adult VSPs to consider adopting and enforcing on their services.
- 13.34 The inclusion of terms and conditions prohibiting the uploading of CSAM does not necessarily mean such terms and conditions are implemented effectively. One of Ofcom's priorities for the year ahead will be to understand how the effectiveness of terms and conditions is measured.

## Watching content and staying informed

### The ability to discover content through platform-wide searches is limited

- 13.35 OnlyFans has a very limited platform wide search function to limit the discoverability of content and requires subscribers to select the specific creators they wish to follow. What this means in practice is that subscribers cannot search for a general topic (i.e. boxing), but must instead search for a specific creator's OnlyFans account name ("handle"). Creators do not post content to be immediately found or viewed by all users. Users must subscribe to the creator's channel before being able to watch videos.
- 13.36 Subscribers cannot post video content to be viewed by general users. They can only share video content with the creators they subscribe to, using unencrypted direct messaging, which is also subject to content moderation.

---

<sup>82</sup> A video of a person in which their face or body has been digitally altered so that they appear to be someone else, typically used maliciously or to spread false information.

### Media Literacy Information relevant to guiding users' safe navigation of OnlyFans is provided off platform

- 13.37 OnlyFans provides most of the tools and information about how it operates, the nature of content on it, and the consequences of uploading violating materials on its off-site Safety and Transparency Center. This is where OnlyFans hosts its Terms of Use for Fans and Creators, Privacy Policy, Acceptable Use Policy, Referral Program Terms, Complaints Policy, and Community Guidelines.
- 13.38 The Safety and Transparency Center also conveys topic specific information on high-risk areas to users, for example a page on Combatting CSAM to inform users of its definition, how the platform addresses it, and what users can do to report suspected CSAM material on the site.
- 13.39 It is unclear from the information OnlyFans has provided how users can become aware of this material or be referred to it while using the platform. For instance, OnlyFans said it encourages users to report any potential violations of its Terms of Service and Acceptable Use Policy, but it did not provide further information on how it encourages users to do this.
- 13.40 OnlyFans reports taking steps to educate users who have posted content which constitutes a minor infraction on of its terms. In these circumstances, the platform shares information with the violating users on how to avoid repeating the infraction in either a push or pop-up notification on the platform.

## Detecting harmful content

### OnlyFans uses a mixture of human manual and machine-based review to detect harmful content

- 13.41 OnlyFans uses a variety of techniques to detect and prevent content which potentially violates its terms and conditions. This includes a mixture of extensive human manual reviewing and technology.
- 13.42 Before content appears on a newsfeed, it goes through an automated image screening and word recognition system that checks whether the content is allowed on the platform, prioritising high-risk material. All content that passes this initial review is then manually reviewed by trained human moderators within 24 hours. The moderators check content to:
- a) ensure that it complies with OnlyFans' Terms of Service and Acceptable Use Policies
  - b) ensure it complies with payment card association rules and regulations
  - c) identify any content which is illegal or which OnlyFans deems to be harmful
  - d) identify any content that may increase potential risks, including customer dissatisfaction complaints or chargebacks, as well as potential fraud.
- 13.43 OnlyFans has a reporting mechanism available to creators and subscribers. The provider states that it reviews social media feeds and has a dedicated monitored mailbox to ensure compliance with its terms and conditions.

### OnlyFans prioritises the reviews of potential CSAM but the user reporting process lacks clarity

- 13.44 OnlyFans told Ofcom that it takes an expedited approach to the review of potential CSAM once brought to its attention. The platform did not provide Ofcom with details on how it achieves this. It is worth noting that when users report content through the on-platform reporting mechanism, CSAM is not one of the 'reasons for reporting' options. It is therefore not clear how priority review of these reports is achieved in practice. Ofcom will be engaging with OnlyFans to learn more about how such content is expedited.
- 13.45 Once moved to review stage, OnlyFans says it aims to begin investigating CSAM reports within a few hours and to complete assessment of whether CSAM is present within 24 hours. Ofcom considers an aim to address CSAM content within 24 hours as a reasonable target for a platform of OnlyFans' size and nature.
- 13.46 OnlyFans did not provide data on how it performs against these time targets. Analysis of performance data can be a helpful way of understanding how effective a platform's protection measures are in practice and can provide useful insight into where improvements can be made.

## Enforcing against harmful content

### OnlyFans takes a wide range of measures if it finds content containing CSAM

- 13.47 If on review flagged content is confirmed to be CSAM, OnlyFans takes the following actions:
- a) Deleting the content;
  - b) Disabling and deleting the user account;
  - c) Taking steps to prevent a user from reopening an account;
  - d) Sending a report to the [NCMEC CyberTipline](#). According to OnlyFans' transparency reports, a report sent to NCMEC may include information identifying the user, the minor victim, and/or other helpful contextual facts to assist in protecting and safeguarding minors. It may be the case that more than one report is sent on a particular user or piece of content – for example, in cases where content is identified from multiple sources. It may also be the case where the same content is identified associated with a user with multiple accounts, we only report that matter once, per NCMEC guidelines and standard operation procedures.
  - e) Creating a hash value for any previously unreported content which it shares with NCMEC and other relevant bodies.
- 13.48 According to OnlyFans' [transparency reports](#), the platform reported 125 users to the NCMEC CyberTipline and contributed 32 hashes to NCMEC from January 2022 to September 2022.
- 13.49 Ofcom is encouraged to see the range of measures that OnlyFans has in place to tackle CSAM that is uploaded to its service. Its relationship with NCMEC and use of hashing

technology is particularly welcome, and we would recommend that all online services familiarise themselves with these safety mechanisms ahead of the Online Safety Act coming into force.



## 14. Smaller adult VSPs

### Introduction

As discussed in **An Introduction to Adult VSPs** Ofcom has grouped together platforms which specialise in content that is not appropriate for users under 18, including pornographic material, and refers to them as 'adult VSPs'. Adult platforms have a minimum age requirement of over 18 for users. In this report we refer to all adult platforms except OnlyFans as 'smaller adult VSPs'. They are as follows: AdmireMe; Fanzworld; Freyja; PocketStars; RevealMe; and Xpanded. Of these, only Freyja was not issued with an information request, due to the point at which the platform was launched. The other platform responses are reported on in this section.

### About the platforms

- 14.1 **AdmireMe's** creators (termed 'VIPs' on the platform) can share adult content either with paying subscribers or by selling one-off pieces of content to users not subscribed to them via their 'Premium Shop' if they have been verified on another social media platform. Creators can become verified with a blue tick by their profile if they have been verified on another social media platform. The platform takes 20% commission from creators' earnings.
- 14.2 **Fanzworld** enables its creators to upload photos, videos and blogs which subscribers can view for a monthly subscription fee. It also has a messaging platform to allow interactions between users and creators. The platform takes 17.5% commission from creators' earnings.<sup>83</sup>
- 14.3 **PocketStars** allows its creators to upload adult content to the platform and provides a messaging service for subscribers to contact creators. The platform's Feature Page features fifteen different randomly selected creators each day. PocketStars receives 20% commission on creators' revenue.<sup>84</sup>
- 14.4 **RevealMe** allows creators (termed 'Models' on the platform) to upload adult content to the platform for paying subscribers. It also offers private messaging, and pre-booked videocalls between subscribers and creators. Creators can pay to have their accounts featured in more prominent positions on the platform. Creators receive 100% of their earnings as the platform generates revenue through advertising and some add-on products.<sup>85</sup>
- 14.5 **Xpanded** differs from the other notified adult platforms in that it offers primarily live webcam services and telephone chat. The area of Xpanded which is notified as a VSP is the livestream content, which sits behind a paywall. 'Livestreaming' is when a creator sends

---

<sup>83</sup> Fanzworld, terms and conditions [last accessed 22 September 2022 16:39]

<sup>84</sup> Pocketstars, FAQ [last accessed 22 September 2022 11:41]

<sup>85</sup> RevealMe, FAQs [last accessed 22 September 2022 11:48]

video over the internet in real time, without first being recorded, edited, and stored. Xpanded says it usually has between ten and forty livestreams at any one time.

**Table 13.1: Key information on the smaller adult VSPs**

	Minimum age requirement	Livestreaming	Private messages	Subscription model
<b>AdmireMe</b>	18	No	Yes (paid for)	Yes
<b>Fanzworld</b>	18	No	Yes	Yes
<b>Pocketstars</b>	18	No	Yes	Yes
<b>RevealMe</b>	18	No	Yes	Yes
<b>Xpanded</b>	18	Yes	Yes	Yes

*Source: Ofcom desk research (including review of platform sites/apps, their terms & conditions, and web articles) and information provided to Ofcom by VSP providers in the Spring 2022 formal information request.*

## Key findings

### Subscriber age assurance measures do not appear to follow Ofcom's guidance

- 14.6 Most of the smaller adult VSPs appear to have age assurance measures that involve potential subscribers entering their own date of birth or ticking a box declaring themselves to be over 18 and agreeing to terms and conditions which state that all users must be over 18 years of age. We state in our guidance that we do not consider such measures to be appropriate forms of age verification for pornographic material. In particular, it should be noted that possession of a payment card alone does not guarantee that a user is over 18, as debit cards and certain pre-payment cards can be held by under-18s.
- 14.7 Some VSPs told us that as their pornographic content is behind a paywall, there is reduced risk of under-18s accessing pornographic content on their platforms.
- 14.8 In addition to the above age assurance measures, **Fanzworld** told us that its third-party billing provider uses procedures to verify the age of all potential subscribers. This includes ensuring that the payment card used belongs to the potential subscriber. Fanzworld's approach to age assurance sounds potentially promising, but we might continue to engage with the platform to gather more detail on how the process is implemented.
- 14.9 **Xpanded** told us that all adult content on its platform is accessed by dialing an 'adult premium rate number,' and it claims that this eliminates the risk of under-18s gaining access to inappropriate content.
- 14.10 One smaller adult VSP told us that it had considered implementing robust access control measures to prevent under-18s from accessing pornographic material, but had decided not to as they believed that adding further restrictions would impede adults from accessing the platform and reduce the profitability of the business.

## Smaller adult VSPs have put measures in place to protect users from CSAM

- 14.11 When considering measures that protect users from child sexual abuse material (CSAM), it is creator rather than subscriber age assurance that we need to think about. Creators can upload content to a platform and are therefore often subject to more stringent checks than those described above.
- 14.12 All the smaller adult VSPs that received our information request have age assurance measures in place for creator sign-up which, if working as described, Ofcom would consider to be robust. However, we have not conducted an assessment of effectiveness, nor have any of the platforms voluntarily provided data or evidence of such an exercise. As discussed in **Our strategic priorities in Year 2**, assessing the effectiveness of measures will be a priority for Ofcom over the next twelve months.
- 14.13 The smaller adult VSPs which provided information to Ofcom told us that no CSAM has been reported or flagged on their platforms. From the information provided it is not clear whether this is because the platforms are successfully preventing it from being uploaded or because their moderation methods are failing to detect it.
- 14.14 **AdmireMe, Xpanded, RevealMe** and **Fanzworld** have not closed any accounts on the grounds that a subscriber or creator was found to be under 18; **PocketStars** did not provide this data. Again, it is not clear whether this is because under-18s are effectively prevented from signing up or because platforms' systems and processes fail to spot it when they do.
- 14.15 With regards to reporting and flagging, only **AdmireMe** and **Xpanded** provided Ofcom with the information we asked for on these systems. Although these platforms prioritise reports of CSAM, neither has CSAM as a specific reporting category - raising questions about how and whether prioritisation is achieved in practice.
- 14.16 All respondents have sanctions in their terms and conditions that include removal of CSAM and suspension of the uploading account if this material is found on the platform. Some platforms have plans in place to report CSAM to authorities if discovered; both **AdmireMe** and **Xpanded** would report this content to law enforcement and Xpanded would also inform the IWF and NCMEC where applicable. The level of detail AdmireMe and Xpanded provided on how these mechanisms work in practice could be improved. **Fanzworld, RevealMe, and PocketStars** did not answer this question.

## Many smaller adult VSPs' responses to our information requests lacked detail, one VSP failed to respond within the appropriate time frame

- 14.17 While Ofcom has been encouraged by the smaller adult VSPs' engagement with the challenge of preventing CSAM from being uploaded to their platforms, many of the platforms' responses to our information request lacked specific details on how their protection measures worked in practice. This has made it harder to form a clear picture of how their CSAM protection measures are working to effectively protect users.

- 14.18 Similarly, some of these smaller adult VSPs did not provide detailed information on what access control measures they have implemented to protect under-18s from accessing pornographic content, how the effectiveness of those measures is assessed, or why they have taken the decision not to implement certain access control measures.
- 14.19 We engaged closely with each of the smaller adult VSPs to clarify their responses and request further detail. We will continue to engage with them to further understand the measures in place and how effective they are over the next 12 months.
- 14.20 **RevealMe**, a VSP provided by Tapnet Ltd, did not respond to our information request by the deadline set. On 29 September 2022 [Ofcom opened an investigation](#) into whether Tapnet Ltd failed to comply with its duties to comply with a statutory information request. Tapnet Ltd provided its response after the investigation was opened. This has impacted on our ability to comment on RevealMe's protection measures in this report.

## Engagement with Ofcom

- 14.21 Early in 2022, Ofcom met with all the smaller adult VSPs to set out our programme of work and information gathering process, as well as to learn about the platforms and answer any questions or concerns they had.
- 14.22 We requested information from the smaller adult VSPs in summer 2022, focusing on their age-verification measures and measures to prevent the uploading of CSAM. Our findings in this section are largely based on those responses. We launched an investigation into one smaller adult VSP, **RevealMe**, for its failure to comply with the statutory requirements around the information request.
- 14.23 **Xpanded**, who has experience engaging with regulators, was able to provide a higher quality response than other smaller adult VSPs. This demonstrates the value in Ofcom continuing to support smaller VSPs in preparation for regulation.
- 14.24 A VSP Provider decided to close two of their smaller adult VSPs after their business owner concluded that they would not be able to comply with Ofcom's expectations around age verification.

## Governance and risk management

### Decision-making structure

#### The decision-making processes within these small platforms appear to be relatively informal

- 14.25 Smaller adult VSPs provided us with limited details about their decision-making processes for the Schedule 15A measures they take on their platforms. **AdmireMe** takes quite an informal approach to decision-making regarding measures, with both employees and users able to suggest measures to be implemented on the platform which are then decided upon by the CEO. **Xpanded's** Managing Directors make the decisions on Schedule 15A measures following input and recommendations from the Technical Operations Manager and the

Finance Director. **PocketStars** told us its decision-making on measures happens through the management team.

### **We do not have enough detail to understand how the platforms assess practicability and proportionality when deciding which protection measures to take**

14.26 Under the VSP Framework, platforms need to consider the practicability and proportionality of measures when they decide which measures are appropriate for their platform to keep users safe.

14.27 We do not hold detailed information on how smaller adult platforms assess which measures are appropriate for their platforms.

### **More work needs to be done on assessing effectiveness**

14.28 We asked the smaller adult VSPs whether they test and measure the effectiveness of their Schedule 15A or creator age verification measures.

14.29 **Xpanded** told us it generates metrics on areas such as: flagged keywords, problematic content reports, breaches of terms and conditions and any account closures that result from these breaches.

14.30 **RevealMe** told us that it generates metrics on user reporting, keyword searches, CSAM reporting, payment activity, viewing hours and data transfer for photos or videos in order to assess the effectiveness of its Schedule 15A measures.

14.31 Collecting information on the number of content reports or subscriber/creator terms of service breaches is important as tracking this number over time can contribute to a platform's understanding of how well its moderation methods are working. It might also be useful to collect and track the reasons for the content being reported so platforms can see if any type of harmful video content is more prevalent on their platform.

14.32 The Directors of **Fanzworld's** parent company (Promo 1) monitor and moderate content on the platform. However, no metrics, research or reports have been provided to a decision-maker to aid in considering the effectiveness of its Schedule 15A measures.

14.33 Similarly, **PocketStars** and **AdmireMe** do not currently assess the effectiveness of access control measures and neither had effectiveness metrics in place in Year 1. PocketStars has recently begun to make use of Google Analytics; from its information request response it is not clear how it makes use of this data.

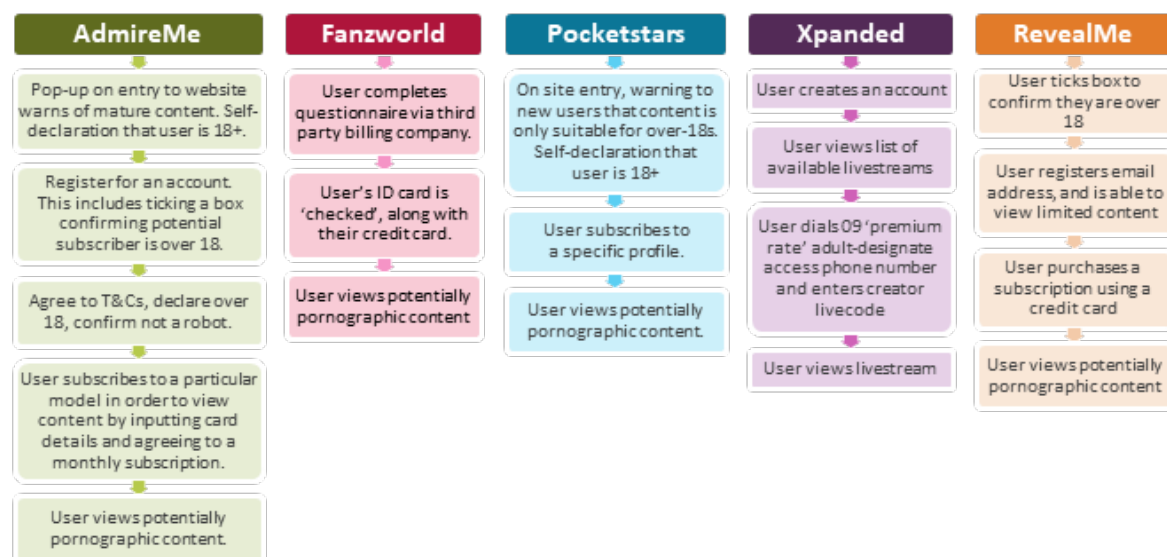
14.34 The responses showed that of the smaller adult VSPs only **Xpanded** is making attempts to assess the effectiveness of the Schedule 15A measures it has in place. Platforms seemed unsure how to test the effectiveness of their age assurance measures. Over the next year, we plan to engage with the adult VSPs to help them to monitor the effectiveness of their Schedule 15A measures moving forward.

## User journey

### Signing on as a subscriber

14.35 As mentioned at the beginning of this section, these smaller adult VSPs are generally subscription-model services. The user journeys of these platforms are broadly similar. Generally speaking, a user signs up for a profile by inputting their details. During this process, they may tick a box declaring they are at least 18 years of age or agree to terms and conditions which state that all users of the platform must be at least 18 years of age. The user then uploads payment details, before selecting or subscribing to a creator profile whose content they wish to view. This content is likely to be pornographic, though this is not always the case. The order of these steps may vary according to the VSP.

Figure 13.2: Subscriber sign-on process



Source: Smaller adult VSP responses to Ofcom Information Request, June 2022

### Xpanded's user journey involves calling an adult-designated premium phone rate number

14.36 The video-sharing section of **Xpanded's** service offers livestream webcam services. After creating an account, the user selects a livestream they wish to view. They then dial an adult-designated premium rate phone number, enter a code for the creator they wish to view, and stay on the line while watching the livestream. Premium rate phone numbers are currently regulated by the Phone-paid Services Authority (PSA) and allow services to charge their customers through their mobile bill. When the user wishes to stop viewing the livestream, they hang up the phone. Xpanded charge their users a minutely fee for these calls.

14.37 **Xpanded** claim that requiring users to access explicit content through a premium rate phone number specifically designated for Sexual Entertainment Services in the [National Telephone Numbering plan](#) prevents under-18s from accessing their service. The extent to which these premium-rate numbers are inaccessible to under-18s is not clear at this stage.

- 14.38 Creators on the platform agree to **Xpanded's** terms and acceptable use policy, which prohibit them from engaging with minors. Xpanded expects its creators to act as live moderators. If they suspect a viewer is underage, they must end the chat session and report the user to the moderation team. The moderation team also have access to the livestreams and Xpanded told us it may take action directly if it suspects a user is underage.
- 14.39 No information was provided about how effective these measures are at preventing under-18s from accessing pornographic material.

#### Fanzworld use a third-party bill provider which requires 2D identification

- 14.40 **Fanzworld** claim to have robust access control measures in place to prevent under-18s from accessing restricted material. This is achieved through the implementation of a third-party specialist billing provider that requires potential users to fill out a questionnaire. Fanzworld stated that there is a "quite rigorous" procedure for accepting card payments. This includes 2D identification requirements, which check the ID of both the card owner and the person using the card to make the payment. Fanzworld did not provide additional detail about how this measure works to effectively prevent under-18s from accessing restricted material on their platform, or how it tests the effectiveness of this measure.

#### RevealMe users can access some content after they have registered with just an email address

- 14.41 **RevealMe** told us it allows users to access a limited amount of content once they have registered with their email address, with further access provided once a subscription has been purchased with a credit card. Users can also send separate payments called 'Gems,' which can be purchased and exchanged between users.
- 14.42 The extent and nature of the content users can access without a subscription was not made clear. It is also unclear how Gems can be purchased, and which users are able to purchase, exchange or receive them.
- 14.43 We will be engaging further with **RevealMe** to better understand the extent to which pornographic material is freely available on their platform, and the range of payment options that are available.

#### AdmireMe and PocketStars require users to confirm they are over the age of 18

- 14.44 **AdmireMe** and **PocketStars** both claim to have measures in place intended to prevent under-18s from accessing pornographic material. These involve the potential subscriber agreeing to terms and conditions stating that all users must be over the age of 18 and ticking boxes confirming that they are over the age of 18.

#### Ofcom will further assess the protection measures on each of these smaller adult VSPs

- 14.45 Ofcom will be engaging with the smaller adult VSPs to better understand their access control measures, user journeys, and the risk of harm to under-18s. Where we have concerns that age assurance measures are not effectively protecting under-18s from



restricted material on particular platforms, we will carefully consider whether enforcement action, or some other action, would be appropriate.

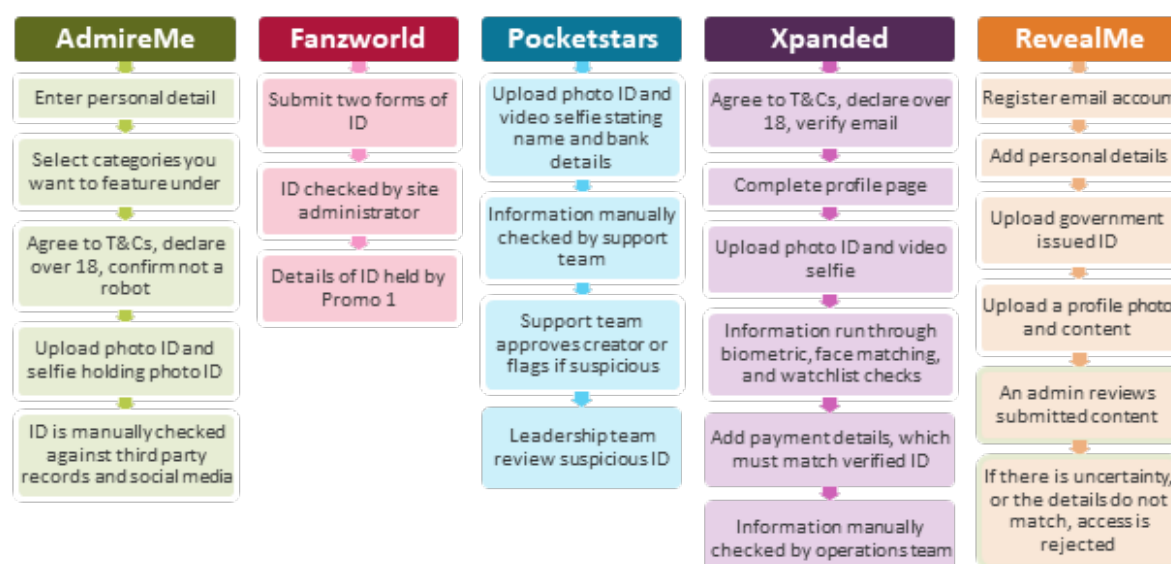
## Signing-on as a Creator

### Age verification systems in place for Creator accounts

14.46 Creator sign-up represents a point of heightened CSAM risk on the adult platform user journey. Creators are the only users who can upload content, so if an under-18 can open a creator account on an adult platform the risk of self-generated CSAM being uploaded increases. It is therefore essential that VSPs are enforcing the terms and conditions that ban under-18s from signing up as creators effectively.

14.47 Smaller adult VSPs' creator-sign up processes are summarised below:

**Figure 13.3: Creator sign-on process**



Source: Smaller adult VSP responses to Ofcom's information request, June 2022

14.48 Overall, the picture for creator age verification across the smaller adult VSPs is more positive than it is for subscriber age verification, with all responding platforms putting added access control measures in place for creators.

14.49 The more stringent age verification measures deployed to prevent self-generated CSAM reflect not only the greater potential risk of harm, but also the fact that the number of age checks required for creators is far lower than the number of age checks required for subscribers. This enables adult platforms to perform age verification for creators manually, with staff members checking each ID themselves and ensuring that it corresponds to other information provided by the potential creator. While rigorous, this form of age verification would be burdensome for VSPs to implement for subscribers as well as creators, as a far greater number of staff would be needed to manually check each potential subscriber's age. Additionally, users are likely to feel that human-based review infringes on their privacy; some of the participants in our research into attitudes towards age verification



told us that the potential impact of age verification on their privacy was a key concern of theirs.<sup>86</sup>

- 14.50 It is important to note that age verification technology is evolving quickly, and so too is the consensus of 'what good looks like' in this space. As a result, Ofcom's assessment of what constitutes robust age verification will develop over time. In our [VSP Guidance](#) we explained that we expect VSPs to stay informed of emerging technological developments and solutions for online safety, and to consider these as part of their ongoing assessment of the measures that are appropriate for their platforms.
- 14.51 Looking ahead to Year 2 of the VSP Regime and further to the Online Safety Act, Ofcom will be paying close attention to the evolution of age verification methods and platforms' efforts to keep up with them as we work to raise standards in this area.

## Agreeing to terms and conditions

### Most smaller adult VSPs have adequate wording in place to ban the uploading of CSAM

- 14.52 **Xpanded** goes further than the others, banning videos of adults role-playing as under-18s, which is encouraging. Going forward we would like to see more explicit prohibition of simulated CSAM and deepfakes on all adult platforms, as well as tighter rules around age verification of anyone appearing in a video uploaded to the platform – not just the person uploading.
- 14.53 We see the inclusion of terms and conditions banning relevant harmful material as a positive step. However, because (with the exception of **Xpanded**) these VSPs do not monitor the effectiveness of the protection measures they put in place, we are unable to confirm that the terms and conditions they create are working to protect users from problematic content.
- 14.54 It is vital that VSPs not only include terms and conditions in their user agreements, but also implement them effectively in practice to protect users on their platforms. This is something Ofcom will be working to gain a clearer picture of in Year 2.

## Media literacy

### Some of the platforms have taken steps to improve the media literacy skills of their users

- 14.55 Ofcom did not request information on the media literacy measures that smaller adult VSPs have in place. Nonetheless, the information we received from their responses indicate that at least two platforms, **AdmireMe** and **Xpanded**, have taken steps to improve user's media literacy and to raise awareness about the media literacy tools available on their platforms.
- 14.56 **Xpanded** promotes video explainers, tooltips, and blogposts to its users on subjects such as avoiding excessive network charges while using Phone-paid Services. Additionally, it has published a plain English summary of its terms and conditions to improve users'

---

<sup>86</sup> [Ofcom's Adult Users' Attitudes to Age Verification on Adult Sites Research, 2022 \(ofcom.org.uk\)](#)

understanding of permissible content as well as tools such as features for users to flag or report content on each page of the site.

- 14.57 **AdmireMe** has taken steps to raise user awareness about platform operations. For example, once a user reports content, they are shown a pop-up notification informing them of the steps the platform will take to review the content and that they may be contacted for further information.

## Detecting harmful video content

### Users can report harmful video content for review by moderators on PocketStars, Xpanded and AdmireMe

- 14.58 Content on **PocketStars**, **Xpanded** and **AdmireMe** can be reported or flagged for review by moderators, and all three platforms say they treat content flagged as CSAM as a priority. **RevealMe** told us that it has a "user-to-user" reporting mechanism but it is not clear from its response what category of user is able to report in this way (i.e., whether subscribers, creators or both can report content that concerns them). Fanzworld did not respond to the questions we put to it on this point; something we are engaging on further.
- 14.59 While Ofcom is encouraged to see that **PocketStars**, **Xpanded** and **AdmireMe** have taken steps to ensure that users can bring problematic content to moderators' attention, none of the platforms provided enough information to demonstrate how these measures work in practice. This means that so far, we have been unable to assess how effective these measures are. We will address this in Year 2 – focusing particularly on how content flagged as CSAM is categorised and prioritised for review.

### All notified smaller adult platforms have some form of proactive moderation in place

- 14.60 **AdmireMe** told Ofcom that it reviews all content that is uploaded and carries out "random daily checks"; **Pocketstars** says its customer service team regularly reviews content; **RevealMe** told us that it uses CSAM media scanning provided by Cloudflare, and that site administrators routinely check uploaded content; and **Fanzworld** has moderators on its platform.<sup>87</sup>
- 14.61 **Xpanded** takes a different approach to moderation because, unlike the other smaller adult VSPs featured in this report, it is a livestreaming service.
- 14.62 **Xpanded** said its moderators review livestreams as they happen and can review past livestreams for up to ten days after streaming if required. The platform expects its creators to take some responsibility for moderation by identifying potentially underage subscribers and reporting them to moderators.
- 14.63 Alongside human moderation, **Xpanded** uses software from a third-party provider - Hive - to automatically moderate creators' profiles and identify inappropriate content, including

---

<sup>87</sup> Cloudflare says its CSAM scanning tool allows website owners to proactively identify and take action on CSAM located on their website, see: [Announcing the CSAM Scanning Tool, Free for All Cloudflare Customers \(cloudflare.com\)](https://www.cloudflare.com/announcements/csam-scanning-tool-free-for-all-cloudflare-customers)

the presence of under-18s. If detected, this content is passed over to human moderators for further review. There is also an automated system to check text against the platform's list of prohibited words and phrases; if problematic content is found it will not be uploaded to the platform and will be flagged for review by human moderators.

- 14.64 Ofcom welcomes the fact that five smaller adult VSPs have taken steps to introduce some form of proactive content moderation to protect users. **Xpanded's** adoption of a variety of moderation methods is particularly encouraging; it acknowledges the need to keep up with the ever-changing nature of online risks and their mitigations.

#### Enforcing against harmful video content

- 14.65 As part of our work to understand how effectively smaller adult VSPs are enforcing their terms and conditions, we asked them to tell us how many videos containing CSAM had been detected on their platforms; and how many accounts they have closed because a creator was discovered to be under 18 in a 12-month period.
- 14.66 As mentioned previously, **RevealMe**, **Xpanded**, **AdmireMe**, and **Fanzworld** all stated that they have never identified CSAM or an underage account on their platforms. **PocketStars** has never identified CSAM on its service but did not provide underage account data.
- 14.67 Because all of the smaller adult VSPs' responses lacked adequate detail, we are unable to comment definitively on whether CSAM and underage creator accounts have not been detected because the protection measures in place are working effectively to prevent this, or because their detection systems are failing to pick it up. Clarification of this point will form one of the strands of our work in Year 2.