

## Your response

Please refer to the sub-questions or prompts in the [Annex](#) of our Call for Evidence.

Question	Your response
<p><b>Question 1: Please provide a description introducing your organisation, service or interest in Online Safety.</b></p>	<p>Confidential? <b>N</b></p> <p>Roblox was founded in 2004 and our mission is to provide a platform for shared experiences among people of all ages. Developers from around the world create experiences on our platform, and we provide the infrastructure to host those experiences safely. While Roblox is often associated with “gaming,” there are thousands of other skill-building and educational experiences available, such as learning how to run businesses, visiting historical landmarks, or viewing art exhibitions.</p> <p>As recently as this month, child safety experts observed in the New York Times that “as other sites try to broaden their appeal to children...they could learn important lessons from Roblox.” Experts have further noted that Roblox experiences “reflect the sort of imaginative play you often find in the playground.” Notably, during the COVID-19 pandemic, Roblox has played a key role for families by enabling children to “socialise” online while remaining physically distant. In a survey, a quarter of teens also said that Roblox provided an outlet to discuss difficult topics, such as their struggles with pandemic isolation.</p> <p>More broadly, there is increasing awareness about the role that videogaming plays in promoting interest in science, technology, engineering, and maths (STEM), and its strong influence in attracting girls and young women to careers in computer science. Roblox offers free STEM skill development for people of all ages via “Roblox Studio,” and provides free curricula for teachers to use in classrooms via our “Roblox Education” platform. Earlier this year, Roblox’s use as an educational tool was described as “the future of education.”</p> <p>Roblox has always maintained a firm and demonstrated commitment to children’s safety. We use a combination of technically advanced tools including both industry standard systems such as PhotoDNA as well as our own internally developed systems to both proactively and reactively detect and mitigate inappropriate content. We maintain a large, highly trained global moderation team. Our Safety Advisory Board is comprised of leading experts in children’s online safety as well as a paediatrician who specialises in children’s media consumption as it relates to their cognitive and behavioural health. This helps advance our goal of promoting children's well-being, on Roblox or off.</p>

<p><b>Question 2: Can you provide any evidence relating to the presence or quantity of illegal content on user-to-user and search services?</b></p>	<p>Confidential? <b>Y</b></p>
<p><b>Question 3: How do you currently assess the risk of harm to individuals in the UK from illegal content presented by your service?</b></p>	<p>Confidential? <b>N</b></p> <p>We currently have two major internal programs in place for risk assessment and risk mitigation:</p> <ul style="list-style-type: none"> <li>● <b>Trust by Design ("TbD"):</b> TbD is an internal program that involves cross-functional teams that engage with our product development and engineering teams during their planning phase to identify risks and offer recommendations on gap mitigation. The product and engineering teams complete an intake form that addresses questions from stakeholders throughout the company. Stakeholders within the TbD team include Information Security, Content and Product Policy, Public Policy, Accessibility, Compliance, Privacy and Legal. TbD provides ongoing consultations pre and post launch as needed to support the business.</li> <li>● <b>The Moderation Policy Advisory Council (IMPACT):</b> IMPACT is a mechanism through which Roblox can conduct internal conversations surrounding high impact/visibility moderation policies. Trust &amp; Safety Moderation and Policy teams require input from stakeholders (PR, Legal, Developer Relations, product, etc.) and senior leaders in order to provide rapid response to changing community needs. IMPACT is the venue for those inputs.</li> </ul>
<p><b>Question 4: What are your governance, accountability and decision-making structures for user and platform safety?</b></p>	<p>Confidential? <b>N</b></p> <p>At Roblox we conduct a comprehensive review of safety, cybersecurity and legal risks called Trust by Design. Trust by Design is a systematic approach that simplifies the process for building safe and trusted products and features on Roblox. We aim to provide a one-stop shop for Product and Engineering teams by offering consultations from all trust related perspectives. The process involves the following:</p> <ul style="list-style-type: none"> <li>● Identify potential risks: We offer a holistic product risk assessment process, with the goal of identifying and evaluating risks from different trust perspectives during the product conception or design stage;</li> <li>● Recommend mitigation strategies: Based on the risk level of each identified risk, we provide recommendations around prevention and mitigation;</li> </ul>

- Provide ongoing consultation: We are here to support the business by providing ongoing consultation on product implementation and continuing to evaluate the trust level of product/feature. The 'Trust by Design' (TbD) team works with product owners to identify risks and offer ideas of potential solutions. We also provide ongoing consultations pre and post launch as needed to support the business.

We have formed a virtual team, the Trust by Design Committee, that consists of representatives from Legal, Privacy, Safety, and Security, to ensure that we review the product and provide recommendations that take all aspects of trust into consideration. This team will involve additional functions such as Public Policy, Operations, tooling, Comms, and Developer Relations, etc. depending on the product needs.

An TbD assessment mainly considers 4 pillars of trust factors: Legal, Privacy, Security, and Safety. Some examples of what each pillar will touch upon include:

Legal: International laws/regulations; COPPA; IP; contract; consent; data collection, usage, storage and retention; privacy settings; parental controls; GDPR; accessibility

Safety: UGC policy; user behaviour; age appropriate design; account protection; support platform; operations

Security: User data protection; company data protection; account security; scraping; API security

#### Safety Risk Review

As part of the Trust by Design program, the Trust & Safety product and policy teams will conduct a deep dive assessment of every product feature for potential safety risks. Relevant Safety team stakeholders will participate in the assessment, including Safety product managers, Product and Content Policy members, and Content Moderation leaders. The initial assessment will include:

- A review of findings from the product interview
- Abuse brainstorming in the form of User Stories
- Safety assessment based on product-as-planned

Following this initial assessment, the team will:

- Collaborate to outline recommendations for risk mitigation in each of the abuse scenarios for the PMs as an "Ideal Safety Scenario"
- Provide a safety assessment based on the product with all mitigations incorporated

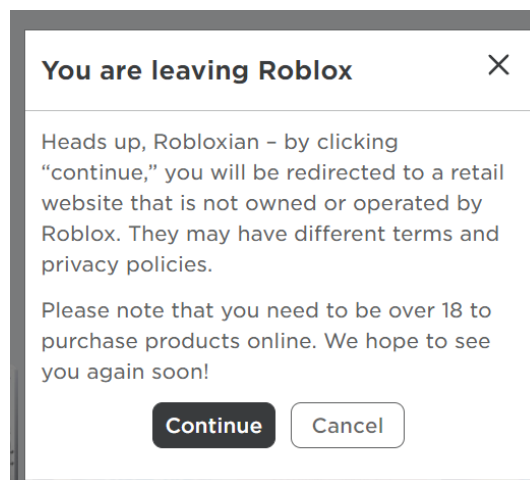
The Safety team will then present findings and recommendations to the Product and/or Engineering Manager and will follow-up as needed. The scorecard will be completed once product managers follow-up with an updated product spec, identifying which - if any - mitigations have been added.

**Question 5: What can providers of online services do to enhance the clarity and accessibility of terms of service and public policy statements?**

Confidential? **N**

Roblox has updated its Terms of Use, Community Standards and Privacy Policy and created [FAQs](#) to help children and their parents better understand our practices. In addition, we are engaging third parties to assist with our endeavour to make our Terms of Use, Privacy Policy, and Community Standards more understandable by the relevant age range of users. The consulting firm we are engaging with will help train our team on developing more child appropriate policies and update not only our privacy information, but also our terms to be more readable for our users.

In addition to our updated policies, which summarise each section in a clear, easy to understand manner, Roblox also provides additional notices to communicate information to children. For instance, Roblox uses just-in-time notices to notify users when they are leaving our platform for a third party site.



Roblox also requires additional consent, or opt-in for our new features like voice chat.

We are working toward making our platform compliant with Web Content Accessibility Guidelines ([WCAG 2.1](#)), which is a US industry standard for web accessibility. Please see our [Accessibility Statement](#) for more information.

**Question 6: How do your terms of service or public policy statements treat illegal content? How are these terms of service maintained and how much resource is dedicated to this?**

Confidential? **N**

To ensure that everyone feels welcome and safe and is treated with kindness and respect, we've developed Community Standards that outline clear expectations for using Roblox. If anyone is found to be in violation of these standards, they may be suspended or removed from the platform. We update our Community Standards periodically as we introduce new product features and as the needs of our community change.

Our Community Standards set clear expectations for how to behave on Roblox. They require that everyone treat each other with civility and respect and help keep everyone safe by prohibiting content and behaviour that may be inappropriate, illegal or harmful.

	<p>We have also established clear guidelines and expectations around the behaviours of our creator community. We also prohibit creators from directing users off-platform (with some exceptions), nor do we permit contests, lotteries or sweepstakes that involve free Robux.</p> <p>Our Trust and Safety team uses these Standards in order to moderate content and respond to abuse reports from users. If any users are found to be violating these standards, they may be suspended or removed from the platform. In some cases, Roblox also works proactively with authorities to report cases of violent threats, child endangerment, or other serious real world harm.</p> <p>Our Community Standards explicitly prohibit “discussing, depicting or promoting illegal or highly-regulated activities or encouraging others to violate their local laws [...], including gambling content.” The Standards are always evolving, maintained by the Policy team, and are updated when necessary to comply with local laws or other changes in our treatment of content on the site.</p>
<p><b>Question 7: What can providers of online services do to enhance the transparency, accessibility, ease of use and users’ awareness of their reporting and complaints mechanisms?</b></p>	<p>Confidential? <b>N</b></p> <ul style="list-style-type: none"> <li>● Just in time notices at the time of or just prior to a new form of collection or processing of personal information</li> <li>● Clear, readable and accessible policies and terms</li> <li>● Explicit reporting mechanisms, supported by help articles for assistance</li> <li>● Follow industry guidelines or standards such as WCAG 2.1</li> </ul>
<p><b>Question 8: If your service has <u>reporting or flagging</u> mechanisms in place for illegal content, or users who post illegal content, how are these processes designed and maintained?</b></p>	<p>Confidential? <b>Y</b></p>
<p><b>Question 9: If your service has a <u>complaints</u> mechanism in place, how are these processes designed and maintained?</b></p>	<p>Confidential? <b>Y</b></p>

<p><b>Question 10: What action does your service take in response to <u>reports</u> or <u>complaints</u>?</b></p>	<p>Confidential? <b>Y</b></p>
<p><b>Question 11: Could improvements be made to content moderation to deliver greater protection for users, without unduly restricting user activity? If so, what?</b></p>	<p>Confidential? <b>N</b>  We believe that we are doing everything possible to ensure an optimal balance between the safety of our community and the freedom of our users.</p>
<p><b>Question 12: What automated moderation systems do you have in place around illegal content?</b></p>	<p>Confidential? <b>Y</b></p>
<p><b>Question 13: How do you use human moderators to identify and assess illegal content?</b></p>	<p>Confidential? <b>Y</b></p>
<p><b>Question 14: How are sanctions or restrictions around access (including to both the service and to particular content) applied by providers of online services?</b></p>	<p>Confidential? <b>Y</b></p>
<p><b>Question 15: In what instances is illegal content removed from your service?</b></p>	<p>Confidential? <b>Y</b></p>
<p><b>Question 16: Do you use other tools to reduce the visibility and impact of illegal content?</b></p>	<p>Confidential? <b>Y</b></p>

<p><b>Question 17: What other sanctions or disincentives do you employ against users who post illegal content?</b></p>	<p>Confidential? <b>Y</b></p>
<p><b>Question 18: Are there any functionalities or design features which evidence suggests can effectively prevent harm, and could or should be deployed more widely by industry?</b></p>	<p>Confidential? – <b>N</b></p> <p>The most effective way to reduce harm stems from the ability to share information from one company/platform to another. We know that “platform-hopping” can be an effective strategy to thwart investigative efforts, but the more we can share across platforms, the better equipped we can be at detecting and mitigating the propagation of illegal content.</p>
<p><b>Question 19: To what extent does your service encompass functionalities or features designed to mitigate the risk or impact of harm from illegal content?</b></p>	<p>Confidential? <b>Y</b></p>
<p><b>Question 20: How do you support the safety and wellbeing of your users as regards illegal content?</b></p>	<p>Confidential? <b>N</b></p> <p>Illegal content is prohibited on our platform, we remove any content which is reported to us and we proactively review all image, video and audio assets before they are allowed on the platform.</p> <p>Roblox goes beyond moderating illegal content as defined by Ofcom. We make sure that the content displayed on our platform is not only legal, but also appropriate for all our users. Our Community Standards describe the type of content we found objectionable and we take proactive and reactive measures to prevent such content from being available to our users.</p> <p>Roblox hosts a <a href="#">Safety and Civility Center</a> which provides advice on how to prevent harm and what to do when something goes wrong. We are in the process of updating this and will be adding additional advice and resources in the coming months.</p> <p>Collaboration with safety organisations and other platforms is key, we have partnerships with over 20 leading global organisations that focus on child safety and internet safety including the WePROTECT Global Alliance, the Internet Watch Foundation (IWF), the UK Safer Internet Centre, Fair Play</p>

	<p>Alliance, Family Online Safety Institute (FOSI), Digital Wellness Lab, Connect Safely, and kidSAFE among others.</p> <p>We are also a member of various industry organisations, such as UKIE and the Technology Coalition with a goal of cross-industry collaboration in the areas of user safety, and child safety. As a member of the Technology Coalition, we are committed to the Voluntary Principles, including transparency on our efforts to combat online child sexual exploitation and abuse.</p> <p>Additionally, we work with individual companies where we see opportunities to share learnings and development efforts. For example, we worked with Microsoft on a cross-industry project to develop an AI-based technique that scans text-based chats and provide better tools for grooming chat detection. We also use PhotoDNA across the platform.</p> <p>And we work closely with other chat, social media, and UGC (User Generated Content) platforms to report bad actors and content, so they can also take appropriate action on their platforms.</p> <p>In consultation with expert organizations like Anti-Defamation League, Tech Against Terrorism (TAT) and The Simon Wiesenthal Center, as well as academics and safety partners from across the globe, we are constantly evaluating our moderation policies. We frequently audit our platform to ensure we are continually strengthening our processes and algorithms to prevent, detect, and block new content or behaviour that violates our Terms of Use. Our aim is to protect our community against emergent discriminatory or harmful content, terms, memes, and symbols.</p> <p>The main cost of these interventions is memberships of the various organisations listed.</p>
<p><b>Question 21: How do you mitigate any risks posed by the design of algorithms that support the function of your service (e.g. search engines, or social and content recommender systems), with reference to illegal content specifically?</b></p>	<p>Confidential? <b>Y</b></p>



<p><b>Question 22: What age assurance and age verification technologies are available to platforms, and what is the impact and cost of using them?</b></p>	<p>Confidential? <b>Y</b></p>
<p><b>Question 23: Can you identify factors which might indicate that a service is likely to attract child users?</b></p>	<p>Confidential? – <b>N</b></p> <p>Our determinations are based upon the self-attestation of our users. Account registration is required to join the Roblox platform. During the registration process, we request the user enter their date of birth. Based on their entry, we increase the level of safety and privacy within their settings by default (e.g., users under 13 have stricter chat filters).</p>
<p><b>Question 24: Does your service use any age assurance or age verification tools or related technologies to verify or estimate the age of users?</b></p>	<p>Confidential? <b>Y</b></p>
<p><b>Question 25: If it is not possible for children to access your service, or a part of it, how do you ensure this?</b></p>	<p>Confidential? – <b>N</b></p> <p>Children of all ages have access to our platform, but there are some features where we restrict access to older children and adults (ages 13 and up). There are a few ways in which we prohibit young users from accessing these features, depending on the overall risk of the feature.</p> <ul style="list-style-type: none"> <li>● Based on the user’s self-attestation, we prohibit access if the user’s age is under 13. Examples of this include: <ul style="list-style-type: none"> <li>○ Access to publishing social media account information</li> <li>○ Stricter chat filtering for those under 13</li> <li>○ Access to developer sites such as devforum.roblox.com</li> </ul> </li> <li>● Based on the user’s proven age through our age verification feature, we prohibit users under 13 from accessing more free-form communication features. Examples include: <ul style="list-style-type: none"> <li>○ Spatial voice chat</li> <li>○ Talent Hub (a site that encourages our developers to seek talent from the developer community)</li> </ul> </li> </ul>
<p><b>Question 26: What information do you have about the age of your users?</b></p>	<p>Confidential? – <b>N</b></p> <p>We collect the date of birth (self-attestation) for all of our users through our registration process. If a user is 13 or older and consents to the use of our Age Verification function, we confirm the date of birth based on the identification the user provides to our age verification vendor.</p>

**Question 27: For purposes of transparency, what type of information is useful/not useful? Why?**

Confidential? **N**

We only collect and retain information that is necessary and useful to fulfil the intended purposes of the feature. Aside from confirming the eligibility of a user for higher risk features, the verified date of birth has been useful in the following ways:

- To ascertain the user’s actual age so that we may provide them with an age appropriate experience on our platform;
- To assist with identifying real users on the platform versus bots, which we are trying to curb and/or eliminate;
- To assist with improving certain product features, such as age appropriate matchmaking.

**Question 28: Other than those in this document, are you aware of other measures available for mitigating risk and harm from illegal content?**

Confidential? **N**

No, but we continue to work with our partners in this developing landscape.