

Your response

Question	Your response
<p>Question 1: Please provide a description introducing your organisation, service or interest in Online Safety.</p>	<p><i>Is this response confidential? – N</i></p> <p>At Nextdoor, our purpose is to cultivate a kinder world where everyone has a neighbourhood they can rely on. Today, Nextdoor is used in more than 295,000 neighbourhoods in 11 countries around the world. Nextdoor is a user-to-user service where neighbours turn in order to access trusted information, give and get help, get things done, and build real-world connections with those nearby – neighbours, businesses, and public services. Nextdoor is the neighbourhood network that brings these stakeholders together to build thriving, vibrant communities.</p> <p>One of our core company values which defines how we bring our purpose to life is to earn trust every day. Trust, and safety, are paramount to how we operate and we are committed to a transparent and inclusive culture. We believe this is foundational to creating a welcoming platform that facilitates constructive neighbourhood connections and conversations. We appreciate this opportunity to provide Ofcom with evidence of Nextdoor’s approach to safety, and to demonstrate the substantially lower risk of harm posed to users on Nextdoor that comes as a result of our approach.</p> <p>Nextdoor is an ads-based revenue model. Our financial results and business are more fully described in our Quarterly Report on Form 10-Q for the period ended June 30, 2022, filed on August 9, 2022, and our other SEC filings, which are available on the Investor Relations page of our website at www.investors.nextdoor.com and on the SEC’s website at www.sec.gov.</p>
<p>Question 2: Can you provide any evidence relating to the presence or quantity of illegal content on user-to-user and search services?</p> <p>IMPORTANT: Under this question, we are not seeking links to or copies/screenshots of content that is</p>	<p><i>Is this response confidential? – Y</i></p>

illegal to hold, such as child sexual abuse. Deliberately viewing such images may be a criminal offence and will be reported to the police.

Question 3: How do you currently assess the risk of harm to individuals in the UK from illegal content presented by your service?

Is this response confidential? – N

As noted, illegal content is explicitly banned on our platform by our [Community Guidelines](#). But we do not rely solely on our Community Guidelines to reduce harm. We have a multi-layered approach to prevent guideline-violating content by adding moments of friction throughout the platform to create accountability and to slow people down and combat hurtful or harmful content, which would include illegal content. Here are just some of our active in-product features:

- **Welcoming Real People:** Nextdoor is a community built on trust and genuine connections. Since launching our first neighbourhoods in 2011, we have ensured that Nextdoor neighbourhoods are made up of real people nearby. In partnership with trusted vendors, we ensure neighbours are connected to real people in the neighbourhoods that matter to them by requiring everyone to sign up with their real names and addresses.. This ensures that conversations and interactions on Nextdoor are between real people creating trust and mutual accountability. We consider that this reduces the risk of harm that can be presented on internet services.
- **Neighbour Pledge:** Upon joining Nextdoor, all neighbours must agree to the Neighbour Pledge, which is a commitment to treat everyone in the Nextdoor community with respect. We also make our Community Guidelines transparent, comprehensive, and easily accessible. These guidelines help promote thoughtful conversations and explicitly forbid racism, discrimination, misinformation, and other types of harmful content and behaviour.
- **Approach to Content Moderation:** Neighbours are able to address any guideline violations by either reporting the content or the author. Further, our approach to content moderation is to set clear guidelines and use a combination of human review and technology to encourage the behaviours that support our purpose.
 - We've always believed it's important to incorporate local context into moderation decisions, which is why we have built our community volunteer programs to empower thoughtful moderation for neighbours, by neighbours. At the same time,

review of potentially harmful content, e.g., misinformation and discrimination, and neighbour behavior is handled by trained specialists on the Nextdoor staff to ensure consistency and so that we can take appropriate action to support the neighbours involved. We work regularly with leading experts, including our [Neighbourhood Vitality Advisory Board](#), to refine our Community Guidelines, iterate on our features and tools, and develop strategic research teams that further our work to create and maintain a welcoming platform.

- **Kindness Reminder:** Our innovative [Kindness Reminder](#) is a critical part of our infrastructure that detects language that may potentially violate our Community Guidelines, and encourages the author to edit their content before they publish. This prompt was developed in conjunction with social scientists, and is designed to encourage positivity and minimise both the creation and visibility of harmful content across the Nextdoor platform. In 2021, neighbours who encountered this reminder edited or withheld their post or comment on average 34.6% of the time. Importantly, this technology has also been deployed across other topics where slowing down the user has benefits.



Hold on ...

Your post contains a word or phrase that many find hurtful. Consider editing before you publish. [Learn more.](#)

Edit post

Publish

Figure 1: Screenshot of the Nextdoor 'Kindness Reminder'.

The features listed above are just a few examples that demonstrate our unique commitment to investing and building a neighbourhood platform whose central value is to earn and maintain user trust.

Question 4: What are your governance, accountability and decision-making structures for user and platform safety?

Is this response confidential? – N

Teams across Nextdoor work together to ensure user and platform safety. These teams include: product, trust & safety, policy, legal, security, ads, comms and customer success. Please see our answers to Questions 3, 8, 10, and 19 for additional information.

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

Is this response confidential? – N

Nextdoor believes that clear and accessible terms of service and community guidelines are essential to ensure that users and providers know what behaviour and content is acceptable, and what is not. At Nextdoor, we have taken a number of steps to ensure that this is the case for our neighbours.

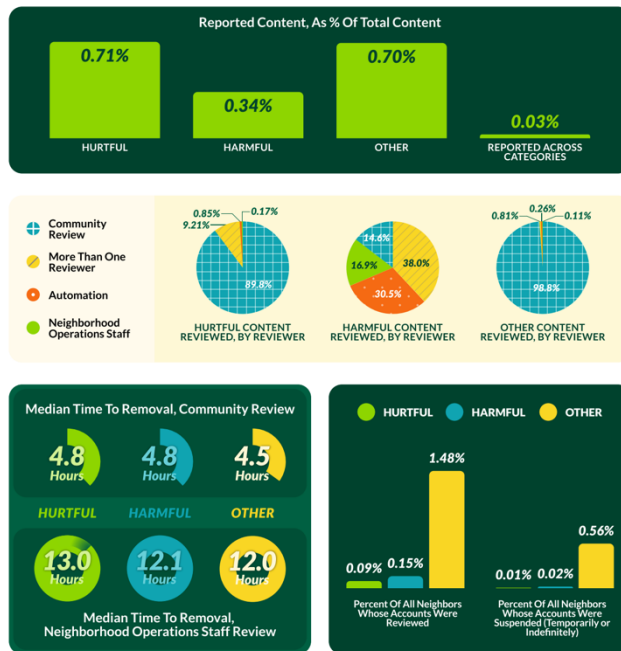
- **In-app stand-alone pledge to follow guidelines:** Upon joining Nextdoor, all neighbours are prompted to read the [Neighbour Pledge](#), which is a commitment to treat everyone in the Nextdoor community with respect which they must agree to to use Nextdoor. We also make our [Community Guidelines](#) transparent, comprehensive, and easily accessible. These guidelines help promote thoughtful conversations and explicitly forbid racism, discrimination, misinformation, and other types of harmful content.
- **In-app Notification & Appeals Process Around Violating Content:** We notify neighbours when a message they have posted has been reported and the community review team (i.e. neighbourhood leads and review team members) has voted to remove it. Rather than removing content completely when this happens, we will instead hide the content so that it is only visible to the author (the poster of the message) and send a notification.
 - Once the author has been notified that their content has been hidden, if they wish to get their content restored, they have two options:
 - Submit an appeal where they explain why they believe their message is consistent with the Nextdoor Community Guidelines and should be restored.
 - Edit their message to remove any problem elements (which will automatically submit the edited message for appeal – the message remains hidden during appeal).
 - A different set of community review team members will then review the appeal and vote on whether to restore the hidden content. If they decide the content does not violate the community guidelines, the appeal is approved and the content will be restored. If, however, they decide that the content is still not appropriate for Nextdoor, or we fail to

	<p>reach sufficient feedback on the appeal to reject or restore, and the appeal is not approved, the content will remain hidden and cannot be appealed further.</p>
<p>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?</p>	<p><i>Is this response confidential? – N</i></p> <p>As noted in our response to Question 2, Nextdoor’s Community Guidelines prohibit content that could be hurtful or harmful to neighbours, which is significantly broader than illegal content. Specifically, Nextdoor prohibits the following content:</p> <ul style="list-style-type: none"> • Uncivil content - insults, belittling remarks, name-calling • Bullying or harassment • Racism • Discrimination • Hate speech, including support for hate groups or hate-based conspiracy theories • Condoning or trivialising violence • Threatening someone’s safety or privacy • Fraud • Spam • Soliciting or offering illegal goods or services • Graphic, violent, sexually explicit or adult content <p>We revisit these Community Guidelines as needed, and meet every two weeks with a cross-functional team to discuss whether changes are needed.</p>
<p>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?</p>	<p><i>Is this response confidential? – N</i></p> <p>At Nextdoor we know how important reporting and complaints mechanisms are to ensuring a positive, safe experience. To that end, we have taken several steps to ensure neighbours make use of these mechanisms wherever appropriate.</p> <ul style="list-style-type: none"> • In-app Reporting and Complaints: Neighbours are able to address any guideline violations by either reporting the content or the author in-app while viewing the piece of content or author - there are dedicated links to file these reports on each piece of content and each neighbour profile. When filing a report, the reporting user is required to provide a guideline violation that serves as the basis for filing the report. We also have a help page dedicated to explaining this to users. • In-app Notification & Appeals Process Around Violating Content: We notify neighbours when a message they have posted has been reported and the community review team (i.e. neighbourhood leads and review team members) has voted to remove it. Rather than removing content

	<p>completely when this happens, we will instead temporarily hide the content so that it is only visible to the author (the poster of the message) and send a notification.</p> <ul style="list-style-type: none"> ○ Once the author has been notified that their content has been hidden, if they wish to get their content restored, they have two options: ○ Submit an appeal where they explain why they believe their message is consistent with the Nextdoor Community Guidelines and should be restored. ○ Edit their message to remove any problem elements (which will automatically submit the edited message for appeal). ○ A different set of community review team members will then review the appeal and vote on whether to restore the hidden content. If they decide the content does not violate the community guidelines, the appeal is approved and the content will be restored. If, however, they decide that the content is still not appropriate for Nextdoor and the appeal is not approved, the content will be removed.
<p>Question 8: If your service has reporting or flagging mechanisms in place for illegal content, or users who post illegal content, how are these processes designed and maintained?</p>	<p><i>Is this response confidential?</i> – N</p> <p>At Nextdoor, our efforts to address guideline-violating (including illegal) content include:</p> <ul style="list-style-type: none"> ● Tools to automatically detect and report harmful content; ● Product features that enable neighbours to report guideline-violating content; ● Volunteer community moderators on Neighbourhood Teams who monitor community discussions and help keep dialogue on the platform civil; and ● A Neighbourhood Operations Team of trained specialists who review content and accounts that have been flagged and take appropriate action to support the neighbours involved. <p>More specifically, on our in-app Reporting and Complaints features: Neighbours are able to address any guideline violations by either reporting the content or the author in-app while viewing the piece of content or author - there are dedicated links to file these reports on each piece of content and each neighbour profile. When filing a report, the reporting user is required to provide a guideline violation that serves as the basis for filing the report. We also have a help page dedicated to explaining this to users.</p> <p>Additionally, Nextdoor has volunteer community moderators on Neighbourhood Teams who are monitoring community discussions 24/7.</p> <p>In 2021, we had 233,615 volunteer community moderators on our Neighbourhood Teams who voted on at least one piece of reported content. These volunteers are active Nextdoor</p>

	<p>neighbours who have access to moderation tools to ensure the guidelines are being followed. When content is reported by a neighbour or by our automated systems, volunteer community moderators can vote on whether or not they think it violates Nextdoor's Community Guidelines. In 2021, our volunteer community moderators reviewed 86.9% of all reported content (1.8% of all pieces of content), and voted to remove 52.5% of reported content in a median time of 4.6 hours.</p> <p>While volunteer community moderators review most types of guideline-violating content, reports of certain types of harmful content, like misinformation and discrimination, are sent directly to and handled by our trained Neighbourhood Operations staff. Given that this content can be particularly sensitive, we rely on our internal agents who have special training to ensure consistent and objective outcomes.</p>
<p>Question 9: If your service has a <i>complaints</i> mechanism in place, how are these processes designed and maintained?</p>	<p><i>Is this response confidential? - N</i></p> <p>Please see our answer to Question 8.</p>
<p>Question 10: What action does your service take in response to <i>reports</i> or <i>complaints</i>?</p>	<p><i>Is this response confidential? - N</i></p> <p>Nextdoor takes a variety of actions in response to reports or complaints, depending on the nature of the report or complaint. In addition to our answers to Questions 8 and 9, below we have included screenshots detailing our moderation efforts, as was published in our Nextdoor Transparency Report 2021.</p>

Moderation on Nextdoor



Automated Reporting & Review

We use automation to help identify content for human review as well as moderate user-reported content.

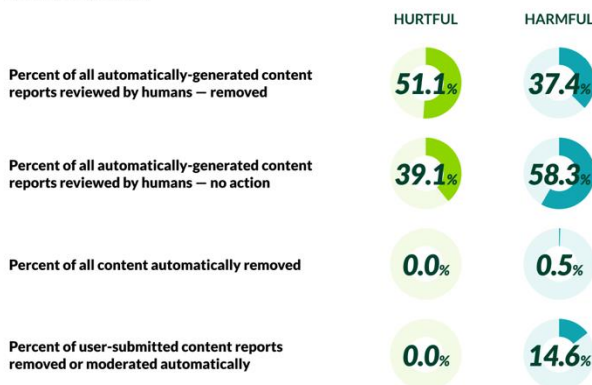


Figure 2: Screenshots from Nextdoor's Transparency Report 2021 highlighting our moderation approach.

Question 11: Could improvements be made to content moderation to deliver greater protection for users, without unduly restricting user activity? If so, what?

Is this response confidential? – N

At Nextdoor, we are routinely thinking about the ways in which content moderation can create the kinds of neighbourhoods our users love, while avoiding overmoderating or creating a "one size fits all" approach.

That is why, as detailed in other answers, we have adopted a strategy that both seeks to catch illegal content as early as possible through both automated and human moderation, while also delegating to some extent wider moderation of legal

	<p>content to individual neighbourhoods (with support from a central team where appropriate) to reflect local communities. We have roughly a quarter of a million community moderators operating on the Nextdoor platform who deliver greater protection for users while avoiding unduly restricting user activity.</p> <p>As well as dealing with content once it is on platform, we have also taken steps to help users post appropriately in the first instance with initiatives such as the Kindness Reminder.</p>
<p>Question 12: What automated moderation systems do you have in place around illegal content?</p>	<p><i>Is this response confidential?</i> – N</p> <p>Please see our answers to Questions 2 and 19.</p>
<p>Question 13: How do you use human moderators to identify and assess illegal content?</p>	<p><i>Is this response confidential?</i> – N</p> <p>Please see our answers to Questions 2 and 19.</p>
<p>Question 14: How are sanctions or restrictions around access (including to both the service and to particular content) applied by providers of online services?</p>	<p><i>Is this response confidential?</i> – N</p>
<p>Question 15: In what instances is illegal content removed from your service?</p>	<p><i>Is this response confidential?</i> – N</p> <p>Illegal content is absolutely prohibited by Nextdoor’s Community Guidelines. It is removed in any and all circumstances in which it is identified. Please see our answer to Question 2 for further details.</p>
<p>Question 16: Do you use other tools to reduce the visibility and impact of illegal content?</p>	<p><i>Is this response confidential?</i> – N</p> <p>As noted above, Nextdoor is focused on removing any content that may be hurtful or harmful to its users, which includes illegal content. Nextdoor is in the process of releasing a new comment filtering feature designed to reduce the visibility of content that violates the community guidelines. When content has been reported and Nextdoor’s automated systems determine that the content is likely to be removed by moderators, these comments will be “filtered”. This means that by default, these comments will not appear to neighbours viewing the discussion. However,</p>

	<p>neighbours who wish to see these comments will have the option to adjust their filter to view all comments.</p>
<p>Question 17: What other sanctions or disincentives do you employ against users who post illegal content?</p>	<p><i>Is this response confidential?</i> – N</p> <p>Please see our answers to Questions 8 - 10.</p>
<p>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?</p>	<p><i>Is this response confidential?</i> – N</p> <p>The effectiveness of any functionality or design in preventing harm will depend on the nature of the service and the user base. Many such functionalities or designs will not necessarily work across different kinds of service. We will provide an example of a functionality we have introduced at Nextdoor.</p> <p>Our innovative Kindness Reminder is a critical part of our infrastructure that detects language that may potentially violate our Community Guidelines, and encourages the author to edit their content before they publish. This prompt was developed in conjunction with social scientists, and is designed to encourage positivity and minimise both the creation and visibility of harmful content across the Nextdoor platform. In 2021, neighbours who encountered this reminder edited or withheld their post or comment on average 34.6% of the time. Importantly, this technology has also been deployed across other topics where slowing down the user has benefits.</p>
<p>Question 19: To what extent does your service encompass functionalities or features designed to mitigate the risk or impact of harm from illegal content?</p>	<p><i>Is this response confidential?</i> – N</p> <p>In addition to the product features outlined in Question 3, we take the following steps:</p> <p>Community Moderation Supplemented by Technology</p> <p>Removing harmful content is key to building stronger neighbourhoods, and at Nextdoor, we rely on a combination of both technology and human-initiated reports to keep our members safe. It is important to note, however, nearly all content moderation reports on Nextdoor, regardless of how they are generated, are reviewed by a human at some point in the moderation process.</p> <p>Community-led moderation is the core of our moderation system. Because of the nature of Nextdoor’s Community Guidelines and the higher standard they place on neighbours, we empower neighbours to moderate their communities by enabling them to report content and members for behaviours and content which contravenes these Community Guidelines, including whether or not a post should be removed for over-posting, violations of commercial guidelines, and general incivility or</p>

arguments between neighbours. Neighbours participation in the community-led moderation process is critical, as understanding the local culture and nuance of a community ensures an online experience that is truly reflective of the offline neighbourhood.

Nextdoor takes different actions depending on the type of content that is reported. For the majority of content that is reported for possibly violating our Community Guidelines, specifically designated volunteer content reviewers in the community engage in content review. Removal automatically occurs at the point a piece of content receives a threshold number of recommendations for removal.

On the other hand, where a neighbour reports content that is illegal or falls into the category of online harms, it is reviewed by our Neighbourhood Operations team for moderation. If we receive a signal either through community reviewers or automated content classifiers that a particular piece of content is especially harmful and/or illegal, we may take stronger action and completely limit its distribution while it is under review by a member of staff.

We have additional technological solutions that we apply as well to identify content that violates our guidelines:

- **Blocklist** - After deeming a certain piece of content to have violated our guidelines, we have the ability to add that specific piece of content to a list so that if that same piece of content is reposted, then our technology will identify the content and flag it for review. This typically applies to, for example, fraud and Covid-19 misinformation.
- **Active account security monitoring** to ensure account takeovers don't lead to abuse on the platform.
- **Content Amplification & Suppression** - For user-generated content, we have internal models that we can apply to content in order to show our members the most relevant and useful content. Similarly, we also use these models to identify whether a piece of content may be potentially harmful and then, if warranted, suppress that content to reduce its visibility.
- **Content labels** - after we determine that a piece of content possibly (though not probably) violates our guidelines, we may apply a "label" that provides contextual information about content that may violate guidelines, for example to point users to health resources from the appropriate in-market health body.

Finally, Nextdoor users have a significant amount of control over what content is shown to them and who will see their content. Specifically:

- Nextdoor provides the option for users to have a chronological feed reflecting what users post and

	<p>when, rather than one curated by algorithms. See this help page</p> <ul style="list-style-type: none"> • Nextdoor gives users the ability to only see posts from certain neighborhoods. See this help page • Nextdoor gives users the option, on a per-post basis, to determine who will see their posts or profile. See this help page.
<p>Question 20: How do you support the safety and wellbeing of your users as regards illegal content?</p>	<p><i>Is this response confidential? - N</i></p> <p>Please see our answer to Question 19.</p>
<p>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?</p>	<p><i>Is this response confidential? - N</i></p> <p>Please see our answer to Question 19.</p>
<p>Question 22: What age assurance and age verification technologies are available to platforms, and what is the impact and cost of using them?</p>	<p><i>Is this response confidential? - N</i></p>
<p>Question 23: Can you identify factors which might indicate that a service is likely to attract child users?</p>	<p><i>Is this response confidential? - N</i></p>

<p>Question 24: Does your service use any age assurance or age verification tools or related technologies to verify or estimate the age of users?</p>	<p><i>Is this response confidential? – N</i></p>
<p>Question 25: If it is not possible for children to access your service, or a part of it, how do you ensure this?</p>	<p><i>Is this response confidential? – N</i></p>
<p>Question 26: What information do you have about the age of your users?</p>	<p><i>Is this response confidential? – N</i></p> <p>While Nextdoor does not have any services that are directed at children, some older teenage users may be interested in joining Nextdoor to identify short-term job opportunities or to offer assistance as a neighbour-for-hire to fellow neighbours for tasks including babysitting, gardening, or other useful household tasks. As a result of this observation, we recently started collecting the age of our users. In the UK, Nextdoor requires that members be at least 13 years of age to use the service, and we are now collecting date of birth to ensure members meet this requirement. Having this information also allows us to build age-appropriate features and privacy controls for our users under 18. In addition, Nextdoor will leverage age data to ensure that neighbours see the most relevant possible content and advertising on Nextdoor. For example, Nextdoor may use age data to prioritise the groups we surface to members (e.g. more prominently featuring groups meant for seniors to members over the age of 65), and ensure that users who are under 18 do not receive age-targeted advertising.</p>
<p>Question 27: For purposes of transparency, what type of information is useful/not useful? Why?</p>	<p><i>Is this response confidential? – N</i></p> <p>We believe that the following would be valuable for transparency purposes, to be included in an annual report with the following information:</p> <ul style="list-style-type: none"> • Copy of terms of service or community guidelines that outline types of content and member activity that may be moderated • How content/members are reported and moderated

	<ul style="list-style-type: none">• Percent of content/overall content and members/overall members that is moderated, potentially by content/member activity type• Transparency around government requests for information - types of requests x raw numbers of responses/rejections/accounts impacted
Question 28: Other than those in this document, are you aware of other measures available for mitigating risk and harm from illegal content?	<i>Is this response confidential? - Y</i>