Transparency Report 2023

February 2024

nextdoor

2023 #LoveYourNeighborhood submission from
Roland W, in San Rafael Heights neighborhood in CA, US

About this report

Nextdoor is committed to creating a welcoming platform that offers neighbors worldwide a place to build authentic connections, both online and in real life. As the world grapples with a loneliness epidemic, the human need for connection is real and lasting. Nextdoor is where neighbors can naturally unite to improve their local communities and lives.

Across the world, there is a universal desire for the connectedness of proximity, even though much of our daily online activities link us to things and people that are physically distant. People everywhere still cherish the sense of belonging to a tangible, real place—caring about their neighbors and community. Nextdoor brings the internet closer to home, fostering true local connections. Meaningful local discussions, community organizations, and businesses inspire neighbors to participate in real-world interactions, cultivating stronger, more trusted communities.

Countless studies show that well-being is higher among people with regular contact with neighbors. We need no convincing that this is a reality. We see it all the time on Nextdoor.

Millions of new neighbors joined Nextdoor in 2023 after completing the registration and verification process. With nearly 85 million verified neighbors on the platform, we continue to prioritize developing features that encourage interactions that support our purpose of cultivating a kinder world where everyone has a neighborhood they can rely on.[1]

In this year's report, we:

- Showcase our innovative use of generative artificial intelligence to foster positive conversations, and facilitate local businesses in connecting with their community.
- Present the evolution of our products aimed at ensuring fairness in content moderation.
- Illustrate the effectiveness of our interventions in minimizing reports of hurtful and harmful content.
- Offer transparency in our approach to handling government requests for neighbor information and their content.

Where noted, we also provide information required by the European Union Digital Services Act.

1. Verified neighbor data as of 9/30/23. Citations to data hereafter are for the entirety of 2023. Nextdoor estimates that over 99% of its neighbors (i.e. users) are 18 and over. Further, Nextdoor estimates that approximately 40% or more of Nextdoor neighbors are 55 and over.

New investments in neighborhood vitality

A neighborhood with vitality can be characterized by a strong sense of community, where neighbors, businesses, and public agencies can depend on each other to exchange valuable information, goods, and services. Nextdoor plays a crucial role in kindling these connections and linking millions of individuals back to the original social network—the neighborhood.

Nextdoor encourages open conversations with varied perspectives, recognizing their role in constructing robust and connected communities. However, as with conversations in person, some conversations may become contentious. To ensure that our products contribute positively to conversations, we use machine learning to anticipate instances where neighbors might violate our Community Guidelines, and then insert moments of friction into the product to foster civil discourse and prosocial behaviors.

Our insights reveal that a small fraction (less than half a percent) of neighbors generate nearly 60% of removed content, mainly in Comments rather than Posts — which is predominantly removed for being disrespectful. Importantly, however, the same small group of neighbors who generate a majority of removed content also post a significant amount of unreported content that is often positive and useful.

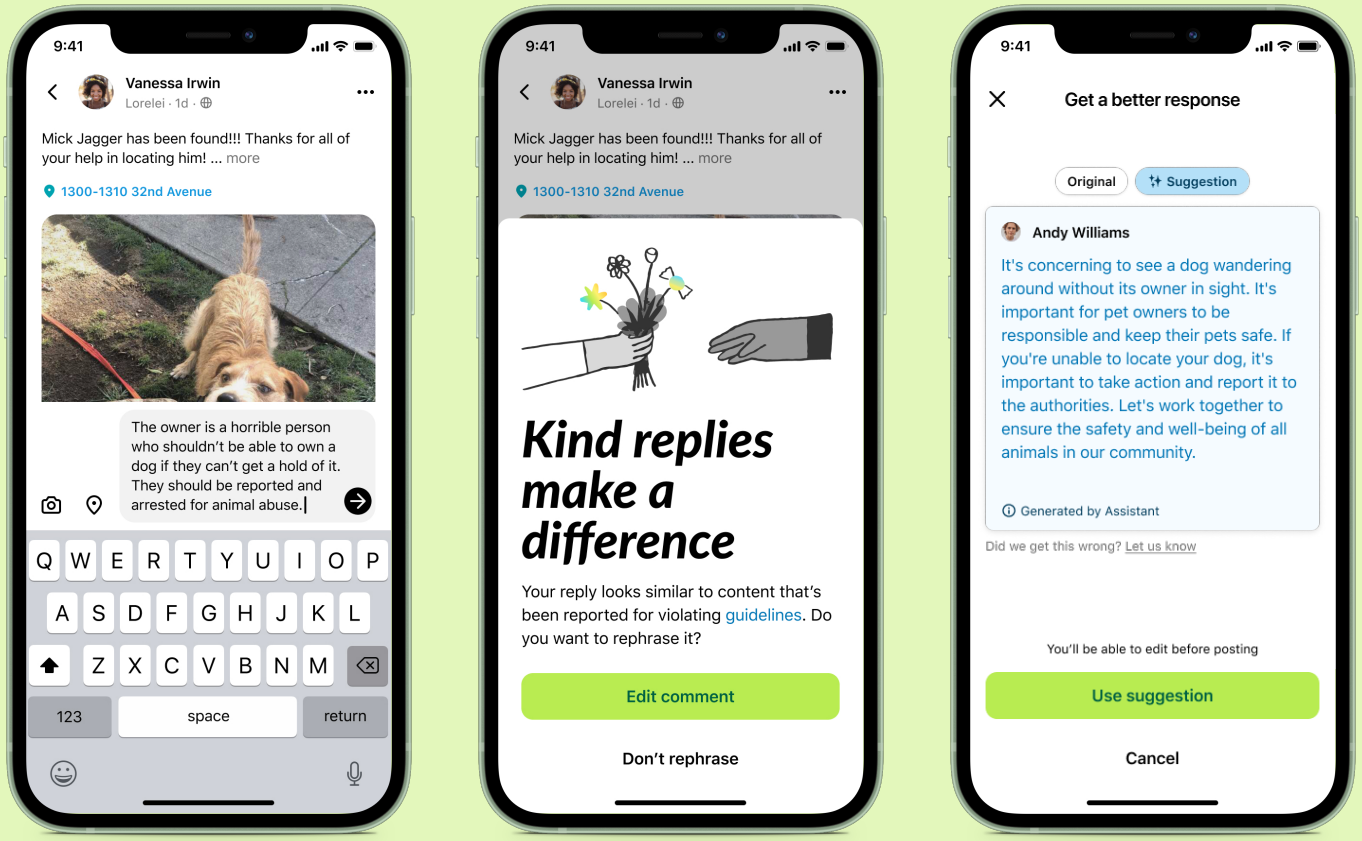
For this reason, we remain steadfast in our commitment to developing cutting-edge product technology that encourages neighborly interactions and enhances neighborhood vitality, especially during conversations with divergent viewpoints. Over the past year, we have iterated on core features prioritizing constructive conversations in Comments, and implementing updates driven by predictive technology and machine learning.

Red Teaming and Deploying Generative Artificial Intelligence to Aid Constructive Conversations in Neighborhoods

Nextdoor's Kindness Reminder is a historic feature designed to detect content that may violate the Community Guidelines and encourages the author to edit their content before they publish. It was the first of our core product features to introduce moments of friction aimed at slowing people down to combat incivility. This feature has consistently yielded positive outcomes over the years, with neighbors who encountered the reminder editing or withholding their content 35% of the time in 2021 and 36% of the time in 2022.

In 2023, Nextdoor introduced generative artificial intelligence (AI) into our product in a variety of critical ways. In particular, we gave neighbors the ability to use generative AI to help write posts that are more likely to drive positive community engagement[2], and made a formative change to the longstanding Kindness Reminder by integrating AI into the product. This new iteration (the Generative AI Kindness Reminder) takes the traditional Kindness Reminder a step further by suggesting more constructive revisions for flagged comments. Leveraging AI technology has enabled us to better analyze tone and context, and gives us the ability to propose alternative phrasings to prevent misunderstandings and conflicts.

2. We found that more than 70% of neighbors who opted to use generative AI to edit their posts accepted the generative AI suggestion, and that small businesses used the generative AI Assistant at twice the rate of other neighbors.

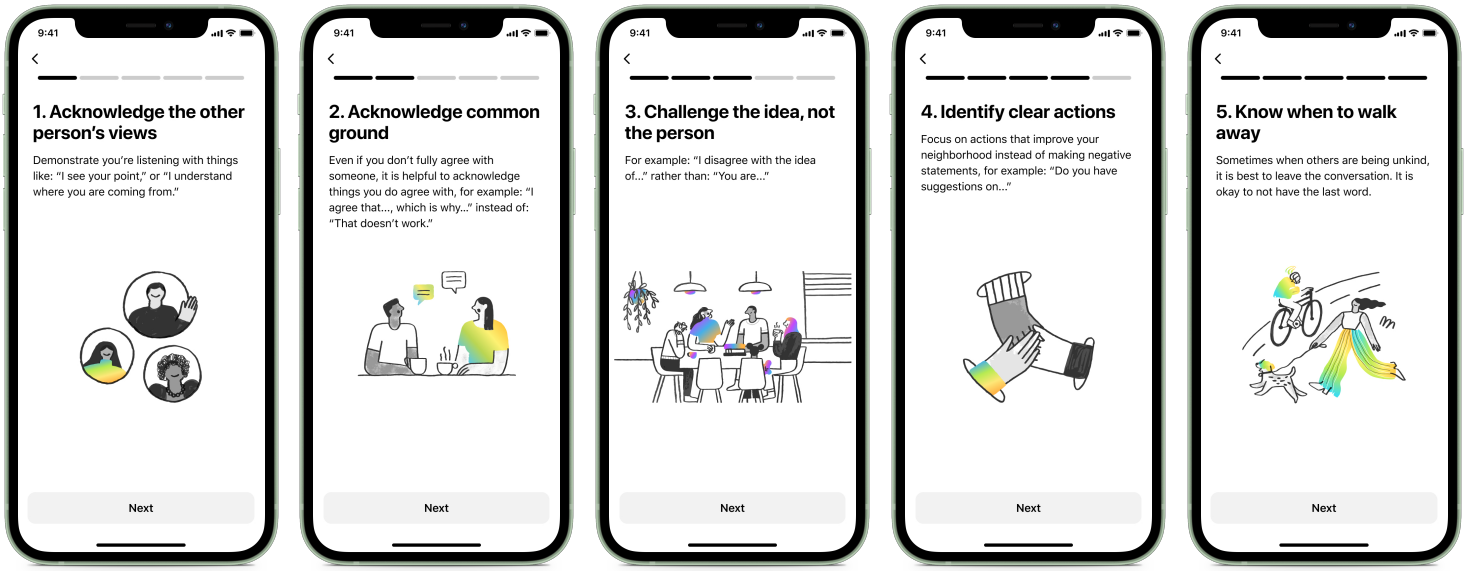


Prior to launch, we ensured adherence to our Generative AI Ethics Principles—Transparency, Privacy, Safety, Accountability, and Inclusivity. We also red teamed the integration with internal and external partners, including our Neighborhood Vitality Advisory Board, and focused on issues such as fine tuning the generative AI model to uphold unique local speech patterns. Finally, we implemented a feedback mechanism for neighbors to report instances where the AI-generated content didn't perform as intended. Notably, only 0.16% of neighbors who encountered the AI Kindness Reminder provided constructive feedback such as changes in meaning, lack of coherence, or a mismatch with their usual tone.

The results are promising. In 2023, 36% of neighbors who saw the AI Kindness Reminder or the traditional Kindness Reminder (where neighbors self-edit without AI assistance) chose to edit or withhold their content, maintaining historically positive results. Further, of those who encountered the generative AI-revised text, 26% adopted the suggestion and published more constructive content. Exposure to the Kindness Reminder and AI Kindness Reminder **reduced the creation of content that violated the Community Guidelines by 15%.**

Kindness Tips:

As noted above, the same small percentage of neighbors who generate nearly 60% of removed content also post a significant amount of content that is not reported, and is often positive and useful. In order to keep these neighbors sharing useful content and to reduce the likelihood of violative content, we introduced Nextdoor Kindness Tips to serve as a supportive tool to remind neighbors who have had content previously removed about best practices for fostering constructive conversations. Developed in collaboration with social scientists, these tips offer five specific, actionable pieces of advice with illustrative examples. Importantly, instead of just taking corrective action on repeat violators, Kindness Tips aims to keep neighbors on the platform by proactively guiding neighbors to reflect on how to engage in open and respectful discussions while aligning with our Community Guidelines.



Moderation by and for real neighbors

Nextdoor's Community Guidelines are designed to keep interactions on the platform safe and productive. Any neighbor is able to report content or a neighbor that may violate Nextdoor's Community Guidelines. For the most part, the Guidelines are enforced by real people who live in the local neighborhoods they moderate. Community moderators are supplemented by paid Nextdoor Operations staff as well as technology, all of which work to detect and review three main categories of guideline-violating content:

- **Hurtful:** Content that neighbors consider uncivil, e.g., insults, rudeness, name-calling.
- **Harmful:** Content that is illegal, fraudulent, or unsafe, e.g., violent, graphic, discriminatory.
- **Other:** Non-local content, spam, content posted in error.

Community Moderators: Localized community moderation ensures that conversations on the platform are reflective of the real world and content reports are handled expediently. Our community moderators are volunteers who are active, local Nextdoor neighbors with access to dedicated moderation tools that they use to enforce Nextdoor's Community Guidelines.

When content is reported by neighbors or detected by our automated systems, community moderators can vote on whether or not they think the content violates the Guidelines. When a consensus is reached by local moderators, our moderation platform takes action in accordance with the recommendation. When local moderators recommend that a piece of content violates the Guidelines and is removed, the content's author is given the opportunity to review and appeal the decision should they choose.

Though local community moderators review most types of guideline-violating content, reports of certain types of harmful content, like misinformation and discrimination, are sent directly to Nextdoor Operations staff for review and actioning.[3] In addition, volunteer community moderators have no insight into neighbor account reports, nor do they have any ability to take action on a neighbor's account. Finally, if a neighbor believes a moderator in their community is not moderating in alignment with Nextdoor's Community Guidelines, then they are empowered to report them to Nextdoor Operations staff for review.

In 2023, Nextdoor's nearly 200,000 volunteer community moderators reviewed 90% of all reported content (1.97% of all content), and removed 55% of reported content in a median time of 5.3 hours. The remaining reported content was reviewed by paid Nextdoor Operations staff or automatically removed, as detailed below in Figures A–C.

Inclusive Moderation Course: In 2020, Nextdoor launched an optional online course designed to educate community moderators about cultural competency and provide additional supportive communication tools. Since launch, nearly 44,000 moderators have enrolled in the online course, with a completion rate of 31.5%, over double the online education industry average completion rate. Not only did attendees find the course educational and improved their ability to moderate without bias, but they also found it personally valuable, with 96% of respondents rating the course from good to excellent.

Nextdoor Operations Staff: While volunteer community moderators review most types of content that violates the Community Guidelines, neighbor reports and reports of certain types of harmful content, like misinformation and discrimination, are sent directly to and handled by our trained Neighborhood 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. (See Figures B–D.)

Use of Automation in Moderation: Nextdoor uses a combination of machine-learning models and heuristic-based rules to identify potentially violative content on our platform and automatically generate reports. These content reports are further reviewed by our community moderators prior to a removal decision being made. Fraud cases, however, stand as a unique category where they undergo moderation by automation. Despite the role played by automated moderation within our ecosystem, it's essential to highlight that the majority of reports originate from our neighbors, and nearly all undergo human review. Please see Figures B and E for data on automated reporting and review on Nextdoor.

Fairness in Moderation: Our philosophy on moderation centers around transparency, consistency, participation, and proportionality. It is vital that neighbors and organizations understand the moderation system, observe its functionality, and have a voice in the process. By ensuring actions align with the harm caused, satisfaction with moderation outcomes increases, promoting adherence to community norms and guidelines.

In 2022, Nextdoor introduced the ability for neighbors to appeal in-app when their content was removed, and expanded in-app appeals to account suspensions in October 2023. We also added the ability for neighbors who reported content to appeal if they believe that content was erroneously upheld on the platform. Results are included in Figure F.

3. Some reports of multi-guideline-violating content are reviewed by both community moderators and our trained Nextdoor Operations staff. When this occurs, a decision by either moderator to remove content is determinative, with appeals available.

Moderation on Nextdoor

Figure A

Reported Content, As % Of Total Content

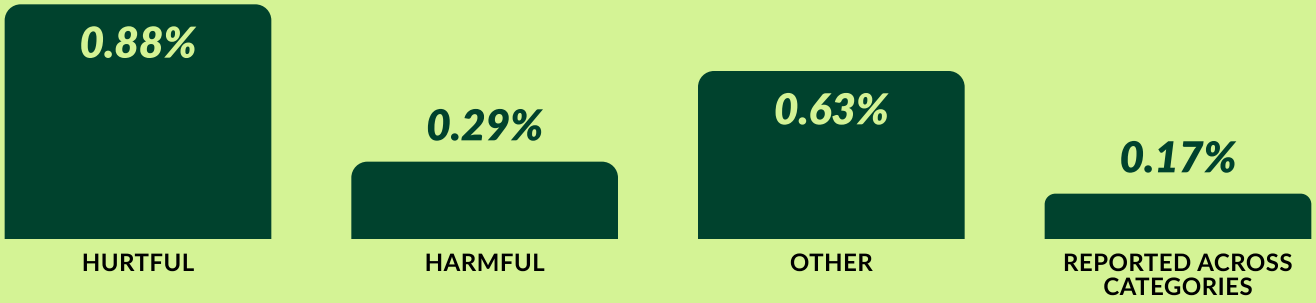


Figure B

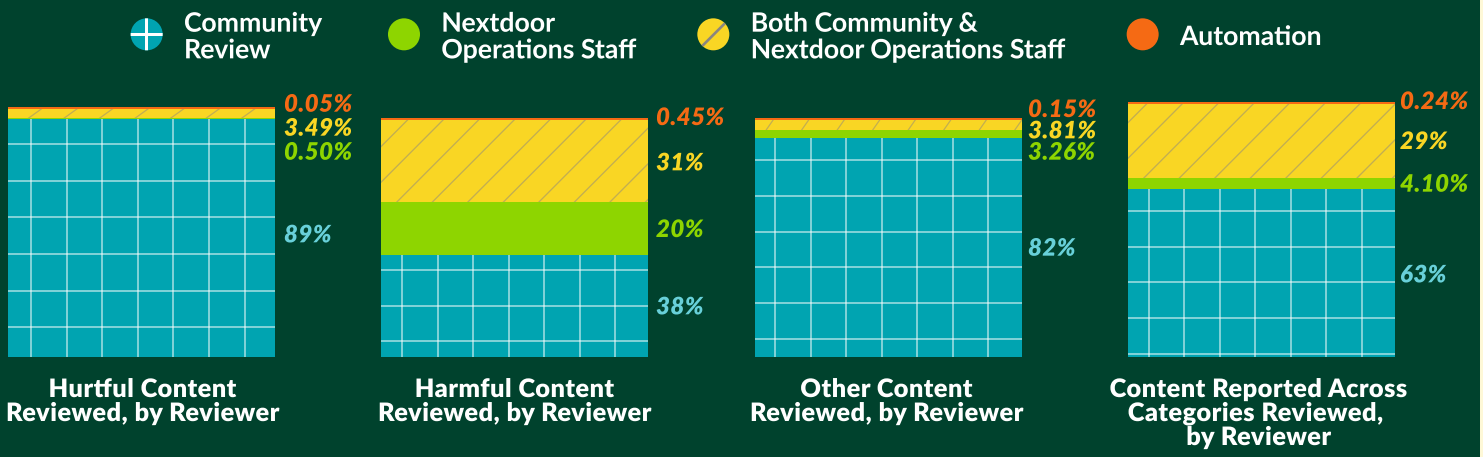


Figure C

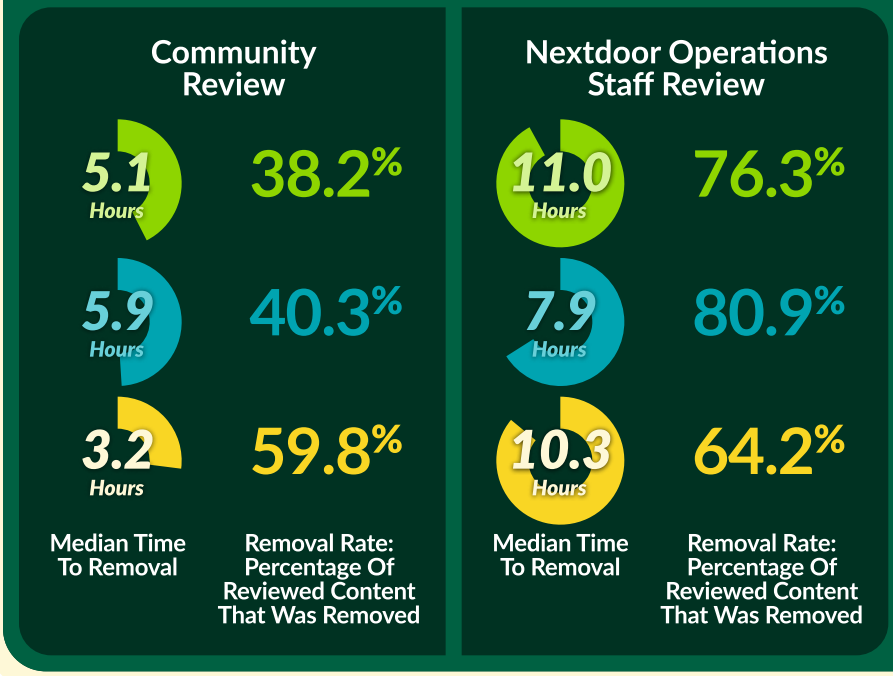
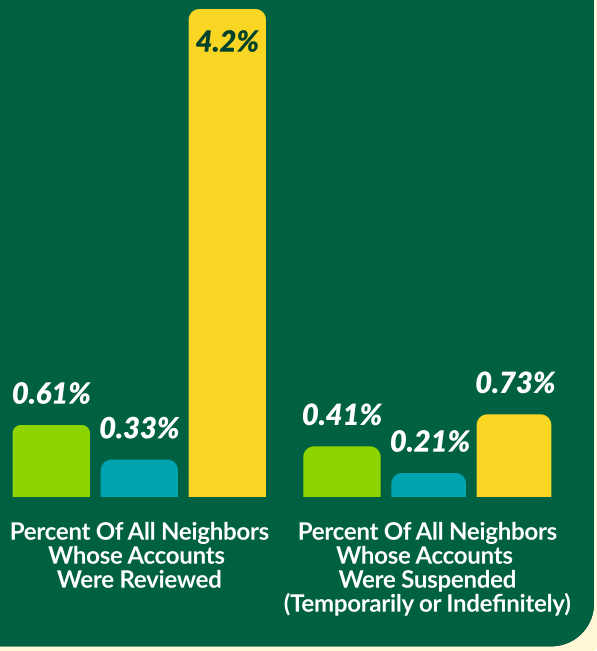


Figure D

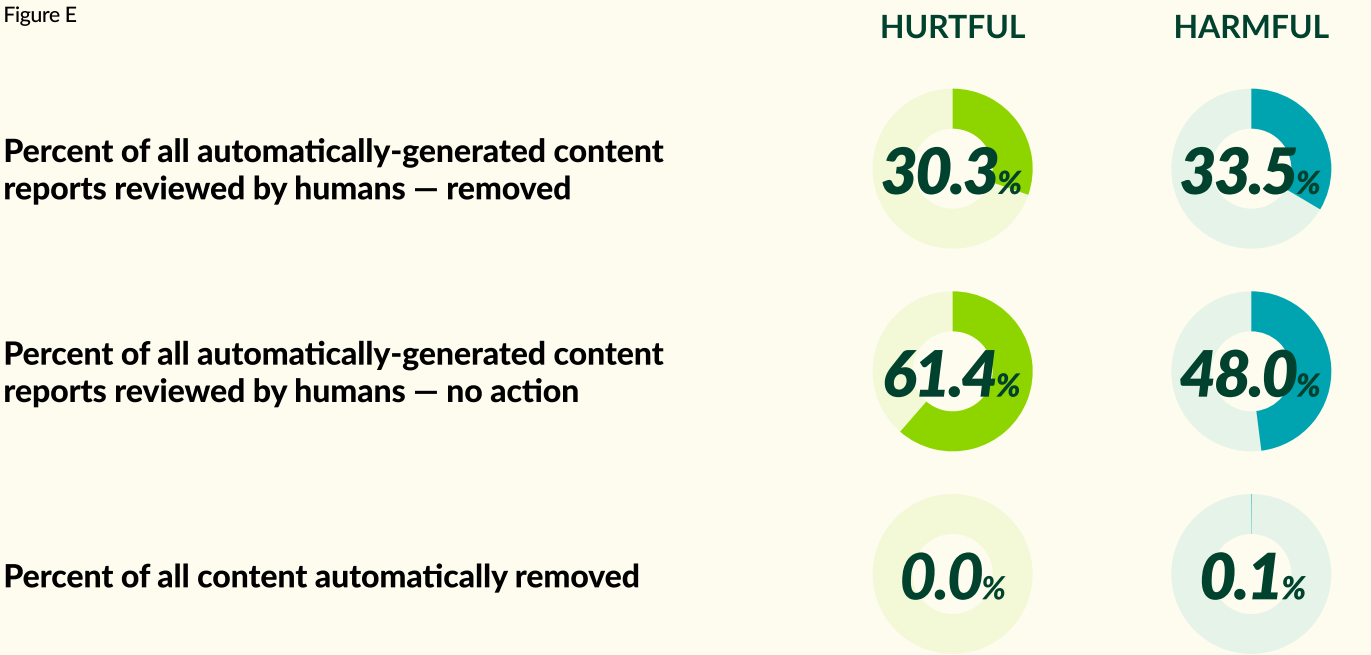


Data is global unless otherwise noted. Data for this report does not include duplicates across reporting categories.

Automated Reporting & Review

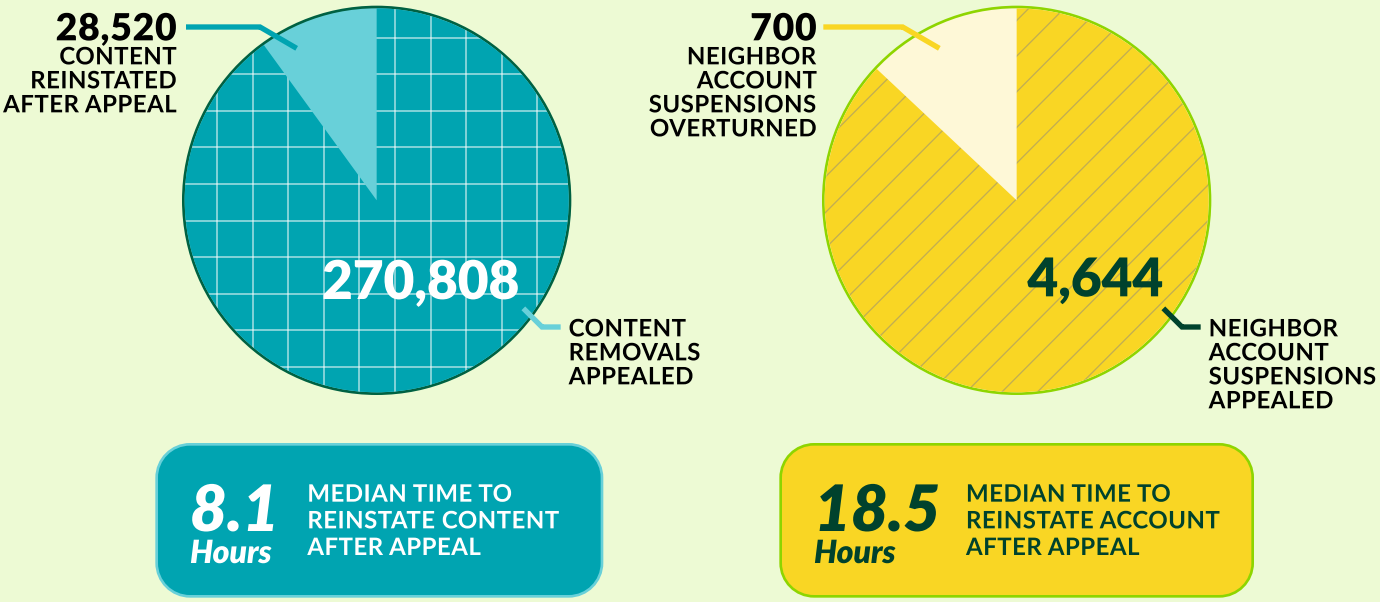
We use automation to help identify content for human review as well as moderate user-reported content.

Figure E



Appeals Metrics[4]

Figure F



Data is global unless otherwise noted. Data for this report does not include duplicates across reporting categories.

4. Appealed suspensions do not include users who were suspended for using a fake name and requested assistance with reinstatement.

Responding to government requests for neighbor information

Government agencies around the world can, by law, request neighbor information for civil, administrative, and law enforcement reasons from Nextdoor. In this report, we've included information about the raw number and type of requests we received from government agencies in 2023 for nearly 85 million verified neighbors.

Prior to providing information to government agencies, each request is reviewed to make sure it satisfies applicable laws.[5] From there, it is determined whether content is available and can be provided.

2023 United States Government Requests for Information

Nextdoor received 27 requests for neighbor information within the U.S. in 2023, and proactively filed six Cybertip reports with the National Center for Missing and Exploited Children (NCMEC).

Figure G

	Raw Number of Requests	Provided Content	Provided Non-content Only	Preserved	No Data Found	Rejected	Number of Accounts Impacted
COURT ORDER	2	0	1	0	0	1	1
EMERGENCY REQUESTS	3	0	3	0	0	0	3
PRESERVATION REQUEST	6	0	0	4	2	0	0
SEARCH WARRANT	7	5	2	0	0	0	7
SUBPOENA	9	0	5	0	2	2	5
CIVIL	0	0	0	0	0	0	0

2023 International Government Requests for Information

Nextdoor received 32 requests for neighbor information from international government agencies in 2023.

Figure H

	Total Request	Police	Council/Agency
UK	22	10	12
NETHERLANDS	7	3	4
SPAIN	2	2	0
FRANCE	1	1	0

5. See Nextdoor's Privacy Policy for more information.

European Digital Services Act

Article 15, section (1)(a)
2023

Requests	Member State	Type of Request	Criminal Inquiry	Benefits Fraud	Defamation	Median Time to Respond (days)
2	SPAIN	PROVIDE INFORMATION ON USERS	2	0	0	3
7	NETHERLANDS	PROVIDE INFORMATION ON USERS	2	4	1	2
1	FRANCE	PROVIDE INFORMATION ON USERS	1	0	0	4

Article 15, section (1)(b)

In accordance with DSA, Article 15, section 1(b), 3273 pieces of content were reported in the European Union in 2023 for "illegal activity" or “illegal or regulated goods or services.” Of the 3273 items, 508 pieces of content were removed for violating a Community Guideline. 81 pieces of content were removed automatically.

Nextdoor Community Guidelines specify the neighbor behavior and activities that are permitted on Nextdoor, and the neighbor behavior and activities that may subject the neighbor or an item of content to being actioned. We have five major guidelines on Nextdoor, and they necessarily include illegal content. As such, all removed content was removed pursuant to our Community Guidelines. The median time to removal was 13 hours.

Article 15, section (1)(c)-(e)

Please see above for further details.

Article 24, section (1)(a)

Not applicable for 2023.

Article 24, section (1)(b)

Nextdoor imposed 45 suspensions on neighbors for posting content that was “illegal or regulated goods or services” in the European Union in 2023.

Article 24, section (2)

Average Monthly Active Recipients: In accordance with Article 24, section (2) of the DSA, Nextdoor reports that the average monthly active recipients of our services in the European Union over the six month period beginning July 1, 2023 and ending December 31, 2023 was below the current 45 million VLOP threshold.

It starts with a wave