

Anti Hate

The new functionality to protect children from online hate speech.

Early parents warning as a key to preventing cyberbullying

Feb 2024

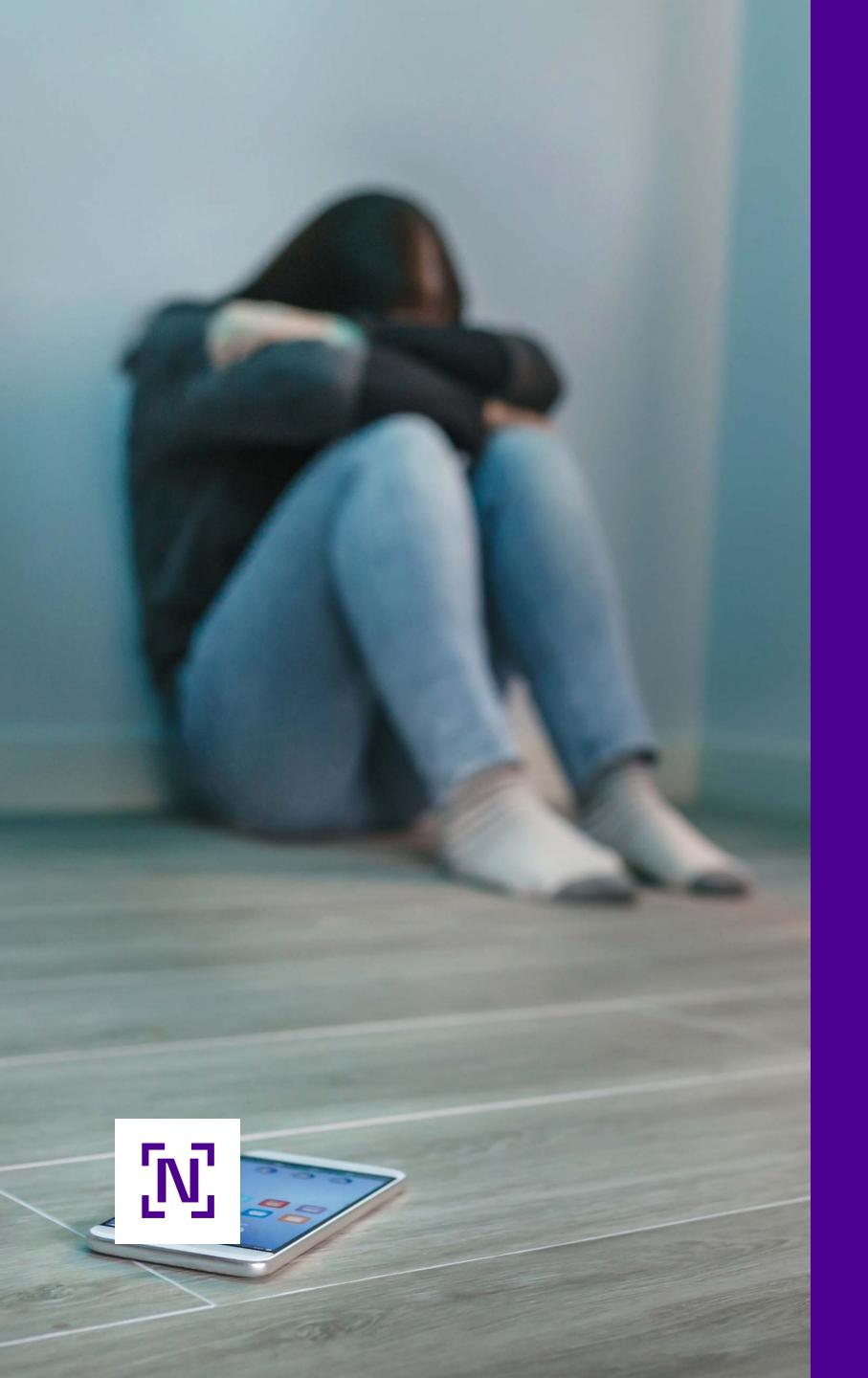






01 What is hate speech?

www.safekiddo.com



What is hate speech?

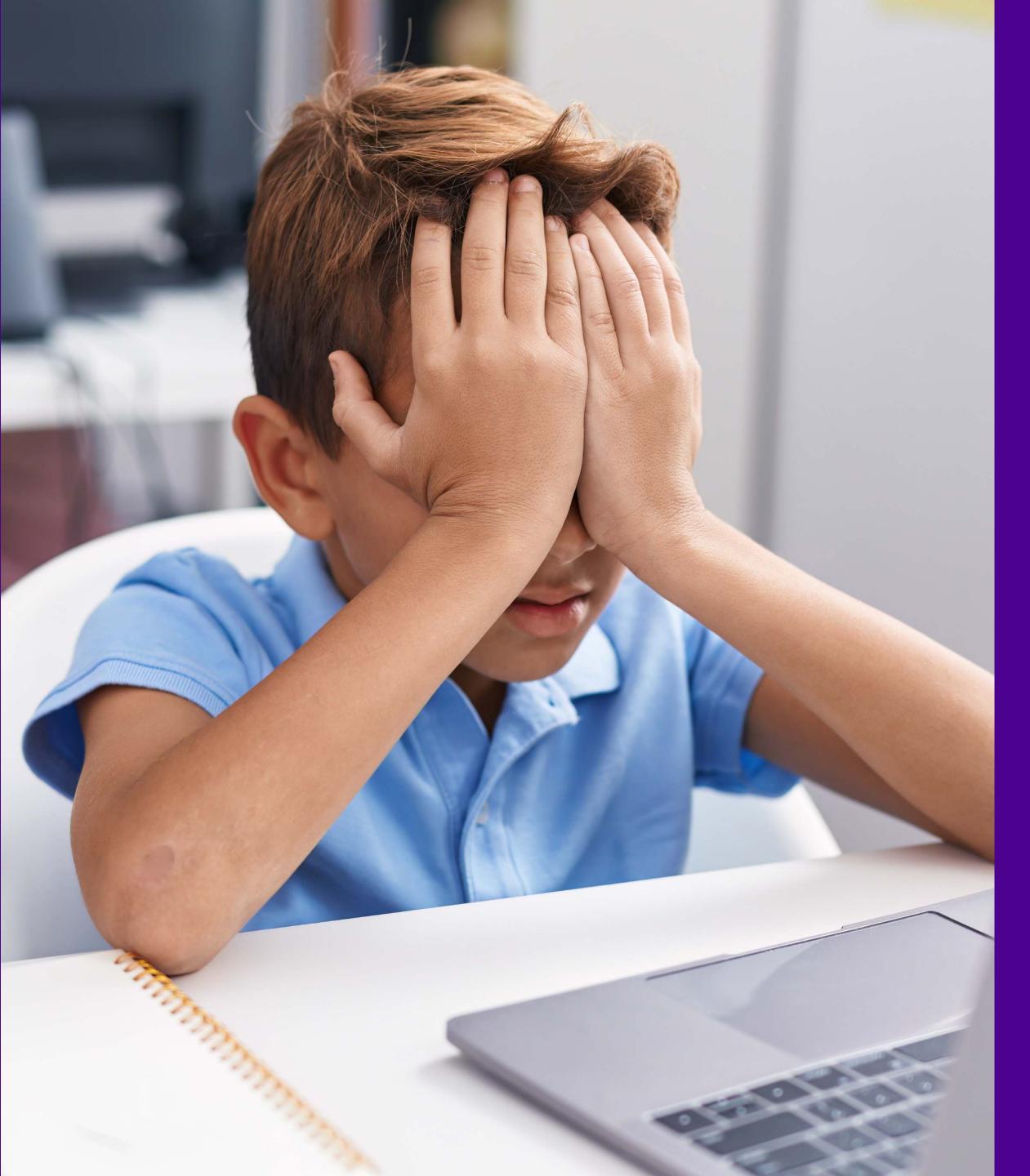
Hate speech, also known as online hatred or trolling, is a behavior that involves deliberately insulting, attacking, and criticizing other people on the internet.

Why is it dangerous?

Online hate speech is particularly dangerous due to its reach and anonymity. Hate speech can lead to emotional and even physical harm to victims, including depression, anxiety, suicide, and other forms of violence.

Since when does it exist?

The phenomenon of online hate speech has existed since the beginning of the internet. However, due to the development of technology and widespread use of social media, the problem of online hate speech has intensified in recent years.





02 The scale of hate speech on the Internet

www.safekiddo.com

mentions of hate speech per month*

*Source: monitoring of the internet and social media by IMM (Internet Monitoring and Media Analysis) for the period of 02.2014 - 02.2017 for the phrase 'hate speech'.





39% of young people aged 13-24 admitted to being a victim of online hate speech, and 45% of them experienced hate speech at school*

*According to a study conducted by the Give Children Strength Foundation in June 2021









03 The impact of hate speech on children

www.safekiddo.com

Hate speech can have a negative impact on children and young people. Here are some of the possible effects of hate speech on children:

Lowered self-esteem

Constant exposure to negative comments and attacks can lead to lowered self-esteem in children, which can affect their sense of self-worth and belief in themselves.

Stress and anxiety

Hate speech can trigger stress and anxiety in children, which can affect their overall emotional state and well-being.



Aggressive behavior

Children who are exposed to hate speech may start exhibiting aggressive behavior in response, which can escalate the situation and lead to the escalation of conflicts.



Online hate as a cause of suicide

Many studies indicate that young people are particularly vulnerable to the negative effects of online hate.

Therefore, it is very important to educate children and young people about the harmful effects of hate speech and provide them with emotional support when needed.

It is also important to monitor children's online activity to identify early signs of hate and prevent its further escalation.



Detecting hate speech at an early stage and parents' response are important.



Online hate can have a negative impact on a child, so it is important for parents to detect it early and take action.

Reacting to hate can prevent escalation and counter cyberbullying.

Discussions about online hate strengthen the bond between parents and their child.



Reacting to hate helps children learn respect, empathy, and tolerance.

Early detection of hate speech helps a child increase their selfawareness and understand the consequences of their online behavior.

Helping a child to cope with online hate supports healthy emotional and social development.



Early Warning System

Detecting hate speech at an early stage and parents responding to it is very important to prevent further escalation of the situation and provide emotional support to the child.

Remember

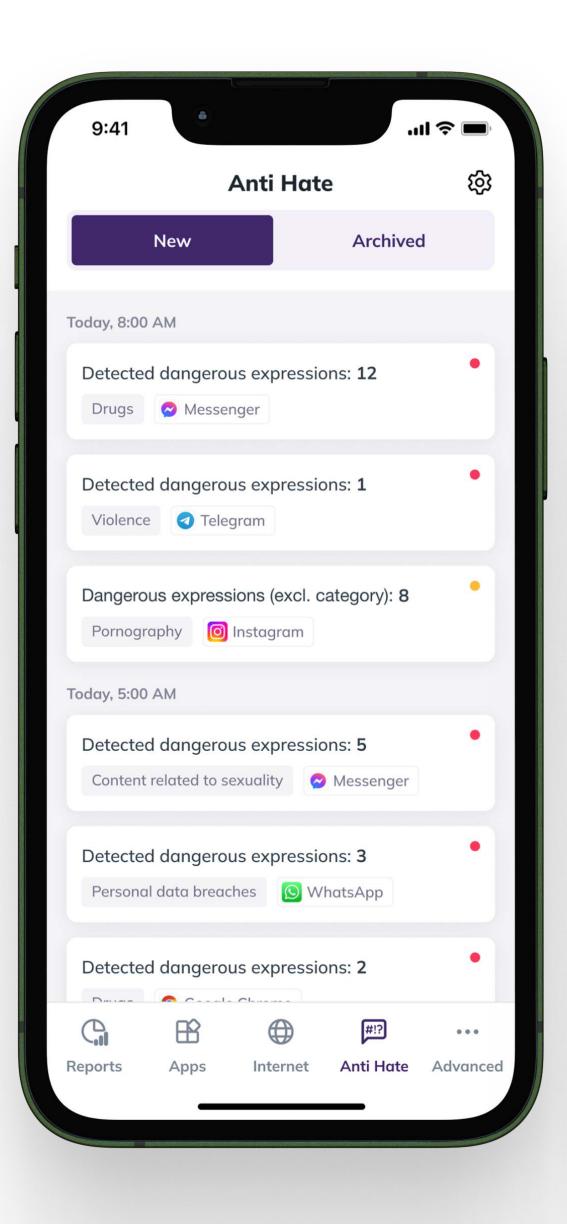
Preventing hate speech is much easier than dealing with its consequences. That's why it's so important for parents to monitor their children's online activity, educate them about the harm of hate speech, and encourage them to use the internet in a safe and responsible way.

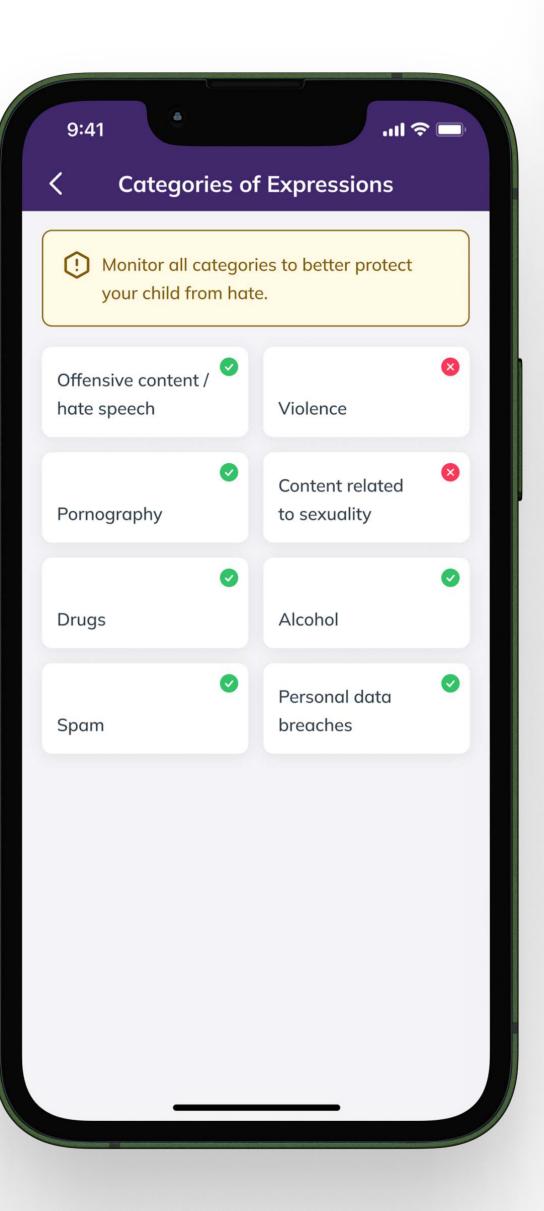


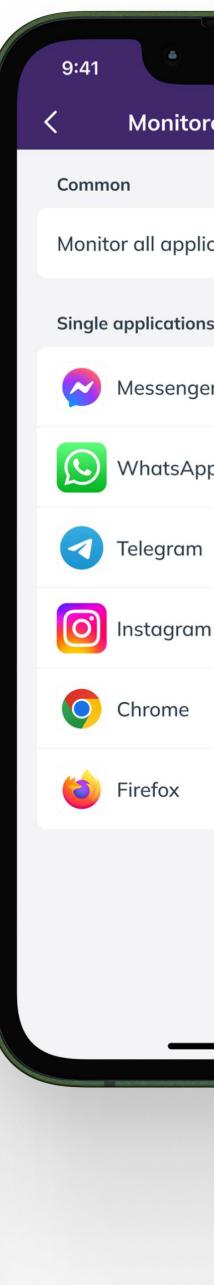


04 Anti Hate function as a solution

www.safekiddo.com







red application
cations
IS
er
р
n

The Anti Hate feature comes to help parents as an early warning system.

Its main goal is to verify:



Whether the child is not affected by prohibited content and whether they are not a victim of hate speech?



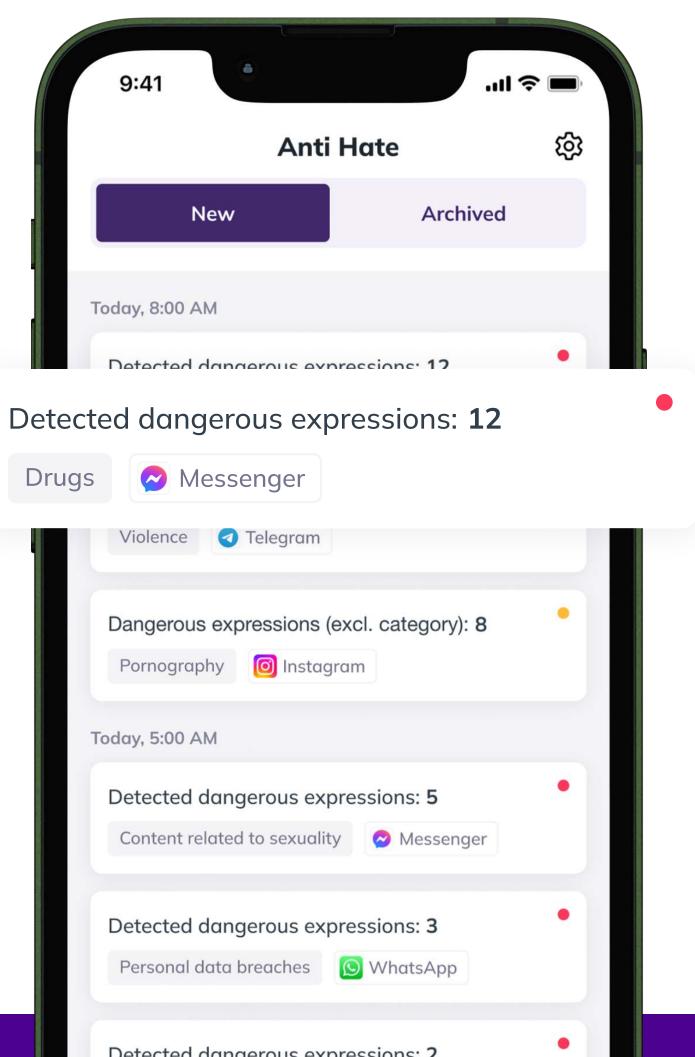
Whether the child is not a perpetrator of hate speech?

Discover the Anti Hate feature

To make your child even safer on the Internet, we have introduced a new feature - Anti Hate

Thanks to it, you can receive notifications about detecting expressions from specific categories in communication with peers as well as in phrases searched by your child in internet search engines.

Safekidde

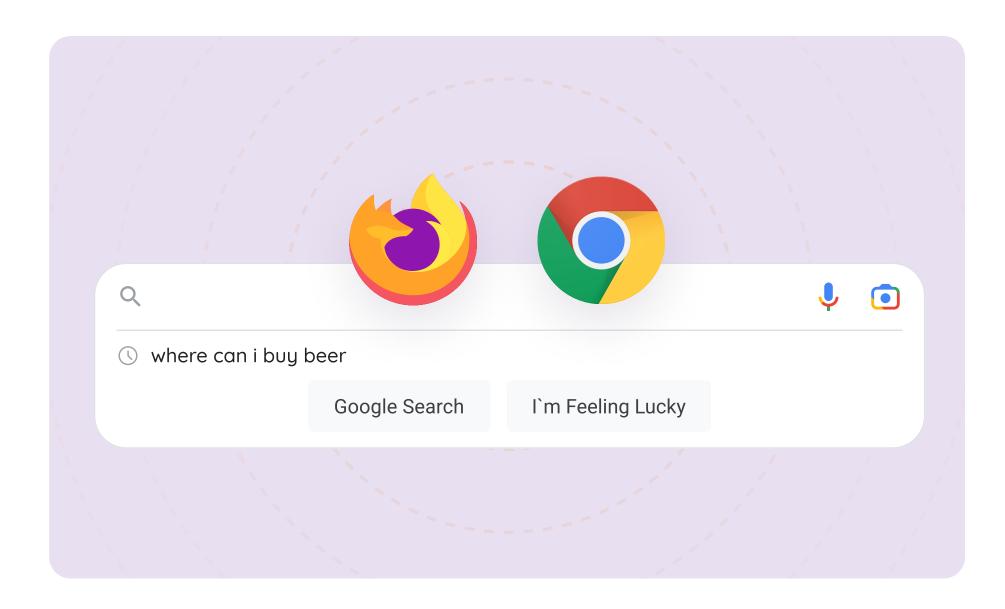


Social communication apps



If your child uses: Messenger, WhatsApp,Anti Hate will also monitor phrases searchedInstagram, Telegram, the feature will monitorby your child in the Chrome and Firefoxsent and received messages for hate speechinternet browsers.and other harmful content.and other harmful content.

Internet search engines





"New" tab

In this tab, you will see the latest notifications.



9:41	? ■
Anti Hate	ණ
New Archived	
Today, 8:00 AM	
Detected dangerous expressions: 12	•
Detected dangerous expressions: 12	
Drugs Messenger	
Dangerous expressions (excl. category): 8	•
Pornography 🞯 Instagram	
Today, 5:00 AM	_
Detected dangerous expressions: 5 Content related to sexuality OMessenger	•
incoderiger	
Detected dangerous expressions: 3	•





Archiving and deleting notifications

You manage received notifications easily.



9:41		·'II 🗢 🔲
An	ti Hate	ණ
New	Arch	ived
Today, 8:00 AM		
Detected dangerous e		•
xpressions: 1	↓ Archive	D elete
Dangerous expressions Pornography		8
Today, 5:00 AM		
he notification has b	een deleted	

Detected dangerous expressions: 3





"Archived" tab

This tab contains notifications automatically archived after 30 days or manually archived ones.



New Archived Today, 8:00 AM Detected dangerous expressions: 1 Drugs Messenger Detected dangerous expressions: 1 Personal data breaches SwhatsApp Cangerous expressions (excl. category): 1 Pornography Solution (Stagram)		Anti Hate	ණ
Detected dangerous expressions: 1 Drugs Messenger Detected dangerous expressions: 1 Personal data breaches SWhatsApp Dangerous expressions (excl. category): 1	New	A	rchived
Drugs Messenger Detected dangerous expressions: 1 Personal data breaches S WhatsApp	Today, 8:00 AM		
Personal data breaches S WhatsApp			•
Pornography 🞯 Instagram	angerous expres	ssions (excl. co	itegory): 1
	Pornography 🧿	Instagram	

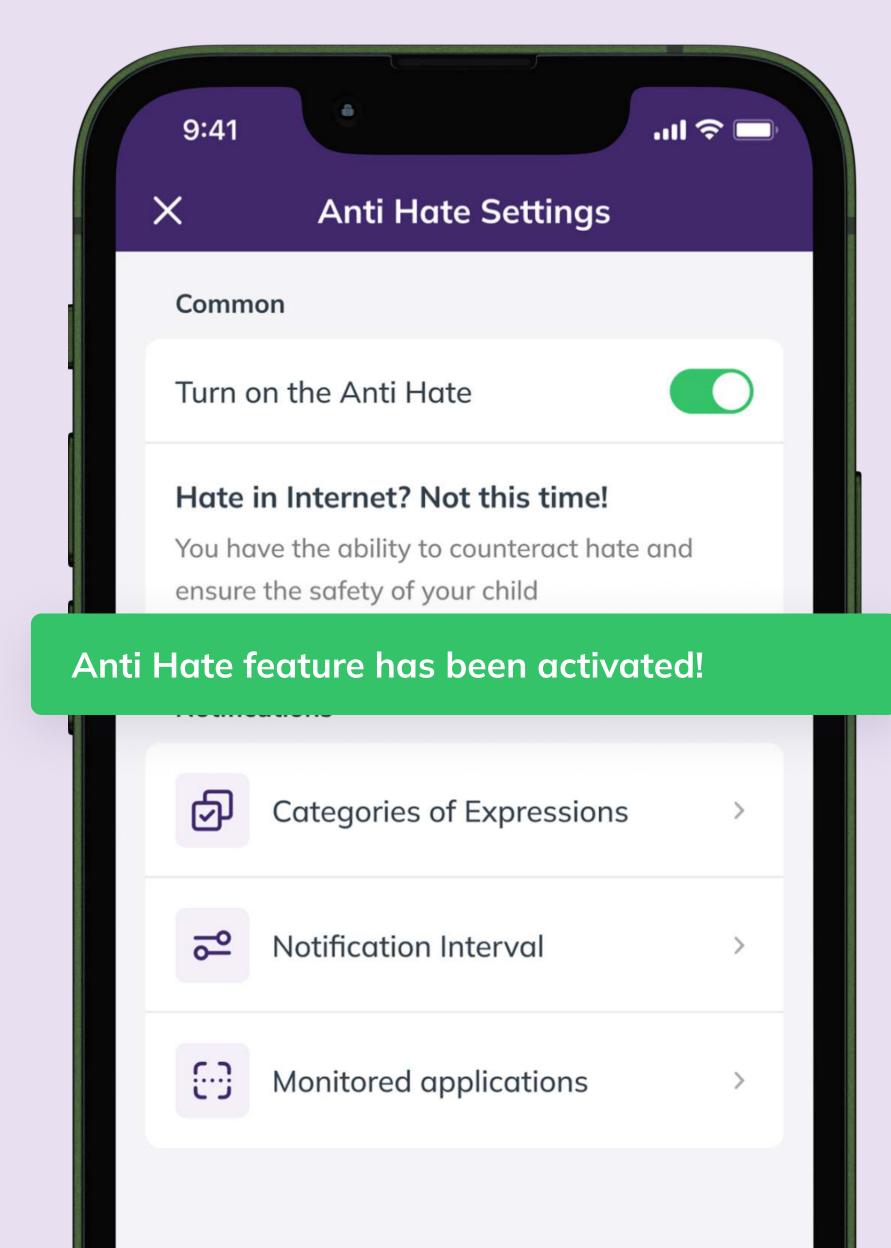




"Enable Anti Hate feature" switch.

Using this switch, you can activate or deactivate the feature.



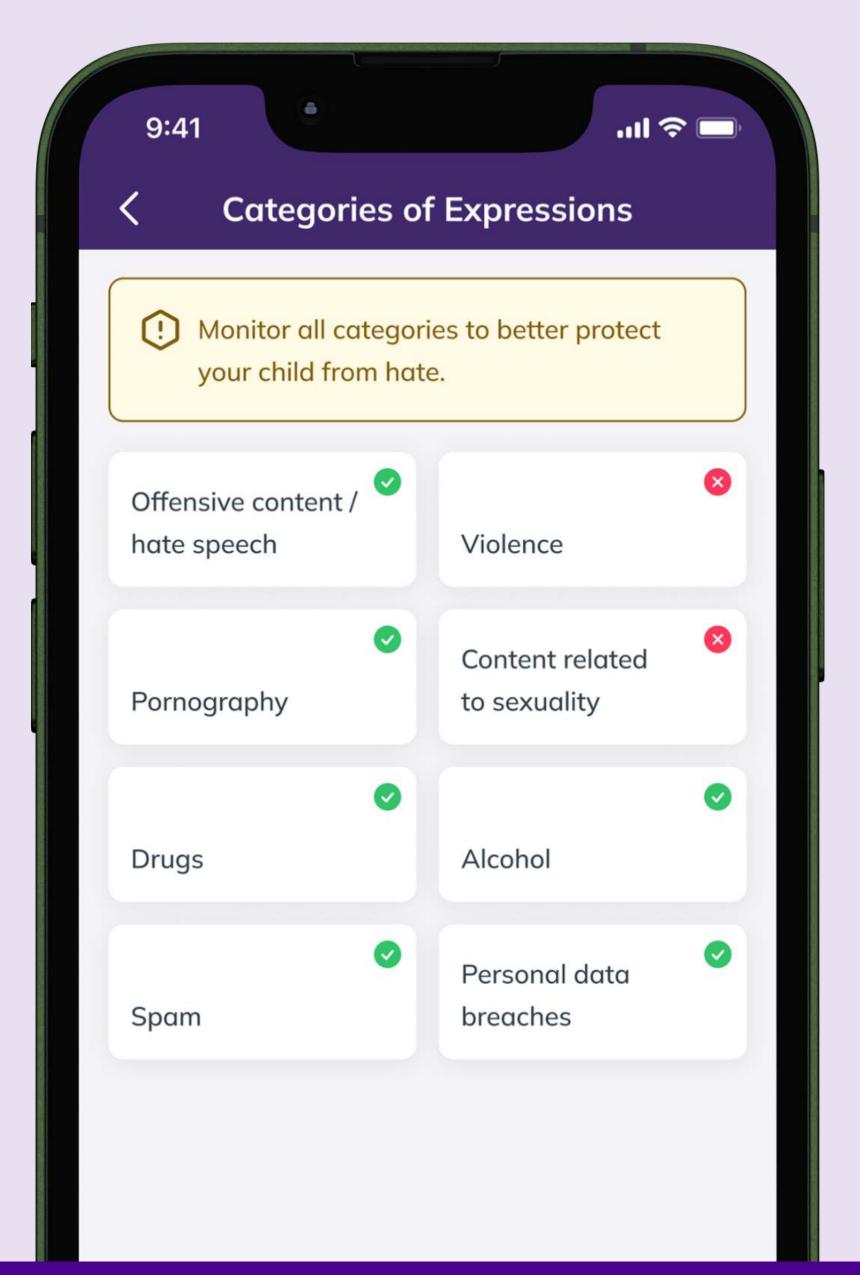




Categories of expressions

You choose the categories for which the app should send notifications upon detection.







Notification interval

You set the frequency at which you want to receive notifications



9:41			.ıll 🗢 🗖
<	Notification	Interval	
Send n	otifications		
◯ E	very 30 minutes		
⊖ e	very 1 hour		
• E	very 3 hours		
() E	very 5 hours		

Prevention is always easier than repair.

Once hate speech has appeared, the damage has already been done, and its consequences can be difficult to repair. That is why it is better to act proactively and try to prevent hate speech before it spreads - the Anti Hate feature is designed to help with this.



Lange August

Website







