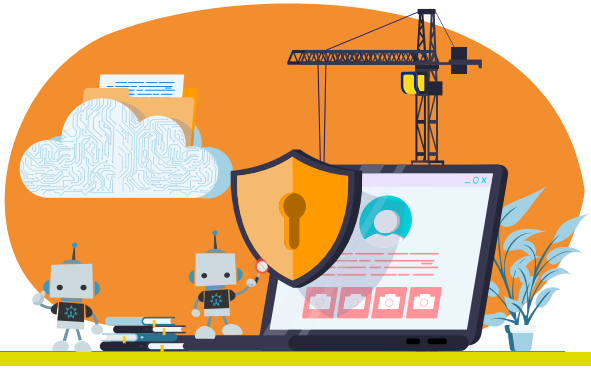
Digital Safety Guide

facebook | Jordan Open Source Association

Digital Safety Guide

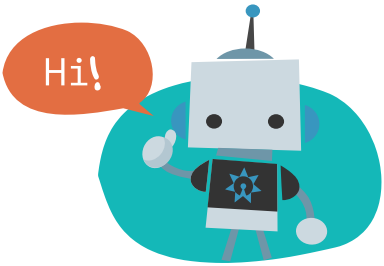
CONTENTS

Welcome to our Digital Safety Guide



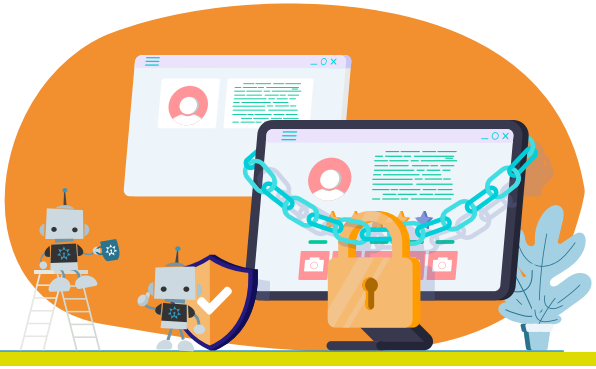
Terminology

Throughout this guide, we will be referring to "FB" as the company, and "Facebook" as the social media application. FB owns the following family of applications:

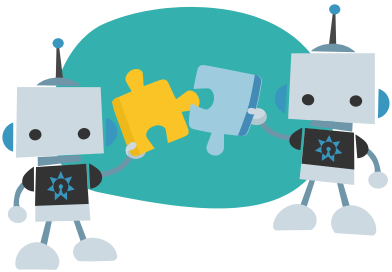


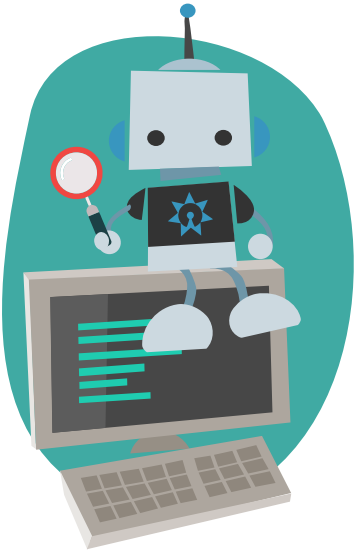
For more information about the company FB, please visit:
about.fb.com

Safety in Five Steps



User safety is paramount to FB. While the company does what it can to protect users on its family of applications, no platform is risk-free and completely safe. For that reason, it is important for you as a user to know what you can do to reduce the risk of falling victim to the most common online threats.

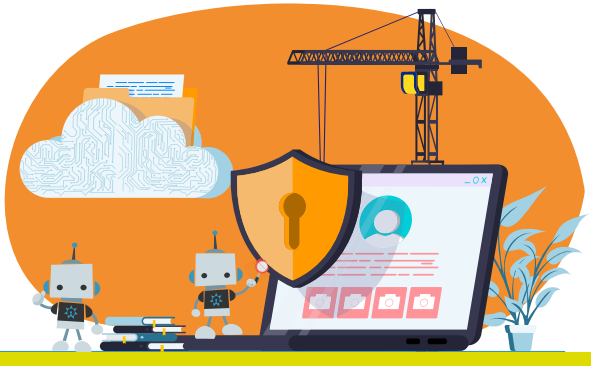




This guide provides tips to protect users from common threats on FB's family of apps. The tips are straightforward and don't require specialised knowledge to follow through. FB designed its platforms with a view to giving people power and control over their own experiences. It›s only a hope that, with the tips laid out in this guide, users will meet halfway to improve their safety online.

To help you better grasp the general idea of digital safety, let's entertain the door analogy.

Safety in Five Steps



1. Activate two-factor authentication (2fa)

Two-factor authentication is akin to protecting your house door with two non-identical locks instead of one. Everytime you want to login, you would need to enter your password, then a code that is sent to your phone, in order to access your account. This way, if someone figures out your password, they still cannot get full access to your account.

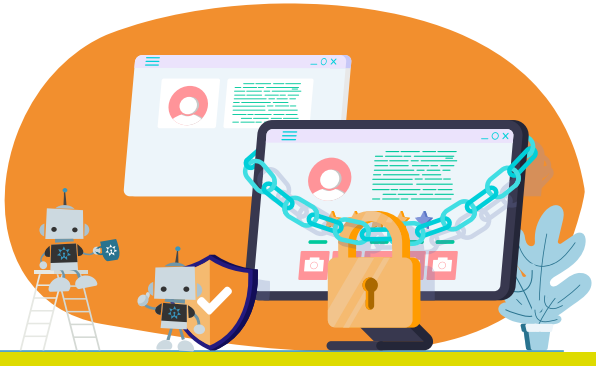
Here's how to activate two-factor authentication:

On Facebook:
Settings & Privacy > Settings > Security and Login > Two-Factor Authentication > Use two-factor authentication.

On Instagram:
Settings > Security > Two-Factor Authentication.

On WhatsApp:
Account > Two-step verification > Enable.

Safety in Five Steps



2. Activate login alerts

A login alert acts like a house alarm. If FB suspects that someone gained unauthorised access to your account, or someone logged in from an unrecognised device, you will receive an email notifying you of unusual activity.

Here's how to activate login alerts:

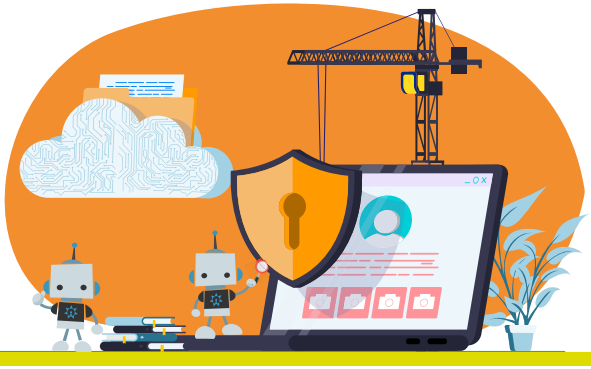
On Facebook:
Settings & Privacy > Settings > Security and Login > Setting Up Extra Security > Get alerts about unrecognized logins.

Here's how to verify login activity:

On Facebook:
go to: Settings > Security and Login > Where You're Logged In

On Instagram:
go to: Settings > Login Activity

Safety in Five Steps

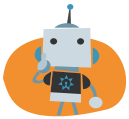


3. Use strong passwords

In the same way you wouldn't protect your door with a loose knot, you shouldn't protect your account with an easy-to-guess and easy-to-crack password.

Here's what you can do:

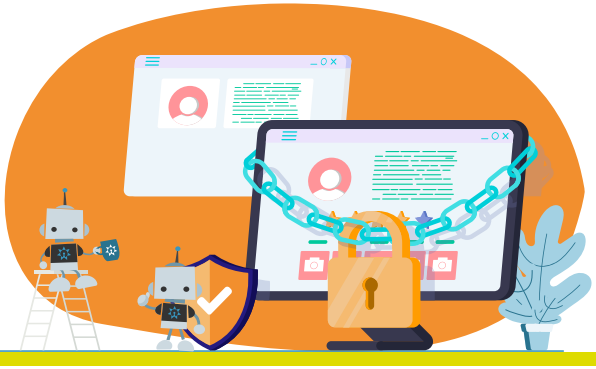
Your password should be longer than 8 characters, contain numbers, and characters.



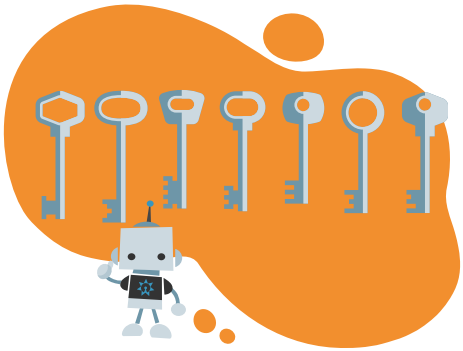
Make it easy for you to remember and hard for other people to guess.

Tip: don't choose the name of your favourite band but a lyric from your favourite song.

Safety in Five Steps



4. Use different passwords



Imagine using the same key to enter your car, your home, and your office. If a thief gets a hold of your key, they can open everything else. The same goes for your online accounts, protecting them with the same password will make a hacker's job easier.

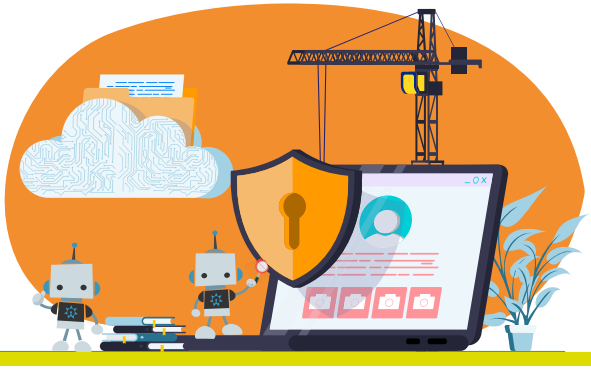
Here's what you can do:

Choose a password that you don't use anywhere else online.

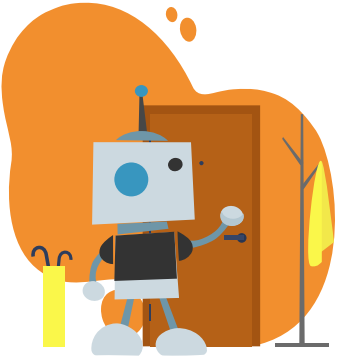
If you worry about remembering all your passwords, use a password manager. A password manager is an application which stores all your online credentials. It works across mobile and desktop.

Never share your password or write it down anywhere. If you have to, store it in a password manager.

Safety in Five Steps



5. Always confirm the source of an email



When someone rings the doorbell, you wouldn't blindly open the door but look out the peephole to make sure the person looks and sounds like a friend or relative you trust. Malicious actors will often reach out to you via a fake email pretending to be Facebook or Instagram. They use these emails for one of two purposes: steal your login details by asking you to fill a form, or download a virus onto your device by convincing you to click a link.

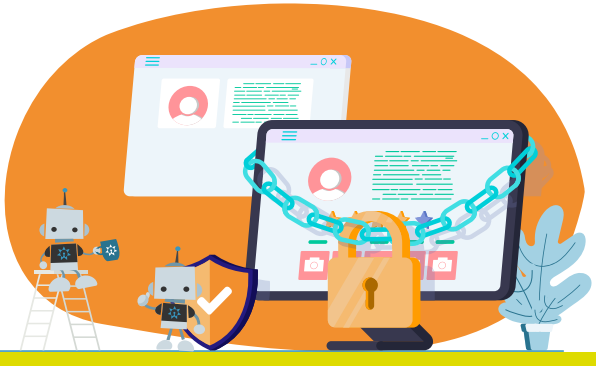
On Facebook:
Settings & Privacy > Settings > Security and Login > Advanced > See recent emails from Facebook.

On Instagram:
Settings > Security > Emails From Instagram.

Sources:

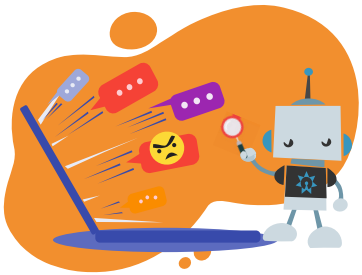
facebook.com/safety
https://www.facebook.com/about/basics/stay-safe-and-secure
https://www.facebook.com/about/basics/manage-your-privac
Instagram.com/safety
whatsapp.com/safetymessenger.com/privacy

Safety for Targeted Users



Targeted users include human rights defenders, activists, journalists, but also politicians and their staff, and media influencers. These individuals are more likely to be targeted for online attacks. This is because access to their accounts holds much more value to the attacker.

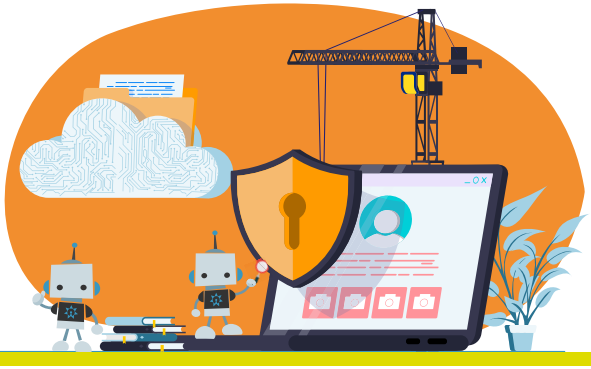
For example, a human-rights defender can be planning their next advocacy efforts – information that is valuable to a government authority. Moreover a journalist running an investigation on corruption could be targeted by a corporate entity.



Changing digital safety behaviour, even the simplest of habits, is not an easy sell. But it's worth noting that a recent study reported that most data breaches are due to human error and not security vulnerabilities. So behaviour change is crucial to protecting your online life.

As a targted user, you need to plan your defenses by identifying potential threats. Back to the door analogy.

Safety for Targted Users



1. Protect against account compromise

OTP authenticators > SMS authentication

Receiving a code via SMS is similar to hiding the second key under the doorway rug – it's almost in plain sight. Using a one-time password (OTP) authenticator application instead is a much safer alternative, akin to protecting the key in a safe.

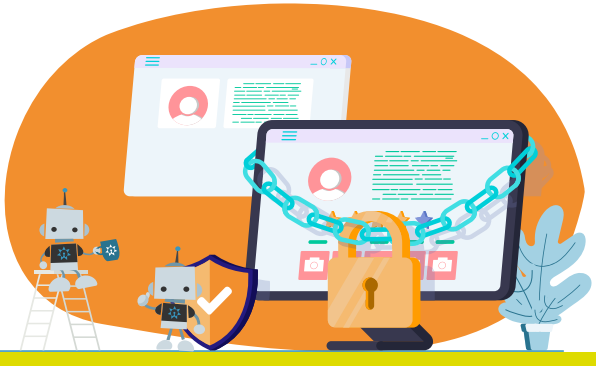
It's highly recommended to download a trusted OTP authenticator, preferably an open source one, and once you activate two-factor authentication on Facebook and Instagram ask to receive the code via an authenticator application.

'Brush your teeth, use a password manager'

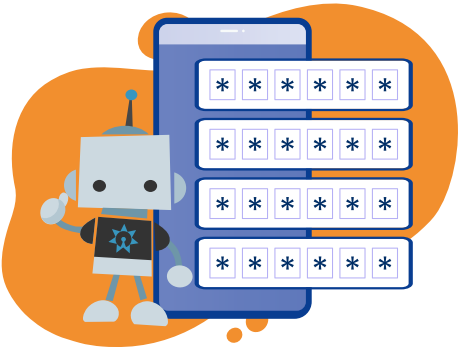
For targted users, using a password manager is as essential as brushing teeth. Attackers don't just guess passwords, they use software to brute-force or crack passwords by trying all possible options (like picking a lock). It's therefore essential for passwords to be strong, and unique so you don't use the same password anywhere else.

A password manager will help you store all your passwords securely, and generate passwords so you don't have to. You only need to memorise the master password to access the password manager.

Safety for Targted Users



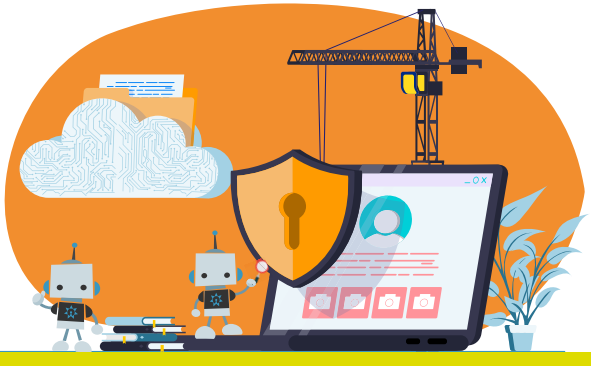
An additional benefit of password managers is storing notes securely. When would that be necessary you might ask? Upon setting up two-factor authentication FB apps will provide you with back-up codes in case you lose access to your phone or the app that generates OTPs for you. You can store these back-up codes securely in the password manager.



Zero-click exploits

Security vulnerabilities are unavoidable. Doors offer indispensable protection to houses, but they're still breakable. FB takes security seriously but when a breach is discovered, FB will notify you if your account has been a target and your account has been compromised.

Safety for Targted Users

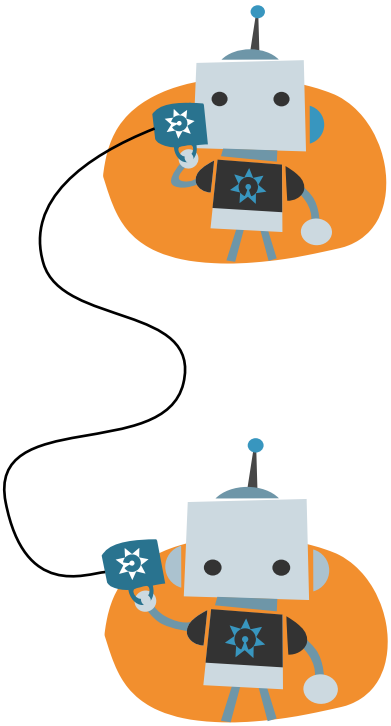


2. Protect against snooping and surveillance

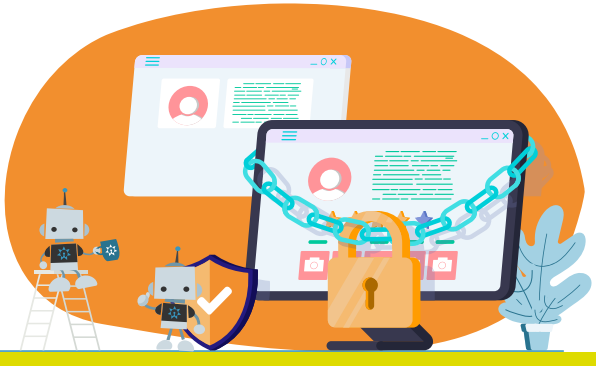
Private communication

Some of your most personal moments are shared on WhatsApp, which is why FB built end-to-end encryption into the latest versions of their app. When end-to-end encrypted, your messages and calls are secured so only you and the person you're communicating with can read or listen to them, and nobody in between, not even WhatsApp.

"End-to-end encryption" means that nobody can spy on a message once you send it and up until it reaches the person you intend to send it to. On WhatsApp, you can turn on "Disappearing messages" so that new messages in the chat are erased after seven days. From the WhatsApp chat, go to the contact's name, then Disappearing messages, "Continue", "On".



Safety for Targted Users



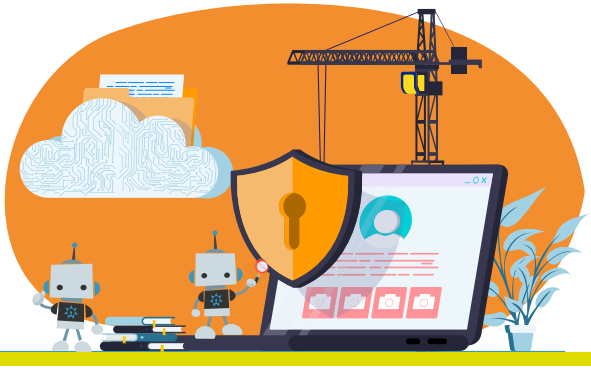
Messenger offers a feature called “Secret conversations” which you can turn on to talk to someone in an encrypted channel.

Conversation data back-up

Messenger conversations and Instagram direct messages are stored forever in FB servers. If you accidentally send sensitive information, you have 10 minutes to be able to delete the message by tapping on the message and selecting the delete option. On WhatsApp you have about an hour to delete the message.

Conversations on WhatsApp are stored on your device – locally. You can also opt to store your chats on a personal cloud provider such as iCloud or Google Drive. It’s important for you to know that the backup is not protected by encryption so you need to make sure that you’re also taking the appropriate measures to protect your cloud backup such as activating two-factor authentication and using a unique and strong password to protect your account.

Common Online Threats

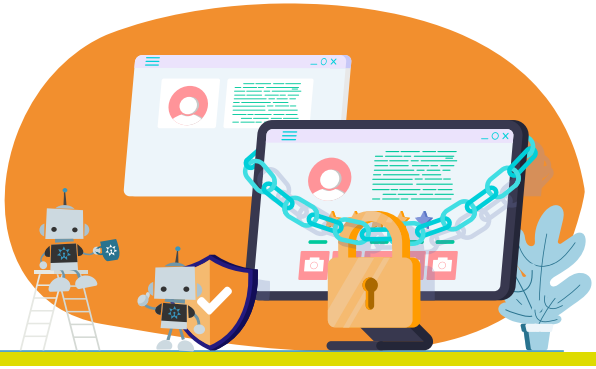


In order to recognise the significance of the safety measures set out above, it's important for users to learn about the most common threats they face in the online world.

1. Hacked and stolen accounts

Malicious actors – often called hackers – can gain unauthorised access to a user's account, often by knowing or guessing a user's password. While FB tries to keep hackers at bay, users also need to take the necessary precautions to protect their passwords and accounts.

Common Online Threats



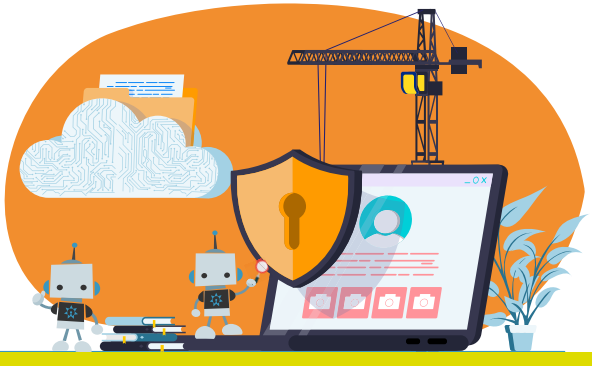
2. Extortion

When a user contacts another user and attempts to obtain money – or other valuables – by threatening to publicly reveal private or damaging information about the victim, this is called extortion. When the user threatens to reveal images of sexual nature, it's referred to as sextortion.

3. Social engineering

Social engineering is the act of manipulating a person into performing a certain action, or revealing personal often sensitive information. Attackers often develop a friendship or employ psychological tricks to convince a person to share photos, share passwords, etc. Instead of finding a security vulnerability in the software, an attacker could pose as an IT support person and pretend to have a noble aim of supporting users with protecting their accounts, and request for the user's password to accomplish this aim.

Common Online Threats



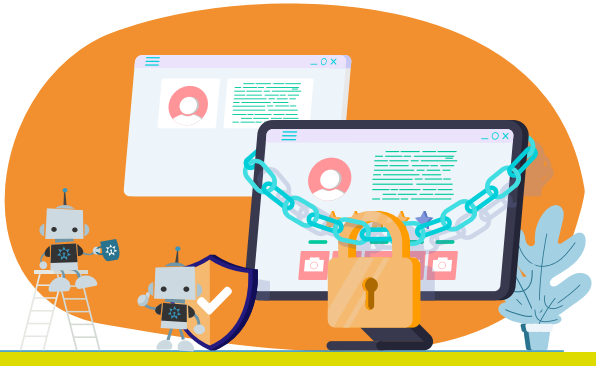
4. Phishing attacks

Phishing is when a user disguises themselves as a trustworthy entity and attempts to obtain (or "phish" for) sensitive information from another user such as their password or credit card information.

For example, someone pretending to be a local grocery business, sends you a message claiming they have a discount at their grocery store. They ask you to place your order via WhatsApp, and to also send your credit card number, expiry date, and security code in order to "process your payment".

Another example, you receive an email that appears to be from Facebook. The email says there's a threat to your account and that you should login right away to resolve it. You click the login link from the email and enter your email and password. There, you've handed your login credentials to the hacker.

Common Online Threats



5. Malicious links

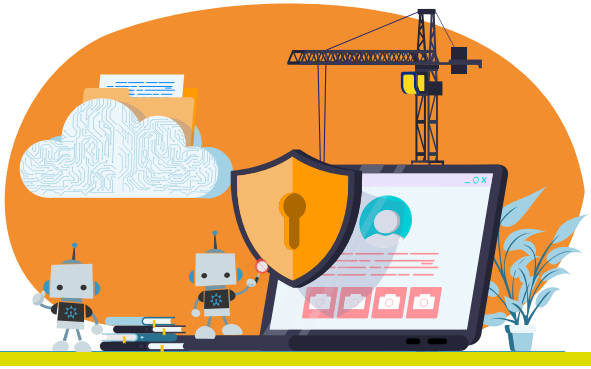
Users can receive links via messages or comments that appear to promote something genuine but they are in fact infected or phishing links. Upon clicking the link, a virus, or malware is downloaded on the user's device. The virus is often used to steal personal data. To check whether the link is infected, copy the URL, go to virustotal.com, and paste it on the search bar. The website will scan the URL and let you know whether the link is trustworthy.

6. Impersonations and fake accounts

Users often spot accounts that pretend to be someone they know, or themselves, or are simply fake. These fake accounts pose a threat as they are often used to gain other FB users' trust and gain followers and friends, and then use the fake account to threaten the user they're impersonating or extort them.

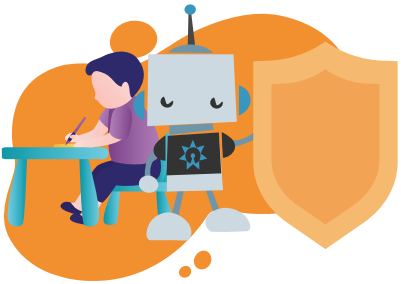
Laid out above are common and general online threats. There are threats that are more present between users in the MENA region and so it's important to spotlight them individually, define their symptoms, and possible ways to mitigate their severity. FB employs policies tools, and technology to combat all threats, and specific ones are highlighted below:

Common Online Threats



7. Child exploitation

What is child exploitation?
FB does not tolerate any behavior or content that exploits children online and it develops safety programs and educational resources to help make the internet a safer place for children. Users play an essential role in stopping child exploitation on FB. Child exploitation online can be categorized into three areas:

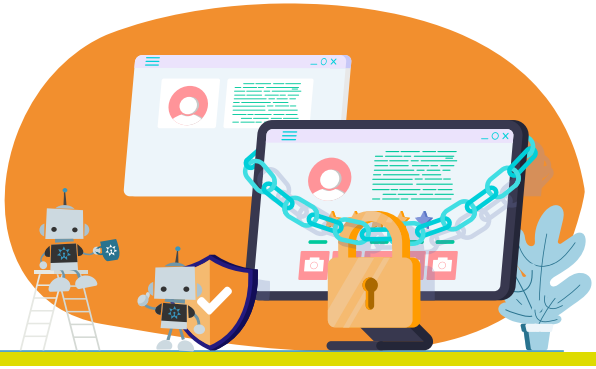


Inappropriate content, which means the dissemination and distribution of child exploitation material. Examples of inappropriate content would be creating a post that contains a picture of a child being physically abused, publishing a video of children being sexually exploited, sharing content containing any sort of child abuse, etc.

Inappropriate conduct, which means the act of enticing or forcing a child to take part in an inappropriate activity. Examples of inappropriate conduct would be to reply to a child's post with an inappropriate comment, stalking a child and making them feel unsafe, harassing a child, etc.

Inappropriate contact, means soliciting a child online often for the purpose of sexual gain. An example of inappropriate contact would be to seek contact with a child online with the purpose to abuse the child online or offline, and/or force them to participate in a sexual activity including sharing or viewing intimate images and/or videos.

Common Online Threats



How is FB fighting child exploitation?

One of FB's most important responsibilities is keeping children safe across its family of applications. FB takes a multilateral approach to combating child exploitation on its platform.

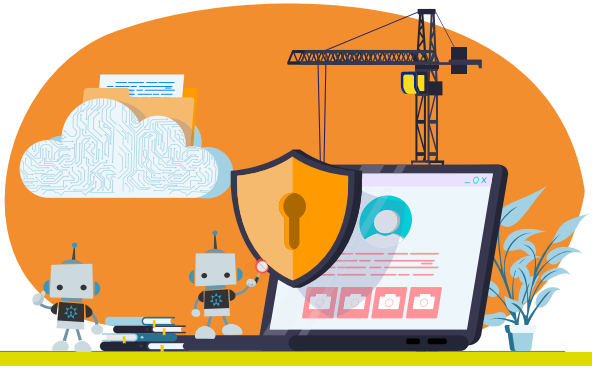
Policies | Polices to provide unique protections for minors:

FB platforms require everyone to be at least 13 years old before they can create an account (in some jurisdictions, this age limit may be higher). It violates our terms of services to provide a false age when creating an account. People can report an account belonging to someone under 13. FB does not allow content that sexually exploits or endangers children. When FB becomes aware of apparent child exploitation, the company reports it to the National Center for Missing and Exploited Children (NCMEC) in compliance with applicable law. FB also limits 13 to 18 years old interaction with other users and age gate the content they interact with.

Tools and Technology | Automatic removals:

FB has designed their platforms to give people control over their own experiences — control over what they share, who they share it with, the content they see and experience, and who can contact them. These tools empower individuals to protect themselves against unwanted content, unwanted contact, and bullying and harassment online. When it comes to users aged 13-18, FB takes extra precautions. The company has designed

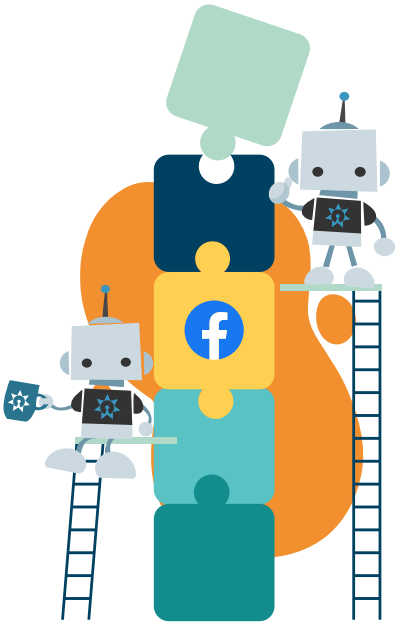
Common Online Threats



many of its features to remind them who they're sharing with and to limit interactions with strangers.

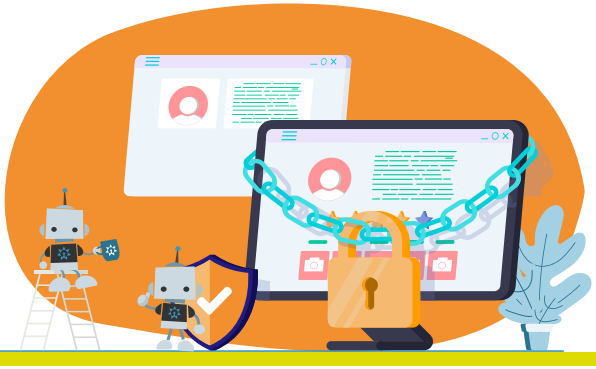
FB is utilising machine learning technology to automatically identify child exploitative content, and proactively removes them from their platform. FB also adds confirmed child exploitative material in a shared image bank so material that matches images already in the bank can be automatically removed

Partnerships | Partnership with Experts:
FB does not tolerate any behavior or content that exploits children online and develops safety programs and educational resources with more than 400 organizations around the world to help make the internet a safer place for children.



FB's work has included using photo-matching technology to stop people from sharing known child exploitation images, reporting violations to the National Center for Missing and Exploited Children (NCMEC), NCMEC works with law enforcement agencies around the world to help victims, and helping the organization develop new software to help prioritize the reports it shares with law enforcement in order to address the most serious cases first.

Common Online Threats



8. Bullying and Harassment

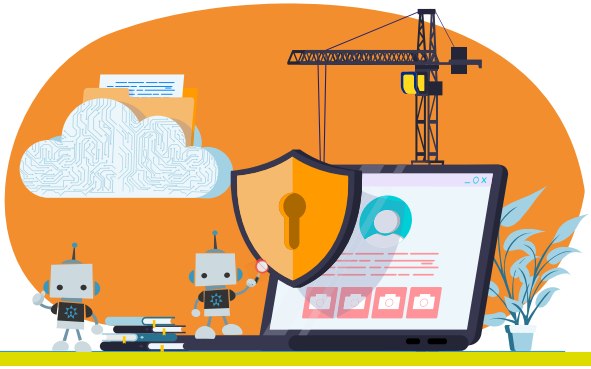
What is bullying and harassment?
As a user on one of FB's family of applications it's important to understand what constitutes bullying. If you're unsure if something's meant to be a joke or is just plain bullying, the question to ask yourself is whether the behaviour hurts you. If the answer is yes, then it's behaviour that constitutes bullying.

Bullying is the act of repetitively tormenting, bothering, and annoying a person. This can also consist of humiliating, and demeaning someone.

Harassment is when the hurtful behaviour is discriminatory in nature and is based on someone's race, religion, gender, sexual orientation, etc.

Examples of cyberbullying can include:
Spreading rumours and secrets about someone, often with the aim of humiliating them
e.g. "you sleep with all the men in your workplace"Sending hurtful messages or threats
e.g. "you're stupid", "you deserve to die"
Pranking someone via video or voice call
Sharing embarrassing photos of a friend with the aim to ridicule them

Common Online Threats



How is FB stopping bullying and harassment?

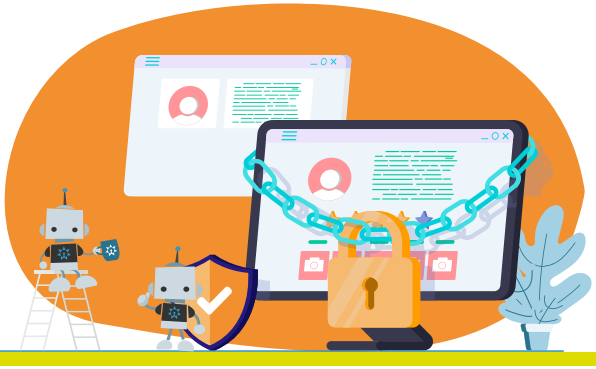
Tools | Giving users more control:

FB has rolled out features to give users control over their online presence. Users on Facebook and Instagram can remove offensive and hurtful comments, report a user with ease, and even report users or comments on behalf of friends or family members being bullied or harassed.

Technology | Notify users when they write offensive captions:

On Instagram when a user attempts to post an image with a caption intending to offend, Instagram automatically detects the text caption as potentially offensive and notifies the user. The notification is a heads-up for the user that the caption resembles other content that has been reported, and that their account might be breaking community guidelines.

Common Online Threats



9. Non-Consensual Sharing of Intimate Images (NCII)

What is NCII?

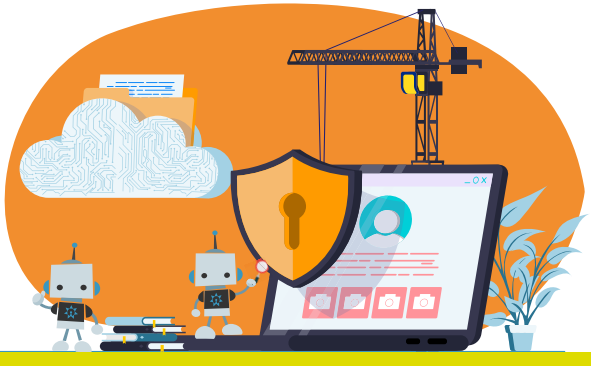
Non-Consensual Sharing of Intimate Images (NCII for short) is when someone's intimate images are shared without their permission. This is also referred to as revenge porn.

When someone threatens to share intimate images unless the victim agrees to do something in return such as engaging in a sexual activity or sending more intimate images, it's categorised as sexual extortion or sextortion.

How is FB stopping NCII?

Technology | Photo-matching technology:
FB leverages machine learning and artificial intelligence technology to automatically find and detect nude images, near nude images, and videos that are shared without consent on Facebook and Instagram.

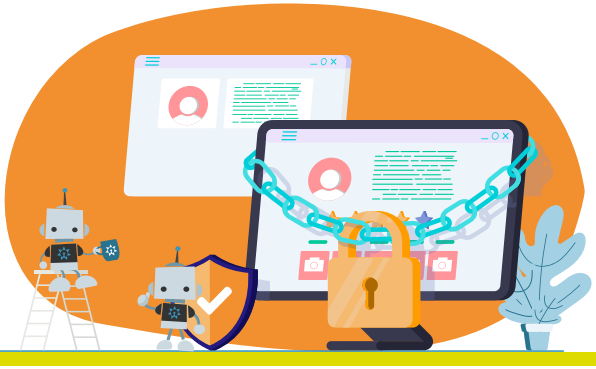
Common Online Threats



Tools | Reporting:
On Facebook, Instagram or Messenger, you can report when someone shares your intimate images without your consent or is threatening to do so by reporting them. You can learn how to report things on Facebook and Instagram (you can also learn how to report messages on Instagram). On Facebook, you also can do this by using the "Report" link that appears when you tap on the downward arrow or "…" next to the post.

Tools | Photo bank:
If you are concerned that someone may share your intimate images on Facebook, Instagram or Messenger but they haven't done so yet, and you have access to the images, you can contact one of our trusted partners here to help you prevent anyone from sharing the images.

Common Online Threats

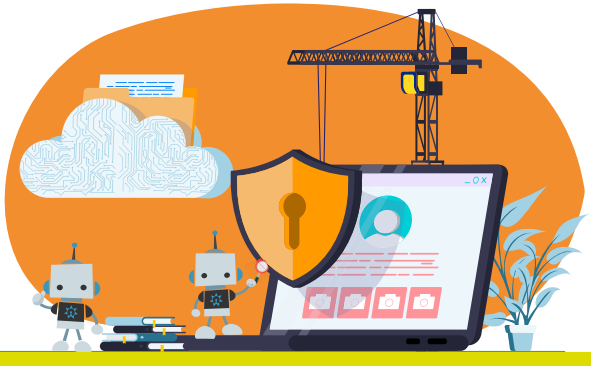


10. Gender-Based Violence (GBV)

What is GBV?
Gender-based violence (or GBV for short) is threats and violent behaviour targeted at women online. This can take the form of: cyberstalking, (s)extortion, doxing, cyberbullying, harassment, hate speech, shaming, and more.

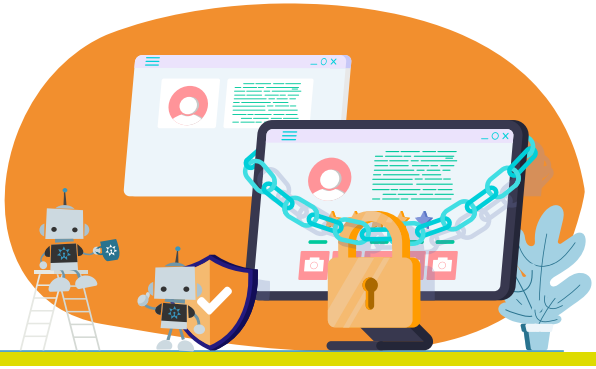
How is FB stopping GBV?
Rules | Pragmatic policies:
FB has rewritten its policies with regards to hate speech for example, in consultation with women's safety experts. The policies are clearer and more pragmatic and that take into account the disproportionate amount of violence women receive online.

Remember the six "R"s



If you find yourself in a situation where someone is bullying or harassing you, threatening you, sharing images without your consent, being violent towards you, remember the six "R"s:

1. Don't **R**etaliate: attackers tend to look for a reaction from their victims..
2. **R**ecord: take screenshots of attacks and save them, they might be useful.
3. **R**eport: report the person's account, and report offensive posts and comments.
4. **R**emove: unfriend, or unfollow the person.
5. **R**estrain: block the person's account so they can't interact with yours.
6. **R**each out: if the threat persists make sure to tell someone you trust. If it's serious, inform local police right away.



Sources

Child exploitation:

https://www.unicef.org/media/66616/file/Industry-Guidelines-for-Online-ChildProtection.pdf
https://www.facebook.com/communitystandards/child_nudity_sexual_exploitation
https://learning.nspcc.org.uk/child-abuse-and-neglect/child-sexual-exploitation
https://www.unicef.org/csr/files/pillars.pdf
https://about.fb.com/news/06/2020/fighting-child-exploitation-online/
https://www.facebook.com/safety
https://www.ceop.police.uk/Safety-Centre/what-is-online-child-sexual-abuse/
https://about.fb.com/news/10/2018/fighting-child-exploitation/

Bullying and harassment:

https://www.unicef.org/end-violence/how-to-stop-cyberbullying
https://www.facebook.com/safety/bullying/
https://about.fb.com/news/10/2018/protecting-people-from-bullying/
https://about.fb.com/news/12/2019/our-progress-on-leading-the-fight-against-online-bullying/
https://www.facebook.com/help/116326365118751?_rdc=2&_rdr

Non-consensual sharing of intimate images (NCII):

https://www.facebook.com/safety/notwithoutmyconsent
https://www.facebook.com/safety/StopSextortion/
https://about.fb.com/news/03/2019/detecting-non-consensual-intimate-images/
https://www.europol.europa.eu/activities-services/public-awareness-and-prevention-guides/are-you-victim-get-help-report-it-we-are-here

Gender-based violence:

https://about.fb.com/news/10/2019/inside-feed-womens-safety/
https://about.fb.com/news/06/2017/giving-people-more-control-over-their-facebook-profile-picture/

facebook

Jordan Open Source Association

facebook.com/safety

https://jordanopensource.org