Meta

Best practices for politicians: www.facebook.com/gpa
Concierge service: www.facebook.com/gpa/help

Security (Instagram and Facebook Messenger)

Account Compromise – How does it happen?

Malicious actors typically gain access to a person's account by using their compromised credentials, mainly email and password;

- Phishing Scams
- Weak password
- Didn't log out

Phishing Scam

Phishing is an online scam where someone impersonates or pretends to be a legitimate organization to gain sensitive information.

*Remember: Facebook or Instagram will **NEVER** send you a private message (DM)

Phishing Scams – So How do you know?

TIPS

- Facebook & Instagram will never direct message you about account violations.

- Facebook & Instagram will never direct message you asking for personal information (like passwords).

- Always look at the sending address:
  - if it looks off at all (i.e. @gmail), it's probably fake.
  - if there are misspellings, it's probably fake.

- If you receive links from strange accounts or people you don't know - don't click on them.

- Anything that is overtly urgent is likely fake.

Account Compromised

If you think your account may have been compromised but you can still log into your account..
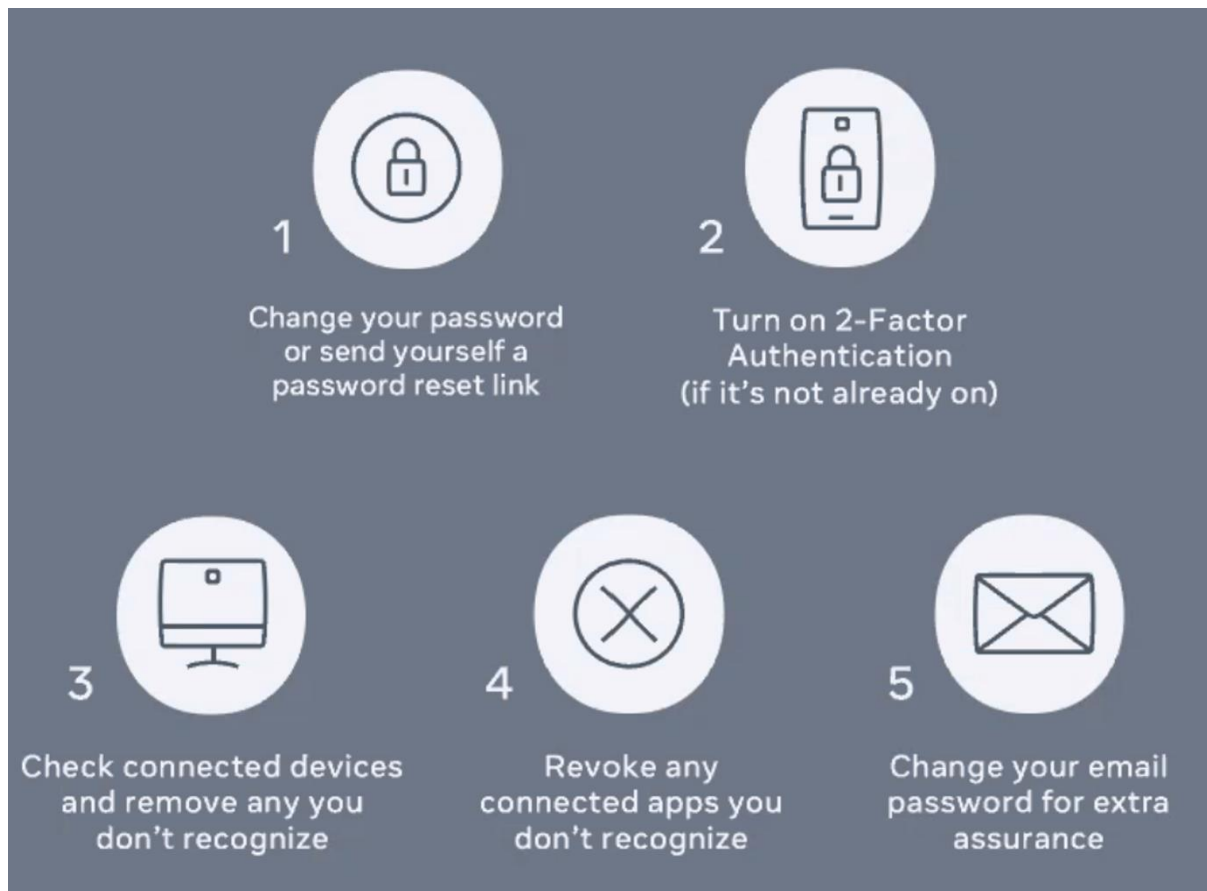
2 Factor Authentication

The best practice

WHAT IS 2-FACTOR AUTHENTICATION?

2-factor authentication is a security feature that helps safeguard your account by requiring something you know (your password) and something you have (your mobile phone or authenticator app) every time you log into your Account from an unknown device.

It's the single most effective step to protect your account from hackers.

Safety

Harassment

Meta's bullying and harassment policies are designed to protect people from abuse but they do permit open and critical discussion of people who are featured in the news or have a large public audience based on their profession or chosen activities.
Behaviour that falls within the category of harassment includes;

- Repeated unwanted interactions
- Death threats
- Revenge Porn
- Hate Speech
- Doxxing - Doxxing is **a form of cyberbullying that uses sensitive or secret information, statements, or records for the harassment, exposure, financial harm, or other exploitation of targeted individuals**.

Instagram - How to manage what people can say to you.

Instagram – Hidden words

WHERE: SETTINGS > PRIVACY > Hidden Words

OFFENSIVE CONTENT
HIDE COMMENTS: Hide comments that you don't want to see.
HIDE MESSAGE REQUESTS: Hide message requests that you don't want to see.

CUSTOM WORD LIST
ADD TO LIST: Ability to create a custom list of words to be filtered out.
HIDE COMMENTS: Applying custom filter to comments.
HIDE MESSAGE REQUESTS: Applying custom filter to message requests.

Manage Interactions with People in Bulk

Instagram – Limits

WHERE: SETTINGS > PRIVACY > Limits

LIMITS: Limits is an account-level tool designed to help you manage intense episodes of interactions by hiding unwanted comments and DM requests with a few taps.

WHO TO LIMIT: Ability to hide interactions from accounts that do not follow you or recently started following you.

HOW LONG TO LIMIT: Ability to set how long the Limits feature is in place.
Limits can be turned off at any-time or be turned off at a specific time.

Manage How people can contact you

Instagram – Reporting potentially abusive DMs

REPORTING: Reporting potentially violating content is crucial towards properly actioning potentially violating content.

REPORTING DIRECT MESSAGES: There are two-ways to report potentially abusive DMs specifically:

1. Holding your finger down on the offensive message will populate a reporting button.
2. Tapping the sender's name, scrolling down to Report>It's Inappropriate>This Chat.

Manage Potentially Abusive Content

Instagram – Reporting bullying & harassment

REPORTING: Reporting potentially violating content is crucial towards properly actioning potentially violating content.

- REPORTING ACCOUNTS: Tap the Account then the ellipsis (3-dots) in the top-right corner. Tap "Report" and then "It's Inappropriate" and then the reason.
- **REPORTING POSTS:** Tap the ellipsis (3-dots) in the top-right corner of the post. Tap "Report" and then "It's Inappropriate" and then the reason.
- **REPORTING COMMENTS:** Swipe right on the comment and tap the "!" and then "Report This Comment, "It's Inappropriate" and then the reason.

Manage How People Can Contact You

Instagram – Manage who can DM you

WHERE: SETTINGS > PRIVACY > Messages

POTENTIAL CONNECTIONS: Ability to choose which inbox potential connections (i.e. followers) messages go.

OTHER PEOPLE: Ability to control where messages from "other people" on Instagram – those not following you – will go.

GROUP SETTINGS: Ability to change who can add you to group message threads.

Manage People's Comments

Instagram – Responsive moderation features

WHERE: VIEW ALL COMMENTS > Ellipsis (Top Right)

COMMENT CONTROLS: Ability to turn off comments or add words to your manual filter.

RESTRICT: Ability to change the visibility of interactions between your account and another to only you and that account.

MUTE: Hide posts and interaction between you and another account.

BLOCK: Ability to limit all interactions between your account and another.

Manage Who Can Interact with You

Instagram – Choose who can interact with your content

WHERE: SETTINGS > PRIVACY > ELLIPSIS (...)

COMMENTS: Ability to choose who can comment on your account.

POSTS: Here you can find the ability to choose who can tag you, as well as manually approve tags.

MENTIONS: Ability to choose who can @mention you.

MESSAGES: Ability to limit who can message you.

Facebook - Bullying and Harassment

Comment moderation automatically filters the words you add to it. You can create a list of all the words you don't want to see.

Manage what people can say to me

Facebook – Proactive page-level settings

WHERE: SETTINGS > PRIVACY > Public Posts

BLOCKLIST FILTER: Will now identify and automatically filter variations of words including misspellings and numbers (i.e. taco = tac0).

MOST RELEVANT COMMENT VIEW: Filters high-quality comments that are relevant to you and surface them to the top.

PROFANITY FILTER: Automatically filters curse and derogatory content from your page based on preferences.

PAGE MODERATION: Ability to hide specific words, phrases or emojis.

COUNTRY RESTRICTIONS: Ability to restrict Page visibility based on people's country origin.

AGE RESTRICTIONS: Ability to restrict Page visibility based on people's age.

Receiving comments you don't like or want to take action about? Facebook provides a number of tools to help with abuse.

Manage who can interact with me

Facebook - Moderating in comments

WHERE: POST > Comments

HIDE: Ability to filter the visibility of a comment to the poster and their friends.

DELETING COMMENTS: Ability to delete a person's comment from your post.

BLOCK: Prevents someone from tagging, finding, messaging or engaging with your page.

Manage my time on moderating

Facebook - Moderation assist

A new feature to help reduce the amount of time spent on moderation by allowing you and your Page admins to set criteria that automatically hides incoming comments that you do not want on your Page posts.

WHERE: Page feed* > PROFESSIONAL DASHBOARD > MODERATION ASSIST

HIDE INCOMING COMMENT IF: Ability to add criteria on what type of comments to hide based on either the comment or the author, such as comments with links, from profiles with no profile picture, from new followers and more. After you add a criteria, all new comments that match it are hidden.

ACTIVITY LOG: Ability to add, edit or delete criteria and also review hidden comments.

*Moderation Assist is only available in new Pages experience and to a select group of users at this time. The product will be available to all users as we continue to test and expand.

Manage potentially abusive content

Facebook - Reporting bullying & harassment

REPORTING: Reporting potentially violating content is crucial towards properly actioning potentially violating content.

- **REPORTING ACCOUNTS:** Tap the Account, then the ellipsis (3-dots) in the top-right corner, then "Find Support or Report Profile" and then the reason.
- **REPORTING POSTS:** Tap the ellipsis (3-dots) in the top-right corner of the post, then "Find Support or Report Post" and then the reason.
- **REPORTING COMMENTS:** Tap the comment and select "Find Support or Report Comment" and select the reason.
- **REPORTING MESSAGES:** Tap the user's name at the top of the conversation, then "Something's Wrong" and then the reason.

Verification

Public verification forms

Both Facebook and Instagram have public application forms whereby any user is able to submit a verification request.

Remember:

Facebook and Instagram will never request payment for verification or reach out to ask you to confirm your verification.

Additional resources

☐ How to hide comments or message requests that you don't want to see on Instagram: https://help.instagram.com/700284123459336

☐ How to moderate your Facebook Page:

https://www.facebook.com/formedia/blog/moderating-your-facebook-page

☐ Reporting content on Facebook:

https://www.facebook.com/help/1380418588640631

☐ Reporting content on Instagram: https://help.instagram.com/2922067214679225

☐ Safety Centre: https://www.facebook.com/safety/

Blue-tick verification

On Facebook: https://www.facebook.com/help/1288173394636262

On Instagram: https://help.instagram.com/398038890351915

Ads

Ads policies: https://www.facebook.com/policies_center/ads

Requirements for ads about Elections or Politics here

Setting up a campaign here

About boosted posts: here

Content Policies

Community Standards: https://transparency.fb.com/en-gb/policies/

Policy Enforcement: https://transparency.fb.com/en-gb/enforcement/