



Addressing
hate speech
in the media:
the role of regulatory
authorities and
the judiciary

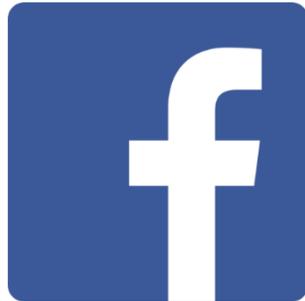
INTERNATIONAL
CONFERENCE

Organised by the Council of
Europe in partnership with
the Croatian Agency for
Electronic Media

6-7 November 2018
Zagreb, Croatia

Facebook Hate Speech Deck

Gabriella Cseh
Public Policy Director, CEE



Gabriella Cseh
Public Policy Director, CEE



Facebook Overview



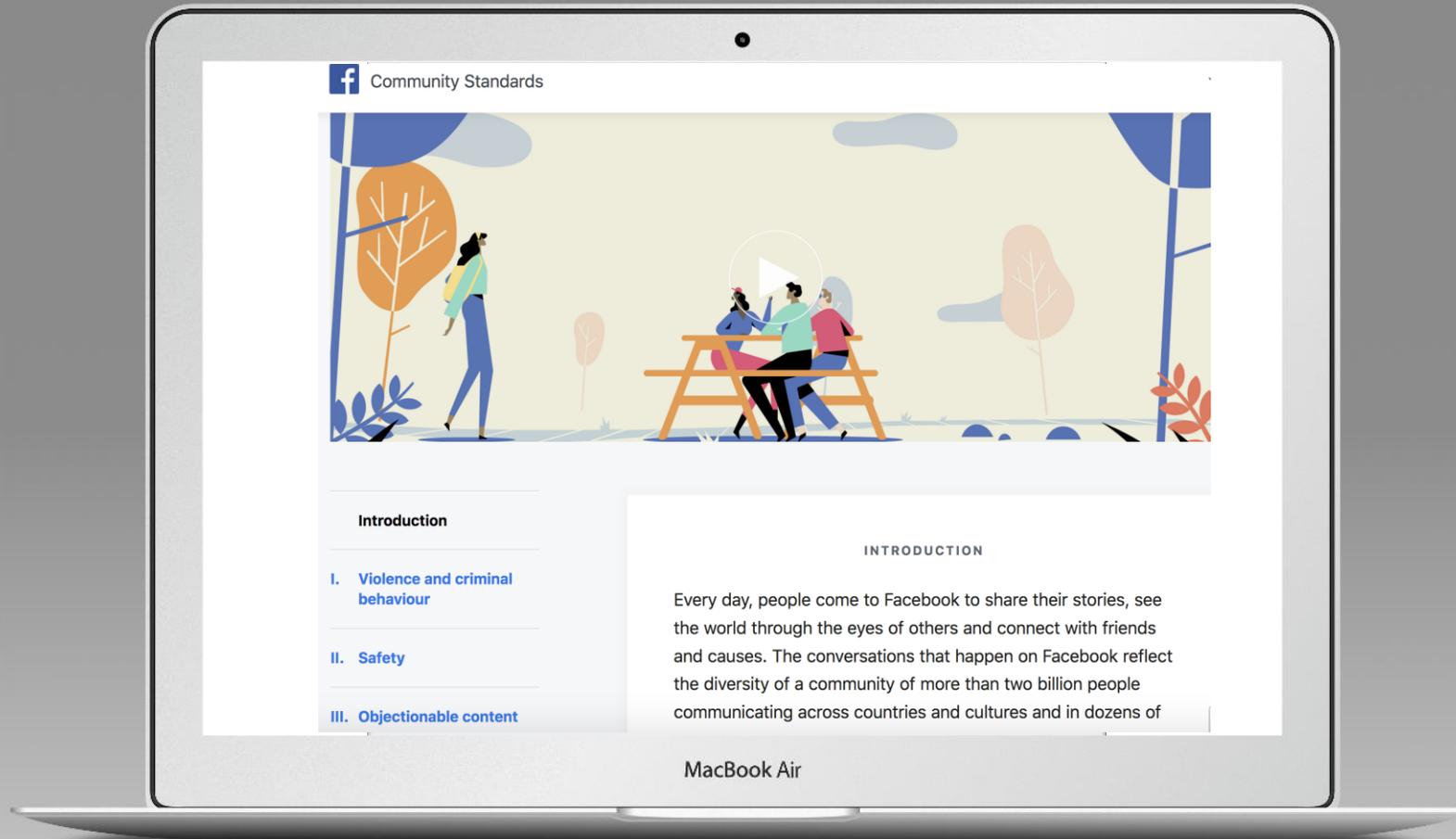
Our Mission

Founded in 2004, Facebook's mission is to give people the power to build community and bring the world closer together.

People use Facebook to stay connected with friends and family, to discover what's going on in the world, and to share and express what matters to them.



Community Standards



MacBook Air

Who we work with



At Facebook

Legal experts
Security engineers
Product and Design
Safety Operations
User Experience team
Data and Analytics
Communications and
Public Policy

Outside of Facebook

Academic Researchers
Local Non-profits
Safety Orgs
Law Enforcement
Lawmakers and Legal
Counsel
Human Rights Activists

Our Policies Govern:



Self Harm

Exploitation

Violent Groups

Hate Speech

Privacy

Harmful Activity

Bullying / Harassment

Sexual violence

Financial Fraud

Commercial spam

Identity Theft

Hacked accounts

Deceased members

Human Trafficking

Sexual content

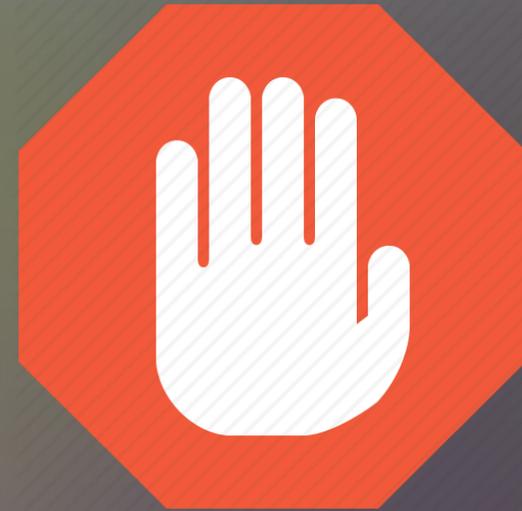
Graphic violence

Cruel Humour



Abusive Content

**Some content has no place on our site
as it is inherently harmful**





Hate Speech

Facebook removes hate speech, which includes content that directly attacks people based on their:

Race

Ethnicity

National origin

Religious affiliation

Serious disabilities
or diseases

Sexual Orientation

Sex

Gender or
Gender Identity



Hate Speech

What do you mean by an attack?

- Any violent speech or support for death, disease or harm
- Dehumanizing speech in the form of comparison to animals, filth or criminals
- Suggestion to use a weapon against a Protected Characteristic
- Calls for exclusion

Hate speech - attack

- We separate attacks into three tiers of severity, as described below.
- **Tier 1 attack:** violent speech, dehumanizing speech, mocking hate crimes
- **Tier 2 attack:** statements of inferiority, physical, mental, or moral deficiency, expressing contempt or disgust
- **Tier 3 attack:** calls to exclude or segregate, use of slur

Open Discussion



Challenge ideas, institutions and practices.

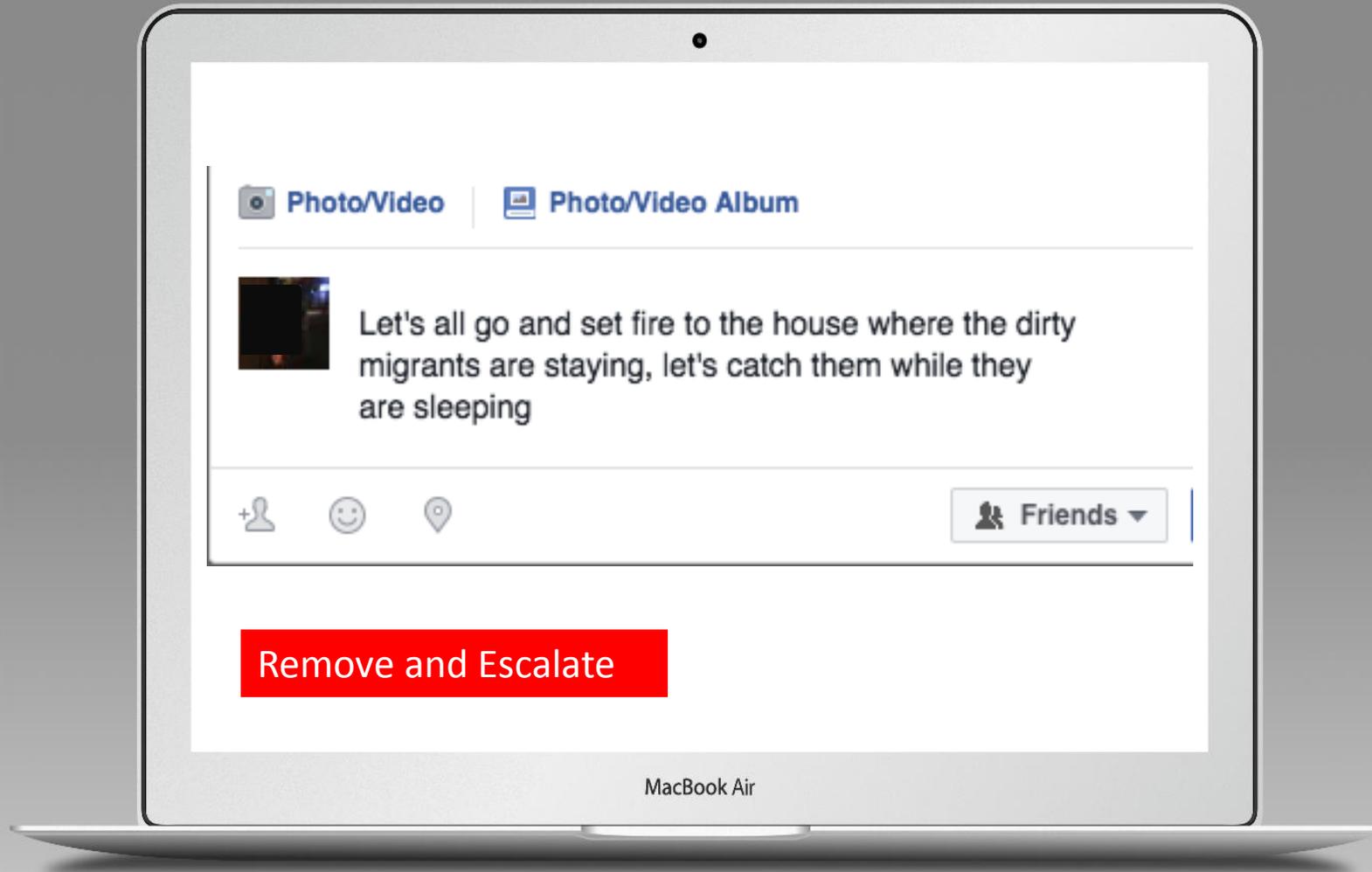
Share content containing for the purpose of raising awareness or educating others

Humour, satire or social commentary

Authentic identity = more responsible when they share this kind of commentary

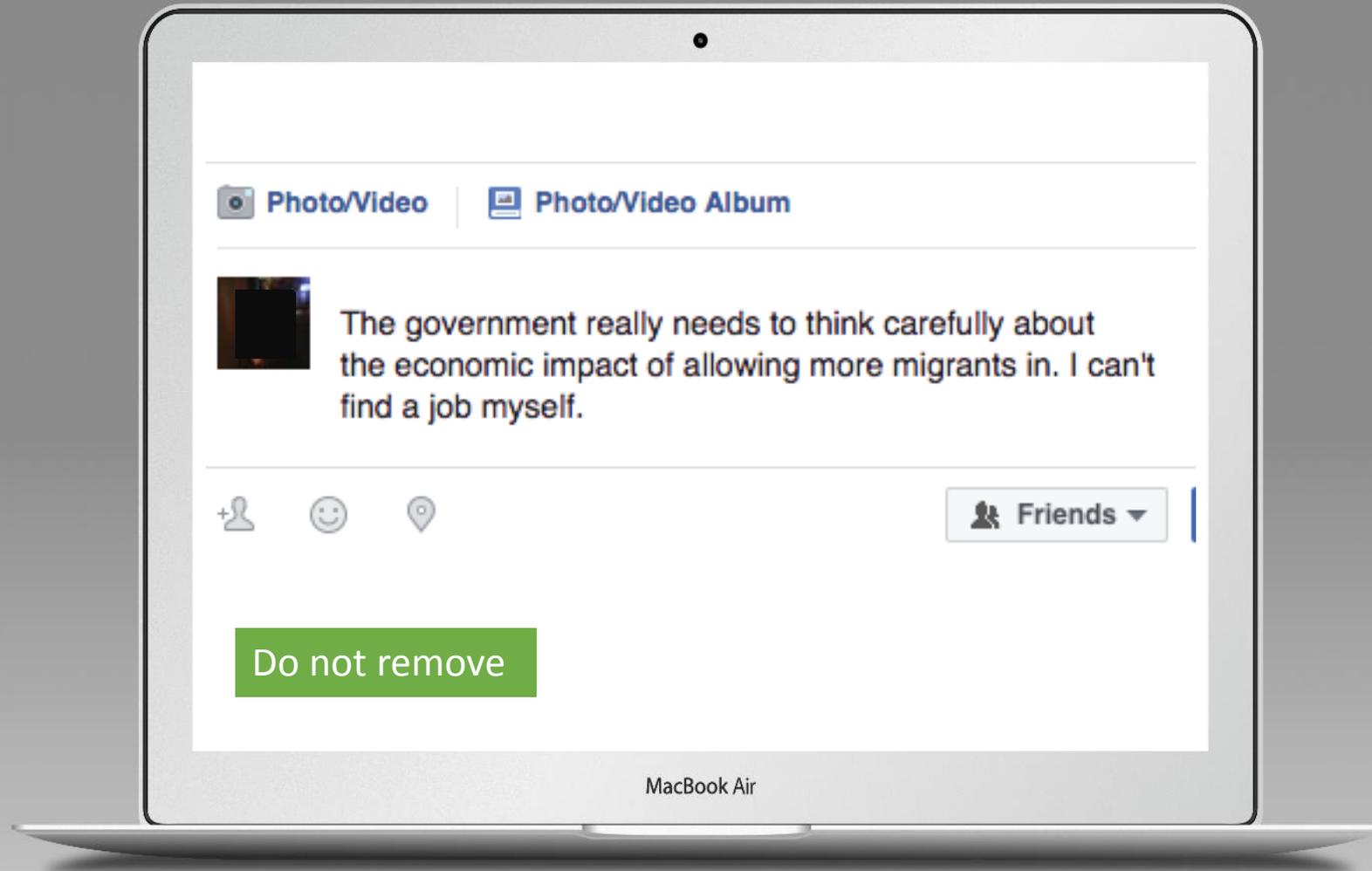
Urge people to be conscious of their audience when sharing this type of content

Practical Application



Remove and Escalate

Practical Application



Practical Application



Remove

MacBook Air



Do not remove

Caption: This is terrible. I can't believe people are saying this. I welcome Syrians here.

MacBook Air



Remove

Caption: Romanian
Gypsies are all
criminals.

MacBook Air



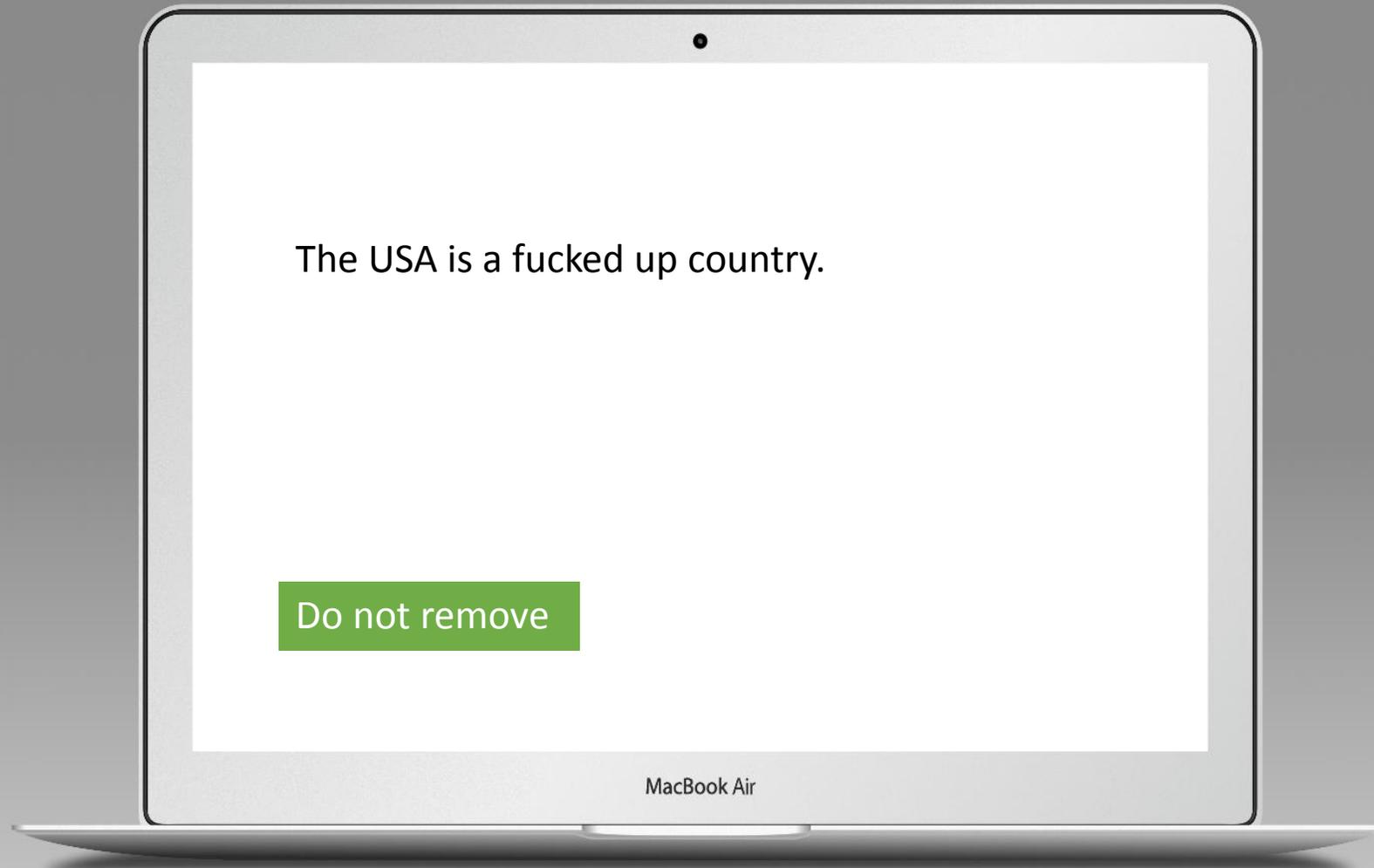
Original Caption: Romanian Gypsies are all criminals.

Do not remove

Re-shared with Caption: I'm originally from Romania. How can someone say this?

MacBook Air

Practical Application



The USA is a fucked up country.

Do not remove

MacBook Air



Americans are lazy shitty people. They should
all be shot.

Remove

MacBook Air



Tools, Reporting & Enforcement

Reporting a post

Reporting is 100% confidential

Every piece of content on Facebook can be reported

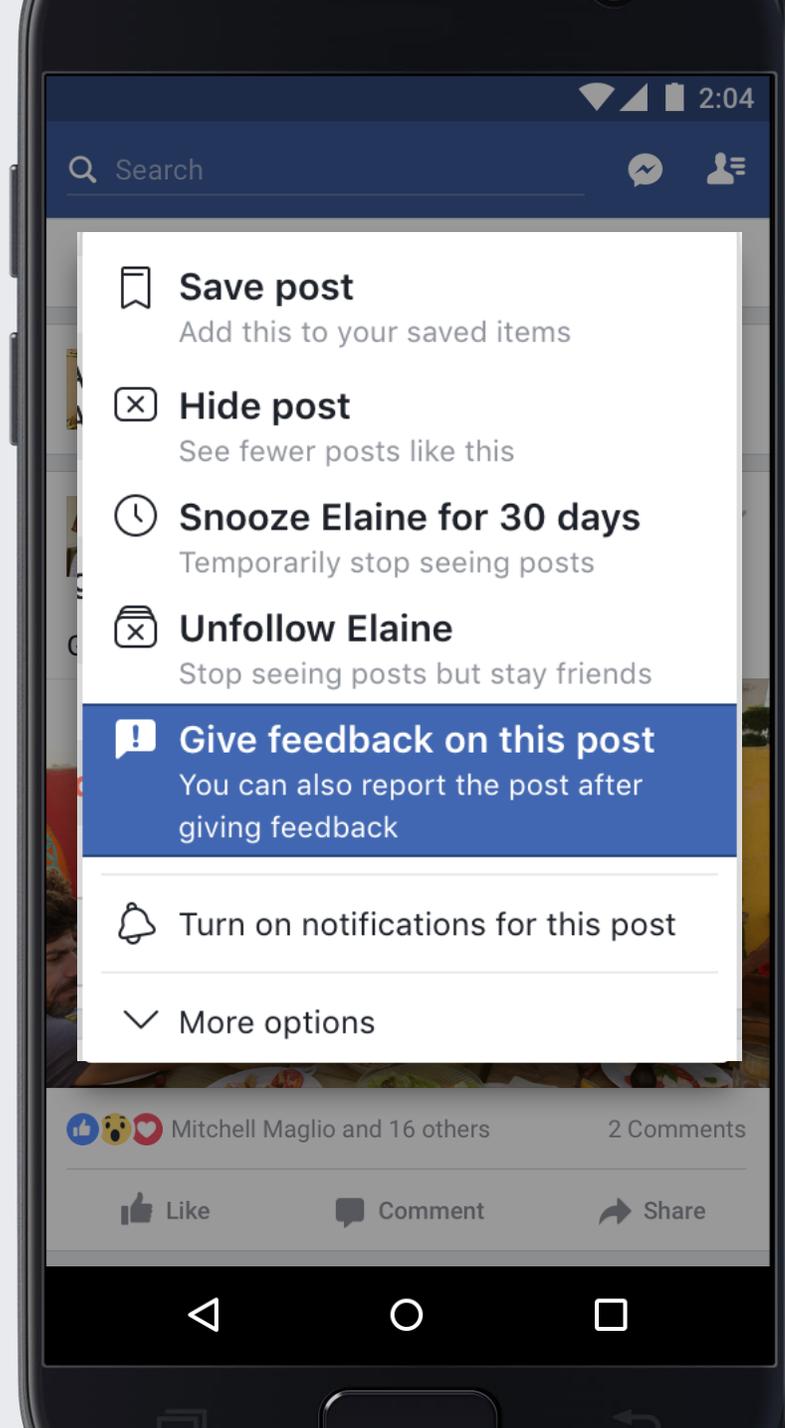
Always report the potentially abusive piece of content itself

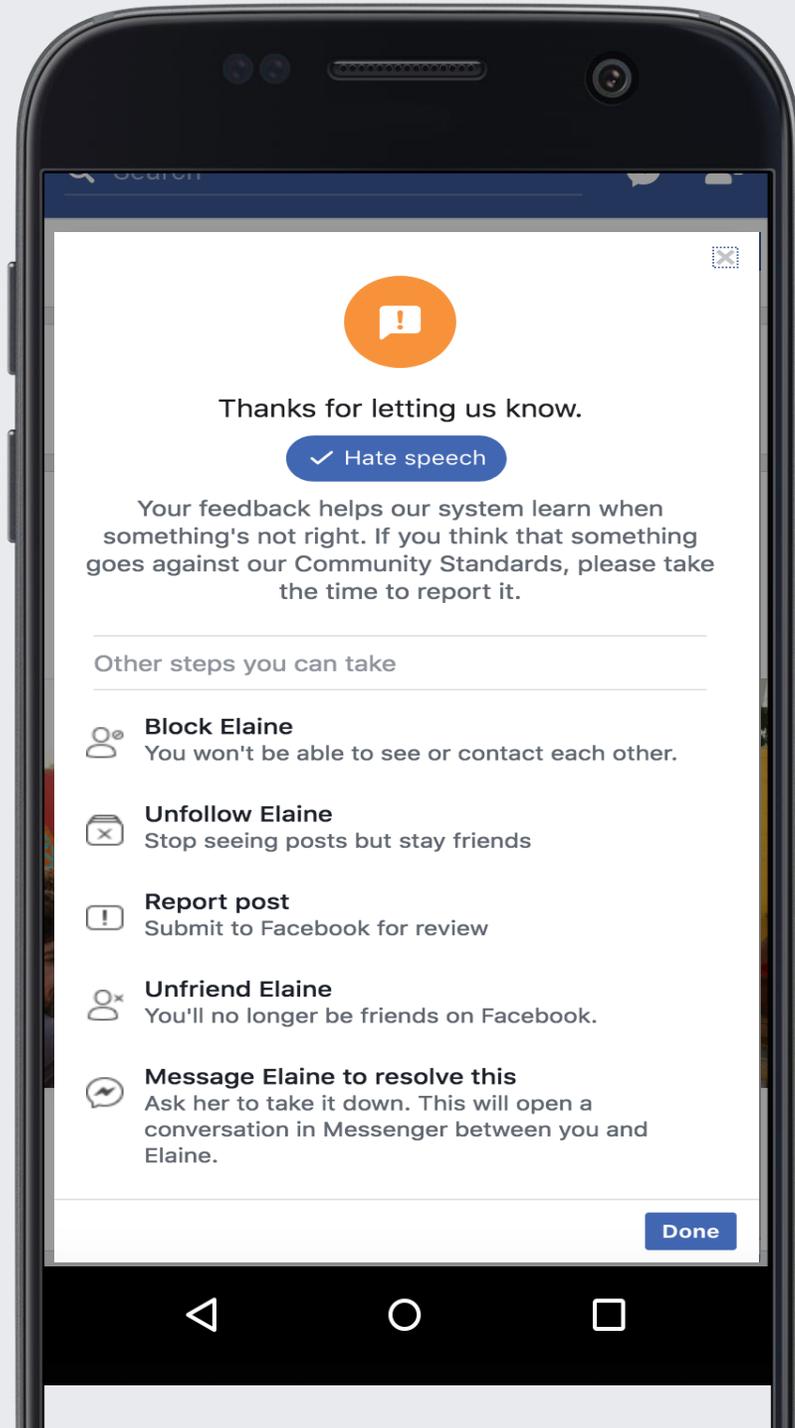
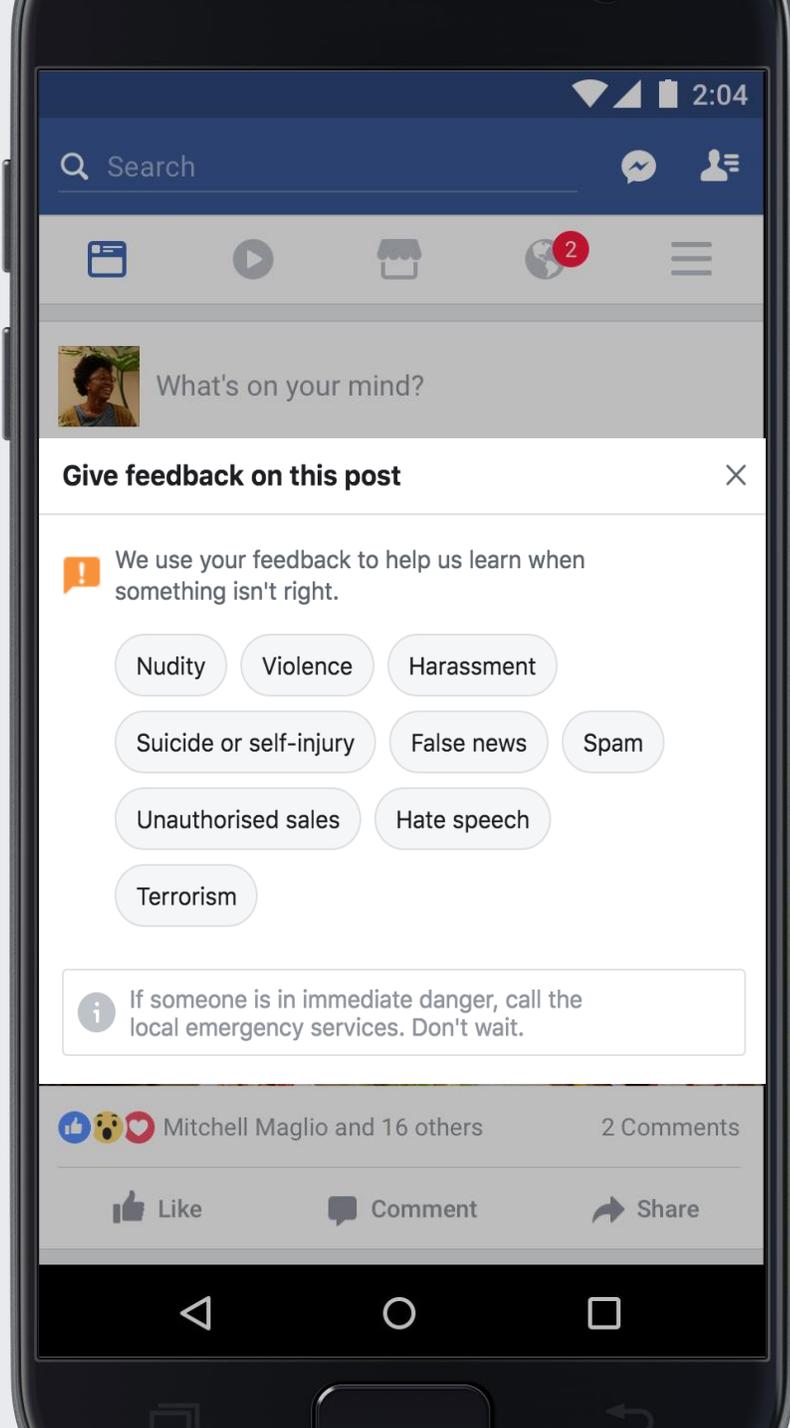
Encourage people to report themselves

Reporting once is enough

Always working to make our reporting mechanisms better







Enforcement of the Policies

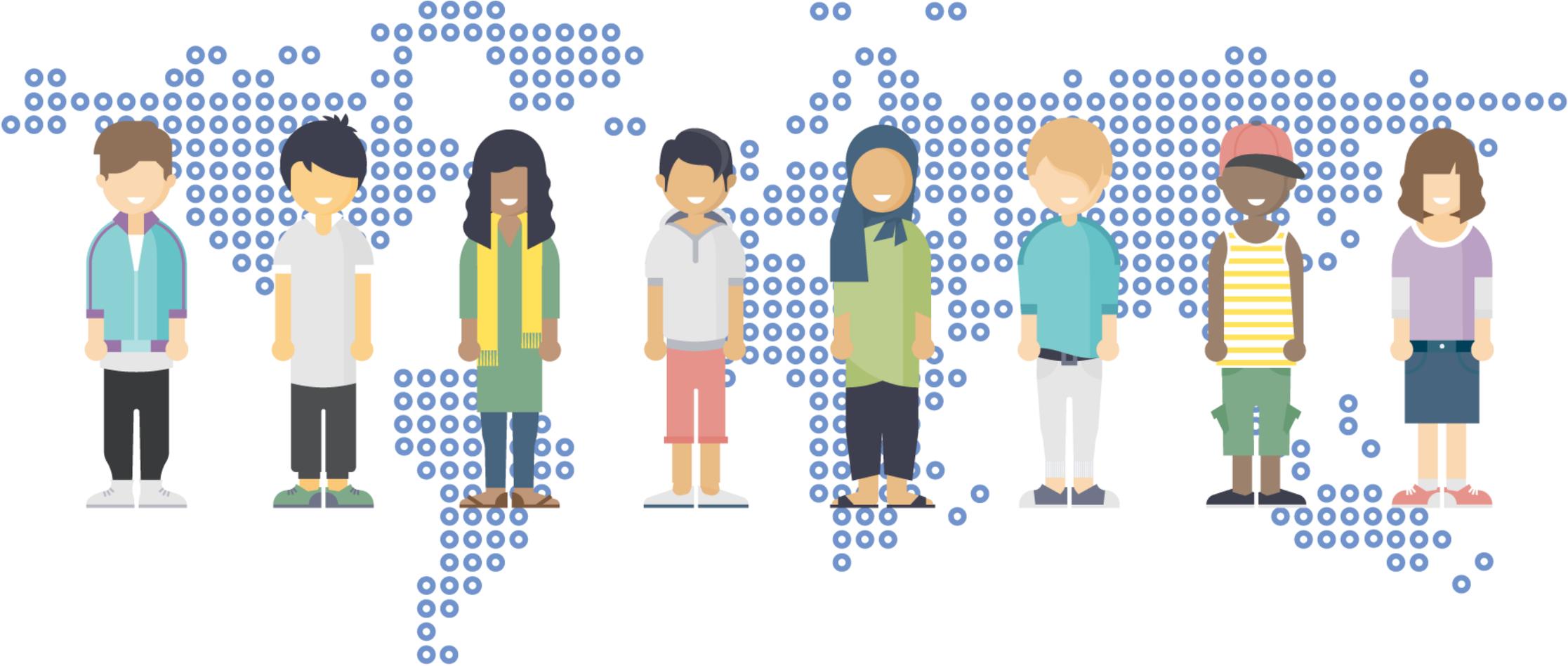


Around the
clock
review

Teams based
in **5** Global
Offices

Covering
50+
languages

We are here to help 24/7



Transparency Report

- Removed or restricted content appears in our Transparency Report

