Tech in politics: Combating viral misinformation

Max Chen
Public Policy Manager, Taiwan
Facebook
Our Mission

Give people the power to build community and bring the world closer together.
Remove bad content and bad actors that violate our policies, including hate speech and fake accounts.
Community Standards

One global set of policies
Cracking Down on Fake Accounts

We use artificial intelligence to identify over 99.6% of the fake accounts we remove before they’re ever reported. 1.5 billion fake accounts in 2nd & 3rd Quarter 2018.

AUTOMATIC
Our security systems run in the background millions of times per second

IDENTIFY FAKE ACCOUNTS
We use artificial intelligence to identify over 99.6% of the fake accounts we remove before they’re ever reported.

HUMAN
More than 30,000 people work on safety and security
Disrupting Bad Actors

**DETECT**
Harmful types of election-related activity are flagged for manual review

**TAKE ACTION**
Security teams investigate suspicious activity and take down violating accounts

**ADAPT**
Continuous improvement with each new lesson learned globally
Disrupting Bad Actors

COORDINATED INAUTHENTIC BEHAVIOR TAKEDOWNS ACROSS FACEBOOK & INSTAGRAM

Jan 2019

Removed 220 Pages, 73 FB Accounts, 29 IG accounts and Banned Twinmark Media for the use of fake accounts, ad farms, and selling access to Facebook pages in the Philippines.

We removed 783 Pages, Groups and accounts from Iran targeting several countries with upcoming elections including Israel, South Africa, Spain and the US.

Feb 2019

Removed 168 Accounts, 29 Pages, and 8 IG accounts targeting people in Moldova ahead of their elections.

March 2019

Removed FB and IG pages as part of networks in the UK, Romania and the Philippines, several networks tied to Iran, and behavior originating in Macedonia and Kosovo.

April 2019

Removed FB and IG pages as part of networks that originated in Pakistan and removed FB Pages for engaging in coordinated inauthentic behavior in India.
Reduce the spread of viral misinformation
Misinformation-Adjacent Issues

CLICK-GAP

UNORIGINAL CONTENT

CLICKBAIT & ENGAGEMENT BAIT
Fact-Checking + Facebook

**IDENTIFY**
Facebook identifies potential false news using various signals and Fact-checkers proactively identify stories on their own.

**REVIEW**
Fact-checkers **review and rate** the accuracy of stories.

**ACT**
Facebook **takes action** by showing false news lower in News Feed and providing context.
Inform our community with additional context, like Related Articles and Context Button
NEWS LITERACY CAMPAIGNS
TRUST INDICATORS
Jasper's Market posted a photo
Thursday at 2:04 PM

Suspicious.

About This Image

Related Articles

FACT-CHECKER A
Fact-Checker
Headline disputing false photo about aliens invading
Impact on Story

**STORY DEMOTED**

**SHARING INTERSTITIAL**

**USER NOTIFICATIONS**

**RELATED ARTICLES**
“Information Operations are coordinated efforts to manipulate or corrupt public debate for a strategic goal.”

— Facebook. (2019)
Coordinated Inauthentic Behavior

Why inauthenticity and coordination matter

A network of people or Pages on Facebook work together to mislead others about who they are and what they're doing

**Coordinated:** multiple accounts working together
**Inauthentic:** intentionally misleading people
**Behavior:** focus on misleading activity rather than content

*Note: CIB policy does not always apply to all suspected IO cases*

*Cybersecurity Policy continuously evaluates policy gaps and adversarial developments in different markets to update and evolve product policies accordingly*