Protecting Election Integrity on Facebook

Facebook Elections

Sean Evins
Politics & Government Outreach
Europe, Middle East, Africa
Protecting Election Integrity on Facebook

**Combating Foreign Interference**
- Monitoring proactively across elections for foreign interference from any country. We work with governments to better understand the threats.

**Cracking Down on Fake Accounts**
- Using machine learning and artificial intelligence to detect abuse, to protect against the creation of fake accounts, and to uncover coordinated behavior that is abusive and counter to our Community Standards.

**Making Advertising More Transparent**
- Working to increase transparency around ads so that the public understands more about who is creating content on our platform.

**Reducing the Distribution of False News**
- Prioritizing informative posts and downranking hoaxes and false news flagged by 3rd parties.

**Supporting an Informed Electorate**
- Helping people connect with candidates and learn about issues important that are important to them.
How We Think About the Problem
How Election Integrity is Under Attack on FB

**KNOWLEDGE**
- **MISINFORMATION**
  Misleading people about the candidates, parties, voting process, etc.
- **FALSE AMPLIFICATION**
  Distorting the public debate by over or understating support for certain POVs
- **MISREPRESENTATION**
  Undermining credibility via account/URL impersonation

**SAFETY**
- **HARASSMENT**
  Bullying people over their political views to silence or intimidate them
- **VIOLENCE**
  Use of FB to instigate, organize, or celebrate offline electoral violence
- **ACCOUNT SECURITY**
  Hacking an account to steal info or shut it down

**TURNOUT**
- **DISENFRANCHISEMENT**
  Discouraging people from voting by demotivating or misleading them

**FOREIGN INTERFERENCE**
- Is a “meta problem” where bad actors exploit many of these attack vectors
Our vision for election integrity is to ensure...

**KNOWLEDGE**
People have accurate information about electoral politics

**SAFETY**
People feel comfortable expressing themselves politically

**TURNOUT**
People are motivated to participate and vote in an election
Confronting the Problem
Proactively Monitoring for Foreign Interference

**MONITOR**

We proactively monitor for harmful types of election-related activity, such as Pages of foreign origin that are inauthentically distributing civic content. We then send these suspicious accounts to be manually reviewed by our security team to see if they violate our Community Standards or Terms of Service.

**DETECT & ACT**

For example, in the Alabama special election last December, our system detected accounts suspected for foreign coordinated activity from Macedonia and sent them to our security team which uncovered they were linked and coordinating with one another. They were quickly deleted.

**LEARN FAST**

This is just the beginning of our work in this space. Our systems are getting smarter and our reviewers are being trained to quickly review this content and only remove it if it violates our standards and policies.
How Facebook Stops Bad Actors

**SET THE STAGE**
- Create multiple Facebook Pages from the same fake accounts

**BUILD AN AUDIENCE**
- Try to generate an online following that they can eventually redirect to false news sites

**DISTRIBUTE CONTENT**
- Publish sensationalized posts and ads or flood the platform with links

**DRIVE TRAFFIC TO THEIR SITES**
- Try to make money off the traffic that goes to their own fake sites from Facebook, where they publish ads or introduce scams or malware

---

**WHAT BAD ACTORS DO**
- Facebook community and machine learning find and report fake accounts
- Systems detect malicious behavior such as posts that redirect users to spammy content
- Before ads appear on Facebook, they are subject to review
- Links to external websites that appear on Facebook are systematically reviewed
- Review and disable fake accounts and associated Pages
- As of May 2017, Facebook demotes sensationalism such as “clickbait” (deceptively titled stories) and links to “ad farms” (pages that contain mostly ads)
- Prohibits these ads from appearing on Facebook; investigates and potentially disables accounts and pages associated with illicit ad activity
- Facebook bans domains that contain malware or other scams
Cracking Down on Fake Accounts

**HUMAN**

We already have about 10,000 people working on safety and security, and we're planning to double that to 20,000 in the next year.

**AUTOMATED**

Each day we block MILLIONS OF FAKE ACCOUNTS at the point of creation.

Our security systems run in the background millions of times per second to help catch threats and remove them before they ever reach users.
View Ads Feature

Making Advertising More Transparent
Reducing the Distribution of False News

**REMOVE**
Remove content and actors that violate our policies

**REDUCE**
Reduce the spread of problematic content

**INFORM**
Inform our community with additional context
False News

FACEBOOK IS WORKING TO REDUCE THE SPREAD OF PROBLEMATIC CONTENT

REDCING THE DISTRIBUTION OF FALSE NEWS
False News PSA

False News Quick Promotion at the top of Facebook News Feed which links to Help Center

Tips for spotting false news.

Print Ads in local paper
Case Study: Clickbait

**INVENTORY**
Clickbait story from a publisher

** SIGNALS **
Headline includes “you won’t believe” or other withholding or exaggerating phrase

** PREDICTIONS **
Likelihood this story is a clickbait story

** DECISION **
Relevance score declines and story gets less distribution

REDUCING THE DISTRIBUTION OF FALSE NEWS
Where do third-party fact-checkers fit in?

**IDENTIFY**
Facebook identifies potential hoaxes using various signals

**REVIEW**
Fact-checkers review & rate their accuracy

**ACT**
Facebook takes action

---

REDUCING THE DISTRIBUTION OF FALSE NEWS
Related Articles

When there is additional reporting from third-party fact-checkers

REDUCING THE DISTRIBUTION OF FALSE NEWS
Notifications

When people are about to share something that has been rated

When something people have previously shared has since been rated

REDUCING THE DISTRIBUTION OF FALSE NEWS
Civic Engagement Product Team

GOVERNANCE TOOLS  VOTER TOOLS  CIVIC INTEGRITY

SUPPORTING AN INFORMED ELECTORATE
Protecting Election Integrity on Facebook

**Combating Foreign Interference**
Monitoring proactively across elections for foreign interference from any country. We work with governments to better understand the threats.

**Cracking Down on Fake Accounts**
Using machine learning and artificial intelligence to detect abuse, to protect against the creation of fake accounts, and to uncover coordinated behavior that is abusive and counter to our Community Standards.

**Making Advertising More Transparent**
Working to increase transparency around ads so that the public understands more about who is creating content on our platform.

**Reducing the Distribution of False News**
Prioritizing informative posts and downranking hoaxes and false news flagged by 3rd parties.

**Supporting an Informed Electorate**
Helping people connect with candidates and learn about issues important that are important to them.
Questions?