Facebook reports on implementation of the Code of Practice on Disinformation

May report
Introduction

This report outlines actions taken during May by Facebook to take forward the implementation of the EU Code of Practice on Disinformation. Facebook's priority during the last month was rolling out our Ad Library Report; publishing our Community Standards Enforcement and Transparency Report; launching our civic engagement products; and ramping up the Elections Operations Center in preparation for the critical period of the elections.

I.   Scrutiny of ad placements

As we noted in earlier updates, our ads policies do not allow the provision of ads which contain 'Low Quality or Disruptive Content', 'Misleading or False Content' or are Circumventing our Systems.

II.   Political and issue based advertising

EU Ad Library Report Launched
In May, we launched our Ad Library Report for political ads within the EU. This report lets people see how many political and issues ads were run in a given country – as well as aggregated advertiser spend and top searched keywords in the Ad Library. It also includes basic information like the name of the Page running the ads, the "Paid for by" disclaimer and a link to the ads in the Ad Library. The report will be refreshed daily, and people can choose to see spend data for the last day, week, month and/or 90-day period. The report will also include aggregate spend by region for each country. Anyone can access and download the report regardless of whether or not you have a Facebook account. The report can be viewed here: https://www.facebook.com/ads/library/report.

We believe the report will be a useful tool to people (especially researchers, journalists and watchdogs) who are interested in understanding high-level activity in the Ad Library since it was launched in late March. Users can then use the Ad Library to deep-dive into specific ads.

Since late March 2019 (when we launched the ads authorisation and Ad Library tools) to May 29, there are 343,736 political ads across the EU, totaling about 19.8 million Euros of political ads spend. The following table provides a country-by-country breakdown taken from the Ad Library Report:

Includes ads related to politics or issues of importance from March 2019 to May 29		
Ad Library Totals	**Total Number of Political Ads**	**Total Amount Spent on Political Ads (EUR)**
Austria	15,414	960,661
Belgium	16,164	1,527,992
Bulgaria	3,682	137,239
Croatia	4,174	109,151
Cyprus	2,831	61,258
Czech Republic	9,117	272,848
Denmark	16,921	1,009,596
Estonia	3,039	95,431
Finland	8,086	401,416
France	16,164	1,527,992
Germany	76,777	3,530,336
Greece	12,907	691,155
Hungary	9,475	591,463
Ireland	8,283	366,987
Italy	22,785	1,825,069
Latvia	3,149	53,205

Lithuania	3,824	105,045
Luxembourg	4,405	69,811
Malta	4,552	116,959
Netherlands	9,296	597,340
Poland	12,220	446,734
Portugal	6,189	102,619
Romania	8,342	602,898
Slovakia	12,515	406,362
Slovenia	3,479	59,017
Spain	35,318	2,756,755
Sweden	14,628	1,368,116
Source: Ad Library Report		

III.   Integrity of services

Transparency & Community Standards Enforcement Report

We take our commitment to transparency seriously and have published our latest Transparency Report for the second half of 2018. Since 2013, we have released information on the nature and extent of the government requests we receive for user data. Each half, we look to improve and expand our scope, which now includes information about the number of content restrictions based on local law, reports on locations where access to Facebook products and services were disrupted, and reports of counterfeit, copyright and trademark infringement.

Also included in this release is our third Community Standards Enforcement Report, covering Q4 2018 and Q1 2019, which shows how much violating content we have detected on our service, so people can judge for themselves how well we're doing at

enforcing our Community Standards. We have updated this report to now also include:

- Data on how much content people appealed and how much content was restored after we initially took action.
- Data on how we're doing at removing attempts at illicit sales of regulated goods — specifically, firearm and drug sales.

In total, we are now including metrics across nine policies within our Community Standards: adult nudity and sexual activity, bullying and harassment, child nudity and sexual exploitation of children, fake accounts, hate speech, regulated goods, spam, global terrorist propaganda and violence and graphic content.

Regarding fake accounts, we aim to remove as many fake accounts on Facebook as we can. We prioritize enforcement against users and accounts that seek to cause harm and find many of these fake accounts are used in spam campaigns and are financially motivated. The amount of accounts we took action on increased in the last six months due to automated attacks by bad actors who attempt to create large volumes of accounts at one time. However, our detection technology has also improved, blocking millions of attempts to create fake accounts every day and detecting millions more often within minutes after creation. (NB: We don't include blocked attempts in the fake account data.)

- We disabled 1.2 billion accounts in Q4 2018 and 2.19 billion in Q1 2019.
- The majority of these accounts were caught within minutes of registration, before they became a part of our monthly active user population.
- 99.8% of fake accounts actioned were found and flagged by Facebook before users reported them.

To help answer a lot of questions about the fake account numbers, we published a blog about how we measure fake accounts.

Community Standards Updates & Independent Oversight

Over the last year, we've taken a number of steps to be more transparent in how we develop our policies and how we measure our efforts to enforce them. When it comes to our policies, we began sharing the meeting minutes from our meetings where we determine updates to our policies, and now provide a change log on the Community Standards website so that each month everyone can see where exactly we've made updates to our policies. Additionally, as part of our efforts to enable academic research, we are awarding grants for 19 research proposals across the world to study our content policies and how online content influences offline events.

Independent external review and input is an integral component to how we improve. In that spirit, we established the Data Transparency Advisory Group (DTAG) — comprised of international experts in measurement, statistics, criminology and governance — to provide an independent, public assessment of whether the metrics we share in the Community Standards Enforcement Report are meaningful and accurate. We provided the advisory group detailed and confidential information about

our enforcement processes and measurement methodologies and this week they published their independent assessment. In its assessment, the advisory group noted that the Community Standards Enforcement Report is an important exercise in transparency and that they found our approach and methodology sound and reasonable. They also highlighted other areas where we could be more open in order to build more accountability and responsiveness to the people who use our platform. These important insights will help inform our future work.

Protecting Facebook Live from Abuse

We've been reviewing what more we can do to limit Facebook Live from being used to cause harm or spread hate. As a result, people who have broken certain rules on Facebook — including our Dangerous Organizations and Individuals policy — will be restricted from using Facebook Live.

We will now apply a 'one strike' policy to Live in connection with a broader range of offenses. From now on, anyone who violates our most serious policies will be restricted from using Live for set periods of time – for example 30 days – starting on their first offense. For instance, someone who shares a link to a statement from a terrorist group with no context will now be immediately blocked from using Live for a set period of time.

We plan on extending these restrictions to other areas over the coming weeks, beginning with preventing those same people from creating ads on Facebook.

We recognize the tension between people who would prefer unfettered access to our services and the restrictions needed to keep people safe on Facebook. Our goal is to minimize risk of abuse on Live while enabling people to use Live in a positive way every day.

Coordinated Inauthentic Behaviour

Coordinated inauthentic behavior (CIB), which is when people or organizations create networks of fake accounts to mislead others about who they are, or what they're doing, to manipulate public debate for a strategic goal. We're constantly working to detect and stop CIB activity because we don't want our services to be used to manipulate people. We're taking down Pages, Groups and accounts based on their behavior, not the content they posted. The bad actors behind CIB activity coordinate with one another and use fake accounts to misrepresent themselves, and that is the basis for our action. We are making progress rooting out this abuse, but as we have said before, it's an ongoing challenge as bad actors become more sophisticated. We are committed to continually improving to stay ahead, which means building better technology, hiring more people and working more closely with law enforcement, security experts and other companies.

In May, we have reported the following CIB cases within Europe and its surrounding neighbourhood:

- **[Removing More Coordinated Inauthentic Behavior From Russia](#)**: On May 6, we removed multiple Pages, Groups and accounts that were involved in coordinated inauthentic behavior on Facebook and Instagram. We found two separate, unconnected operations that originated in Russia and used similar tactics, creating networks of accounts to mislead others about who they were and what they were doing.

    - We removed a total of 118 Pages, Groups, and accounts from the two networks in Russia.

    - The first network was a group of individuals operating fake accounts to run Pages and Groups, disseminate their content, and increase engagement, and also to drive people to an off-platform domain that aggregated various web content. They frequently posted about local and political news including topics like the military conflict in Eastern Ukraine, Russian politics, political news in Europe, politics in Ukraine and the Syrian civil war. We had already disabled many accounts run by the people behind this operation for various violations, including impersonation, and some of the activity we found was linked to accounts we removed in prior enforcement actions. We identified some of these accounts and Pages through follow-on investigations of the accounts and Pages we removed earlier this year ahead of the elections in Ukraine. We have shared information about our analysis with law enforcement, policymakers and industry partners.

    - The second network emanated from Russia and focused on Austria, the Baltics, Germany, Spain, Ukraine and the United Kingdom. The individuals behind this campaign - which were also active on other internet platforms - engaged in a number of deceptive tactics, including the use of fake accounts to join Groups, impersonate other users and to amplify allegations about a public figure working on behalf of intelligence services. They also posted content about local politics including topics like immigration, religious issues and NATO. We took additional steps of pre-briefing governments and law enforcement in affected countries for this second group. We identified these accounts and Pages through our ongoing internal investigations into suspected Russia-linked coordinated inauthentic behavior. Like the other operation, we've shared our analysis with law enforcement, policymakers and industry partners.

- **[Removing Coordinated Inauthentic Behavior From Israel](#)**: On May 16, we removed 265 Facebook and Instagram accounts, Facebook Pages, Groups and events involved in coordinated inauthentic behavior.
    - This activity originated in Israel and focused on Nigeria, Senegal, Togo, Angola, Niger and Tunisia along with some activity in Latin America and Southeast Asia. The people behind this network used fake accounts to run Pages, disseminate their content and artificially increase engagement. They also represented themselves as locals, including local

news organizations, and published allegedly leaked information about politicians.

- o The Page administrators and account owners frequently posted about political news, including topics like elections in various countries, candidate views and criticism of political opponents.

- o Although the individuals behind this network attempted to conceal their identities, our investigation found that some of this activity was linked to an Israeli commercial entity, Archimedes Group. It has repeatedly violated our misrepresentation and other policies, including by engaging in coordinated inauthentic behavior. This organization and all its subsidiaries are now banned from Facebook, and it has been issued a cease and desist letter.

- o We identified these accounts and Pages through our internal investigations into suspected coordinated inauthentic behavior. We have shared information about our analysis with industry partners and policymakers.

- **Removing More Coordinated Inauthentic Behavior From Iran:** On May 28, we removed 51 Facebook accounts, 36 Pages, seven Groups and three Instagram accounts involved in CIB that originated in Iran. The individuals behind this activity — which also took place on other internet platforms and websites — misled people about who they were and what they were doing.

  - o They purported to be located in the US and Europe, used fake accounts to run Pages and Groups, and impersonated legitimate news organizations in the Middle East. The individuals behind this activity also represented themselves as journalists or other personas and tried to contact policymakers, reporters, academics, Iranian dissidents and other public figures.

  - o A number of these account owners also attempted to contact authentic Instagram accounts, some of which later posted content associated with this activity.

  - o The Page admins and account owners typically posted content in English or Arabic without a focus on a particular country, although some Pages focused on the US or UK. The posts from these Pages and accounts discussed topics like public figures and politics in the US and UK, US secessionist movements, Islam, Arab minorities in Iran and the influence of Saudi Arabia in the Middle East.

  - o Based on a tip shared by FireEye, a US cybersecurity firm, we conducted an internal investigation into suspected Iran-linked coordinated inauthentic behavior and identified this activity. We have shared our analysis with law enforcement, policymakers and industry partners.

Dublin Elections Operations Center: In the final and critical weeks leading up to the election, our Election Operations Center in Singapore, Dublin and Menlo Park allowed our global teams to better work across regions and rapidly respond to issues and escalations. It provided 24/7 support in all EU official languages, and facilitated an additional layer of defense against fake news, hate speech, voter suppression and other Community Standards violations.

- Over 500 full-time employees working on elections and over 40 regional and functional teams across the company were plugged into the Elections Operations Center, monitoring and/or reporting issues, potential threats, and suspicious activities.
- The Election Operations Center, for example, was able to quickly investigate information received from the campaign group, Avaaz. They had shared several reports of what they perceived to be Facebook Community Standards violations and [coordinated inauthentic behavior](#) (CIB) across Europe. Where violations were found based on our own internal investigation, we were able to quickly take action. As a result, we removed a number of fake and duplicate accounts that were violating our authenticity policies, as well as multiple Pages for name change and other violations. We also took action against some additional Pages that repeatedly posted misinformation and continued to monitor for further violations. But we did not find evidence of CIB from Avaaz's reports.
- To demonstrate how the Elections Operations Center works, we also opened the doors of our Dublin Elections Operations Center on May 2 to TV crews and print reporters from across the EU for briefings, broadcast interviews and tours. On May 14, we provided a virtual tour and briefing of the center to policymakers, regulators, NGOs and think tanks in Brussels and across Europe.

Election Security Escalation Channel: It is important for us to collaborate with external partners, as we recognise that no single organization can do it alone. To that end, we have launched an Election Security Escalation Channel for trusted security and intelligence partners to directly report potential cybersecurity and information operations threats. This channel now includes partners from ten EU countries and will continue operating even after the election. We will continue engaging and adding new partners to this channel.

Escalation Channel for Political Pages: We launched a dedicated escalation channel for admins of government and political pages. Called the Government, Political and Advocacy Concierge (GPAC), the escalation channel allows political pages admins to report tech issues, bugs, imposter accounts, violating and suspicious content, and other issues.

Safety & Security Information Units: We also ran information units on people's News Feeds about tools and resources to help political parties, politicians and candidates stay secure on our platform.

- **Two-Factor Authentication:** In late April and early May, we ran a security information unit that was seen at the top of people's News Feed across the EU 28 countries reminder users to turn on two-factor authentication. Two-factor authentication is an extra layer of security designed to ensure that no one else can access an individual's account, even if someone knows the password.

- **Government, Political and Advocacy Concierge (GPAC)**: In May, we ran an information unit that was seen at the top of people's News Feed which directed people to our dedicated escalation channel for admins of government and political pages. This escalation channel allows political pages admins to report tech issues, bugs, imposter accounts, violating and suspicious content.

Health Integrity effort

The mission of our Health Integrity team is to minimize bad health related experiences for people and harm to society.

Highlights for this month include:

- Demoting Pages which spread vaccine misinformation - now, when a Page violates our vaccine misinformation policy, they will see their distribution reduced in News Feed. These Pages are already removed from Pages You May Like Search visibility reduced, and are subject to Page Like Friction - a warning notice displayed before people connect with the Page.
- Perceived health misinformation survey results - these research findings support our prioritization of miracle cures and treatments, and have identified new health issues. 55% of users across 10 countries surveyed reported seeing health content on FB. Over a third of those who had seen health content reported seeing health misinformation on topics including vaccines, miracle cures, weight loss, cancer and sexual health. Ads and Pages were the two most prevalent perceived content sources.

IV.   Empowering consumers

Civic Engagement: Supporting an informed electorate is an important part of Facebook's election integrity work. People are already using Facebook to talk about politics and issues that matter to them and to communicate with their elected officials. We want to support them and make it easier to connect them to reliable information about the European elections.

- **Facebook Voter Information Reminder:** In the days leading up to the election, on May 20-22, we ran a Voter Information Reminder across all 28 EU countries. The reminder was displayed at the top of people's News Feed and included a button to see more information about the upcoming election in their

country. The button would redirect people to either the [EU election website](#) or the official election website of the country.

- **Election Day Reminder:** On election day, we ran an Election Day Reminder across all 28 EU countries, reminding people that it is voting day. The reminder was displayed at the top of people's News Feed and included a button to see more information, which would redirect people to either the [EU election website](#) or the official election website of the country. The reminder also included a "Share That You've Voted" button, prompting people to engage and spread the word about voting.

- **Instagram Story EU Elections Sticker:** We created two unique Instagram Story stickers for the upcoming European parliamentary election, including one that says "I voted" in 24 different languages. We worked with the European Parliament on the design of the stickers. The stickers were made available in all 28 EU countries during the week of elections, from May 22 to 26. The sticker allowed Instagram users to share the historic moment with their friends and followers. The stickers included a reminder that people could tap and be directed to the [european-elections.eu](#) website for more information on how to vote.

Digital Literacy: Digital literacy is an important part of the work we do to help keep our community safe and.

- **Pan-European:** We launched a media literacy campaign in partnership with Full Fact (UK), and other fact-checkers - Maldita (ES), Newtral (ES), Correctiv (DE), TheJournal (IE), Pagella Politica (IT), Demagog (CZ), Nieuwscheckers (NL) and Ellinika Hoaxes (GR). The campaign ran in all 28 member states by either local fact-checking organizations or Full Fact and provided updated lessons on [how to stamp out false news](#).

Three questions to help stamp out false news

Stop, think, and check...
- ☑ Where's it from?
- ☑ What's missing?
- ☑ How do you feel?

fullfact.org/toolkit

WHERE'S IT FROM?

A trusted source is your safest option.

If you don't know the source, check out the about page or ask yourself why they're sharing the story.

①

fullfact.org/toolkit

WHAT'S MISSING?

Get the whole story not just a headline.

One click can help you spot false news. Read the whole story and watch out for images, numbers, and quotes that don't have sources or that might have been taken out of context.

(4)

fullfact.org/toolkit

- **Pan-European**: Facebook and Freeformers have partnered with over 20 in-country NGOs and training organisations to deliver the ongoing Digital Skills Training Programme directly to 75,000 citizens across seven countries in Europe: Germany, France, Italy, Spain, Poland, the United Kingdom and most recently, The Republic of Ireland. The highly engaging face-to-face events are comprised of a series of workshops that are designed to develop the confidence and the skills these citizens need in the digital economy, and to become more enthusiastic to learn more - including in the media literacy space to ensure they are safe online.

- **Poland and Digital Youth Forum**: Facebook worked with the Digital Youth Forum in Poland on a three-day educational event for youth that aims to promote safe, informed and innovative use of new technologies as an alternative to risky online behaviours. 400 teens (13-17 years old) participated in the Forum and over 10,000 students from 148 schools all over Poland also followed the event via online streaming. "Przestrzeń" from Facebook, the Warsaw community hub, also held digital skills and safety workshops for 94 teens as part of this initiative.

- **Germany**: Facebook Germany is finalizing preparations for the launch of an inaugural event on 17 June with the newspaper ZEIT, as part of their cooperation programme "ZEIT für die Schule", supporting students in

identifying false news and finding reliable information. The event will present awards to 5 winning schools participating in the programme. The whole collaboration is being held under the patronage of the German Minister of Education, Anja Karliczek.

- **Sweden**: Facebook hosted a delegation of public and private representatives from all levels of the educational sector in Sweden - including school authorities, the Teachers' Union, companies, MPs and the Research Institute for Innovation Sweden - at its headquarters in California to discuss media literacy and disinformation. The session focused on sharing best practices and discussing common efforts that could help increase people's resistance against disinformation, propaganda and online hateful content. The group discussed the need to work together on strengthening media literacy and digital skills, with schools/education sector being the key to resilience when it comes to disinformation, and Facebook providing the platform for teachers and students to come together in a way that is beneficial to their education.

Newsrooms & Journalists Training: In May, our News Partnerships team completed the last of 10 training events across the EU for newsrooms and journalists, to help them prepare for the European elections. We trained over 400 journalists on how to tell digital stories on our platforms and how to spot false news and maintain integrity in digital reporting.

EU Elections Integrity Campaigns: From early April until late May, we ran an awareness-raising campaign across the EU - including Germany, France, Spain, Italy, Poland, Belgium, the Netherlands, Sweden, Denmark, Ireland and the UK - to inform the general public about our election integrity efforts in the run up to the European Parliament elections. The information campaign focused on three areas:

- The reduction of misinformation
- Cracking down on fake accounts
- Political ads authorisation and ads transparency tools

Campaigns ran on our platforms, as well as offline in print newspapers and digital billboards.

**We're working to keep Facebook secure.
At election time. And all the time.**

We've tripled our team of content reviewers, system engineers and security experts. Using artificial intelligence, this team is able to proactively identify 99.6% of the fake accounts we remove – before they're ever reported.

We're also working with independent fact-checkers and machine learning to tackle the spread of false news. If anything on Facebook is rated as false news, we significantly reduce its distribution in News Feed.

What's more, we're also requiring publishers of political ads to go through a rigorous authorisation process to ensure the people running ads during the EU elections are who they say they are. Plus, we're adding labels to political ads so you know exactly who paid for them.

See more details about our ongoing efforts:
facebook.com/EuropeanElections

facebook

V.   Empowering the Research Community

Social Science One: Last month we announced that our partners at Social Science One and the Social Science Research Council (SSRC) announced the first researchers who will gain access to privacy-protected Facebook data. That includes more than 60 researchers from 30 academic institutions, across 11 countries including Italy, Germany, the Netherlands, Poland, Denmark and France. To ensure full independence of the research and the researchers, Facebook did not play any role in the selection of the individuals or their projects and will have no role in directing the findings or conclusions of the research.

Content Policy Research on Social Media Platforms: As part of our efforts to enable academic research, we have awarded grants for 19 research proposals around the world to study our content policies, specifically around hate speech and preventing offline harm. The purpose of this research is to help us develop better content policies, assess possible interventions, and understand the mechanisms by which online rhetoric influences offline events. We received 184 proposals from 38 countries; a selection committee comprising members of the integrity research and policy teams evaluated them. Grantees in Europe include researchers from the

University of Amsterdam, Erasmus University Rotterdam, Nova University Lisbon, and King's College London.

Data Transparency Advisory Group (DTAG): The DATG was formed last year to provide an independent, public assessment of whether the metrics we share in the Community Standards Enforcement Report (CSER) provide accurate and meaningful measures of Facebook's content moderation challenges and our work to address them. This is an independent body made up of international experts in measurement, statistics, criminology and governance.

- The DATG published its first full assessment. In the report, the group noted that the CSER is an important exercise in transparency and they found our approach and methodology sound and reasonable. They highlighted areas where we could be more open in order to build more accountability and responsiveness to the people who use our platform and provided steps which we could take in order to improve the metrics and transparency about Community Standards enforcement practices. These important insights will help inform our future work.

Manipulated Media Research: To stay ahead of the type of adversarial media manipulation we saw after Christchurch - when some people modified the video to avoid detection in order to repost it after it had been taken down - requires research that can be turned into technical innovation. To that end, we've invested $7.5 million in new research partnerships with leading academics from three universities - The University of Maryland, Cornell University and The University of California, Berkeley - to help improve image and video analysis technology. Our partners will research new techniques to 1) detect manipulated media across images, video and audio, and 2) distinguish between unwitting posters and adversaries who intentionally manipulate videos and photographs.