

U.S. 2020 Facebook and Instagram Election Study

User Guide (First Edition)

July 27, 2023

Table of contents

Part 1: Study overview	7
Introduction	7
Citation	8
Data access	8
Project roles and collaboration overview	8
Support	8
Data overview	9
Samples	9
Participants	9
US Adult active users	10
Data types	10
NORC-collected data	10
Individual-level survey data	10
Public records data	10
External passive web-tracking data	11
Meta-collected data	12
User privacy	12
Data disclosure	13
Disclaimers	14
Part 2: Datasets	15
Metadata of Domains Shared on Facebook	15
Overview	15
Facebook Feed Intervention Experiment Participants	16
Overview	16
Participants	16
Data categorization	16
Experimental design	17
Aggregation	17
Disclosures	17
Platform limitations	17
Reshares	18
Content engagement	18
Data issues	18
Filtering	18
Instagram Intervention Experiment Participants	20
Overview	20
Participants	20

Data categorization	20
Experimental design	21
Aggregation and filtering	21
Disclosures	21
Platform limitations	21
Missing values	21
Data issues	22
Filtering	22
<u>Passive-Tracking Participants' Daily Views of Facebook Posts with Civic News Domains</u>	<u>23</u>
Overview	23
Participants	24
Data categorization	24
Aggregation	24
Disclosures	25
<u>Passive-Tracking Participants' Views of Facebook Posts with Civic News Domains</u>	<u>25</u>
Overview	25
Participants	26
Data categorization	26
Aggregation	26
Disclosures	27
<u>Exposure to Facebook Posts with Civic News Domains</u>	<u>28</u>
Overview	28
Data categorization	29
Aggregation	29
Disclosures	29
Platform limitations	29
Categories and classifiers	30
<u>Exposure to Facebook Posts with Civic News URLs</u>	<u>31</u>
Overview	31
Data categorization	32
Aggregation	32
Disclosures	32
Platform limitations	32
Categories and classifiers	33
<u>Engagement with Facebook Posts with Civic News Domains</u>	<u>34</u>
Overview	34
Data categorization	35
Aggregation	35
Disclosures	35
Platform limitations	35

Categories and classifiers	36
<u>Engagement with Facebook Posts with Civic News URLs</u>	<u>37</u>
Overview	37
Data categorization	38
Aggregation	38
Disclosures	38
Platform limitations	38
Categories and classifiers	39
<u>Potential Exposure to Facebook Posts with Civic News Domains</u>	<u>40</u>
Overview	40
Data categorization	41
Aggregation	41
Disclosures	41
Platform limitations	41
<u>Potential Exposure to Facebook Posts with Civic News URLs</u>	<u>43</u>
Overview	43
Data categorization	44
Aggregation	44
Disclosures	44
Categories and classifiers	45
<u>Coexposure to Facebook Posts with Civic News Domains</u>	<u>47</u>
Overview	47
Data categorization	48
Aggregation	48
Disclosures	48
<u>Coexposure to Facebook Posts with Civic News URLs</u>	<u>50</u>
Overview	50
Aggregation	51
Disclosures	51
Platform limitations	51
Categories and classifiers	51
<u>Daily Ideological Segregation of the Audience of Facebook Posts with Civic News Domains and URLs</u>	<u>53</u>
Overview	53
Active users	54
Data categorization	54
Aggregation	54
Disclosures	55
Platform limitations	55
Categories and classifiers	56

Weekly Ideological Segregation of the Audience of Facebook Posts with Civic News Domains and URLs	57
Overview	57
Data categorization	58
Aggregation	58
Disclosures	58
Ideological Alignment of Users' Facebook Networks	59
Overview	59
Data categorization	60
Aggregation	60
Disclosures	60
Facebook User Attributes	61
Overview	61
Data categorization	61
Aggregation	62
Disclosures	62
Exposure to and Engagement with Facebook Posts	63
Overview	63
Data categorization	64
Aggregation	64
Disclosures	64
Data issues	64
Filtering	65
Part 3: Data Dictionary	66
Glossary of terms	67
Appendix: Experimental design	67
Experimental studies	67
Recruitment	67
Platform interventions	68
Control: Platform Interventions	68
Chronological Feed	68
Reshare Holdout	69
Like-Minded Sources Demotion	69
Appendix: Platform data types, classifiers and categories for purposes of these studies	69
Overview of platform data collected	69
Primary definitions	69
Platform activity data types	70
Content interactions	71
Time spent	72
User connections	73

Friends and followers	73
Other Facebook connection types	73
Connection actions	74
User attributes	74
Stated age and gender	74
Predicted political ideology	75
Predicted user location	75
Political interest	75
Content and entity attributes	75
Surfaces	76
Ideological segregation index	77
Favorability score index	77
Content and entity classifiers and categories	77
Aggregation levels	78
Appendix: Meta classifiers and categorization methods	80
Civic classifier	80
COVID-19 classifier	80
Topic classifier	81
News classifier	82
US predicted ideology classifier	83
Content and entity categories	84
Facebook and Instagram policy violation categories	84
Political figures	85
Misinformation Repeat Offenders	85
External classifiers and categorization methods	86
Classifiers	86
Content with slur words classification method	86
Incivility classifier	87
Civic events	89
Non-classifier-based content or entity categorization methods	89
Misinformation and Misleading information	89
Appendix: Paper-to-dataset mapping	92

The U.S. 2020 Facebook and Instagram Election Study documentation includes:

1. Study user guide

An overview of the study. Includes information on data access, collection, limitations, disclosures and categorization.

2. Datasets

A list of the released datasets and the collaborative papers that reference each. Includes

details on data structure, data categorization, participant demographics, experimental design and disclosures for each dataset.

3. Data dictionary

For each dataset, a data dictionary lists all variables included in that dataset. For each variable, it includes the variable name, description, type, group, map keys, aggregation methods and disclosures. The data dictionaries are included in the dataset's study page on ICPSR (they are separate files from this document).

4. Appendix

[The Appendix](#) for descriptions of platform data types and the text classifiers used in the study. A separate Glossary file includes descriptions of terms and classifiers.

This document is intended to be an iterative codebook, updated as additional data is released.

Part 1: Study overview

Introduction

The U.S. 2020 Facebook and Instagram Election Study (U.S. 2020 FIES) is a partnership between Meta and academic researchers to understand the impact of Facebook and Instagram on key political attitudes and behaviors during the U.S. 2020 election.

The academic research team is led by academic collaborators Professors Talia Stroud and Joshua A. Tucker. Professors Tucker and Stroud selected 15 additional researchers to collaborate on this effort, based on their expertise.

As results are published in peer-reviewed journals, Meta will make the replication datasets available to qualified academic researchers until June 30, 2031 to reproduce or extend the analyses. The release of datasets will be staggered, as not all papers will be published at the same time. Datasets for U.S. 2020 FIES papers that have already been published are documented in this codebook. The codebook will be updated on an ongoing basis as papers are published.

Please see the [ICPSR Social Media Archive project repository](#) for further information about the datasets, including information about applying for data access.

Citation

U.S. 2020 Facebook and Instagram Election Study. Inter-university Consortium for Political and Social Research [distributor]. <https://socialmediaarchive.org/collection/US2020>

In addition to this project wide citation, please cite the DOIs provided on ICPSR's website for each U.S. 2020 FIES dataset you use.

Data access

Meta will deposit the replication datasets with the Inter-university Consortium for Political and Social Research (ICPSR) Social Media Archive (SOMAR) until June 30, 2031, so that qualified academic researchers can reproduce or extend the analyses in published papers. The release of datasets will be staggered, as not all papers will be published at the same time. Applicants must indicate the dataset(s) to which they seek access in order to answer their research question(s). Appropriate projects will use these data to study **the role played by social media in elections in the United States** and must be for academic research conducted in an accountable, ethical manner. For more details, see the announcement for US2020 dataset applications on the [US2020 collection portal](#).

Applications for dataset access must be submitted through the ICPSR SOMAR [Infoready platform](#).

Project roles and collaboration overview

The U.S. 2020 FIES involves three groups: Meta, external academic researchers, and the National Opinion Research Center (NORC) at the University of Chicago, an independent, nonpartisan research institution contracted by Meta to execute parts of the research. The academic collaborators participated in the project without receiving any payment from Meta. For details on the roles and responsibilities of each group in the partnership, see [Project Roles & Data Transparency](#).

The papers published as part of this study were written by the academic collaborators with input from Meta researchers. Final control rights on each paper rest with the paper's lead academic authors.

Support

ICPSR maintains a general [FAQ site dedicated](#) to onboarding and access support, in addition to recorded training content found on the data's landing page.

If these resources do not answer your questions, or you have technical questions about the data itself, you can email somar-help@umich.edu. Include "U.S. 2020 FIES" in the email subject line and describe the problem in the body of the email, including mentions of the dataset(s) with which you need assistance.

Data overview

The U.S. 2020 FIES relies on Facebook and Instagram platform data collected by Meta, as well as data collected by NORC at the University of Chicago. The data collected by NORC included survey data and public records data in addition to mobile and desktop internet browsing behavior data from participants who consented to such data collection. The data collected by NORC is described in more detail on the [US2020 NORC study page](#) in SOMARs collection for the US2020 project.

Samples

Participants

Participants are people who were active on Facebook and/or Instagram as of August 2020, based in the United States and 18 years or older, and who consented to take part in the study.

Participants were recruited by invitations on Facebook and Instagram sent to a stratified random sample of active users. Accounts excluded from the sample include, using reasonable efforts, Meta employees, advertisers, predicted fake accounts on Facebook, and Instagram creator accounts and business accounts.

Participants responded to up to six waves of a survey and provided informed consent to have their platform activity data tracked for the purpose of this study (see 'Data Types' below for more detail). 353,533 participants were part of the study. This number represents all participants who took part in the first of the six survey waves and did not withdraw from the study with a request to have their data deleted.

The participant sample was stratified by covariates such as a user's number of active days, predicted location, and demographics to improve representativeness. We created survey weights to calibrate the participant samples to better estimate the number of adult monthly active users and reduce bias.

US Adult active users

US adult active users are Facebook and/or Instagram active users whose predicted home country is the United States, whose stated age on Facebook or predicted age on Instagram indicates they are 18 years of age or older, and who used Facebook or Instagram during a specific measurement period, such as within the previous 30 days (monthly) from a given date. Usage refers to when a user scrolls, clicks, types or navigates on the Facebook or Instagram app or website.

Aggregated statistics on the platform activity of U.S. users was produced by Meta for purposes of this study (See 'Data Types' below for more detail).

Data types

NORC-collected data

For this project, NORC collected several different data types about participants that are described below. In all cases, the data was collected by NORC and scrubbed of direct identifiers before NORC uploaded the data to Meta's Researcher Platform, a virtual clean room environment that is isolated from Meta's internal servers, to be combined and included in the study datasets. For more detailed descriptions about this data see the US2020 NORC study page in SOMARs collection for the US2020 project.

Individual-level survey data

Participants were invited to complete up to six surveys deployed over six waves: Wave 1 was following recruitment (August 31–September 12); Wave 2 was prior to experimental interventions (September 8–September 23); Wave 3 was prior to the election (October 9–October 23); Wave 4 was after the election (November 4–November 18); Wave 5 was post-election (December 9–December 23); and Wave 6 was post-inauguration (February 16–March 2).

The surveys asked questions about the study's key outcomes: political polarization, political knowledge, people's attitudes towards government and democracy, and political participation, including whether and how they vote. More details about the survey can be found on the US2020 NORC study page in SOMARs collection for the US2020 project.

Public records data

Two kinds of public records data were collected for this study: public voting records and public campaign donation data.

For the public voting records, NORC provided information to their subcontractor that owns the voter file. The subcontractor matched the voter file information to participants, and returned the data to NORC. NORC then attached the non-personally-identifying information to the survey responses.

The voter file data provided included information about participants who lived in states that legally allowed the use of the records for our study and consisted of:

- Indicator variables of whether the participant voted in 2020, 2018, and 2016
- Descriptor variables of the method used to vote (in person, absentee, etc.) for 2020, 2018, and 2016

The voter file data is limited to the following 30 states and territories: Alabama, Alaska, Arizona, Arkansas, California, Colorado, Delaware, DC, Florida, Idaho, Indiana, Iowa, Louisiana, Maryland, Massachusetts, Michigan, Mississippi, Nevada, New Jersey, New Mexico, New York, North Carolina, Ohio, Oklahoma, Pennsylvania, South Carolina, South Dakota, Vermont, Washington, and Wisconsin.

For the campaign donation data, NORC attempted to match participants to public FEC data and the Stanford Database on Ideology, Money in Politics, and Elections ([DIME](#)). NORC then attached that information to the survey responses. The data obtained from these sources includes:

- FEC records of how much and to which campaigns to which the user donated in amounts of \$200 or more
- Records of how much and to which campaigns users made smaller donations, as recorded by activist organizations (such as ActBlue, WinRed)

External passive web-tracking data

Select datasets share aggregated passive web-tracking data for a subset of experiment participants. These participants consented to have their mobile and/or desktop internet browsing behavior tracked by one of NORC's vendors (listed below) for the purpose of this study.

Participants in the platform intervention experiments were invited to provide NORC with passive data collection about the number of visits to different web domains (e.g., cnn.com) on their mobile and desktop devices, as well as time spent on apps on their mobile devices.

This data was collected by NORC's vendors, MDI Global and RealityMine, by using a Virtual Private Network (VPN) that consenting participants were asked to install on their devices. The

study collected website visits at the domain level (i.e., no URL-level granularity) and time spent at the application level. The domain-level and application-level data was then linked to the survey data by NORC.

Meta-collected data

For this project, Meta collected data summarizing participants' activity on Facebook and/or Instagram during the time periods specified in each study, and aggregated data on the platform activity of U.S. users during the same periods. This includes, for example, interactions with different types of content (i.e. views, reactions, comments, reshares and clicks) and entities (e.g. Facebook users, Pages, and groups; Instagram accounts) on these platforms. Go to [Appendix: Platform Data Types](#) for a complete list and description of platform data types and derived metrics.

User privacy

The research design across studies is built with privacy in mind. Steps were taken to remove direct identifiers from participants' individual-level data and limit publicly available quasi-identifying attributes of participants to only include those required to complete the studies. Data points aggregating over the US user population had a minimum of 100 potential contributing users for each quasi-identifying attribute(s) used in the aggregation.

Privacy and security practices have also been implemented in the access, storage and processing of the data. Data for the studies was stored on Meta's "Researcher Platform," a platform controlled through a Virtual Private Network (VPN) that is separate from Meta's business servers. Academic collaborators were restricted from downloading/exporting data outside the platform, and logging and monitoring was implemented to log researcher access of the data. Academic collaborators were required to sign and abide by the Researcher Data Agreement (RDA) in which they committed to certain privacy protective data practices, including a commitment not to make any attempt to re-identify individuals.

Participants in experiments had the option to exit the study prior to its conclusion. Those participants who declined to remain in a study could choose whether or not to delete the data that had been collected about them for the study. If they chose to delete the data, the participants were entirely removed from datasets on Researcher Platform and were not included in any analysis. If they did not choose to delete the data, all available data was included in the analysis.

Privacy and security practices will also be applied in the ICPSR SOMAR researcher environment, where replication data will be made available to additional qualifying researchers upon the

publication of the papers for reproduction and extension of the analyses. In order to gain access to the data in SOMAR, researchers will be required to submit an application and sign an agreement with ICPSR, and commit to certain privacy protective data practices, including a commitment not to make any attempt to re-identify individuals. Researchers will be restricted from downloading/exporting data outside the platform, and ICPSR will log and monitor activity on the platform.

Data disclosure

Many of the variables, classifiers, and other categorization methods described in this codebook were developed or utilized in novel ways for the purposes of U.S. 2020 FIES and may not be similarly operationalized by Meta in its core products or services. Accordingly, these datasets may only be used in accordance with their approval in research projects related to elections in the United States. Data types and related definitions for purposes of this study may differ from similar data types tracked or published by Meta in other contexts.

These variables were created using internal Meta company data sources available between 2020 and 2021. Data governance and sources may have changed since the study period and may not be representative of current operations at Meta. Due to potential data issues, such as logging errors in the data sources from which these variables are created, these variables may suffer from inconsistent data quality and incompleteness, which may lead to fluctuating accuracy. Within this codebook, we disclose known discrepancies.

Meta utilized a variety of quality assurance measures on the data represented in this codebook. This included a peer code review by both Meta researchers and the academic collaborators who participated in the original studies that produced the data shared in this codebook. Together, they conducted assessments of the descriptive statistics of the variables to identify any anomalies that may indicate errors in the underlying data.

Through this quality assurance process, as well as during the initial research process that produced these datasets, Meta discovered issues and fixed or reduced their impact where possible. Issues that were discovered but could not be fixed have been disclosed within this codebook by variable. Because of inherent limitations of quality assurance processes and imperfections of any dataset, it is likely that unidentified gaps still exist and thus cannot be disclosed.

Researchers using this data are responsible for conducting thorough data cleaning processes and are responsible for ensuring that their analyses are accurate. As mentioned above, it is expected that researchers may find issues with the data when conducting their analysis. We encourage researchers to share any findings with us. This may include, but is not necessarily limited to, data quality, validity, or fidelity issues. If issues arise, please contact us via somar-help@umich.edu.

Given the historical nature of the data and lapsed retention periods, we may not be able to fix the issues identified but, in such cases, will work to disclose them in the “Updates and Errata” section of this codebook.

Disclaimers

Meta works diligently and utilizes a variety of quality assurance measures to improve the accuracy, quality, and reliability of the data it shares for research purposes. However, given the volume of data released and the imperfection of any quality assurance process, inaccuracies persist. Meta makes no representation or warranty, express or implied, including without limitation, any warranties of fitness for a particular purpose or warranties as to the quality, accuracy or completeness of data or information. By accessing this data, researchers acknowledge that the data may contain some nonconformities, defects, or errors. Meta does not warrant that the data will meet the researcher's needs or expectations, that the use of the data will be uninterrupted, or that all nonconformities, defects, or errors can or will be corrected.

Part 2: Datasets

Metadata of Domains Shared on Facebook

About

Table	metadata_domains_shared_facebook
Level	Domain
Data	Domains shared on Facebook
Description	Information about the domains shared on Facebook
SOMAR Page	https://socialmediaarchive.org/record/13

Overview

Each row of the dataset corresponds to a domain (e.g., nytimes.com, foxnews.com, etc.). Each column of the dataset corresponds to a domain attribute, based on an internal or external classification method.

Domains included in this dataset were either:

- Shared to Facebook 20 or more times by US adult users with a predicted ideology score
- Categorized as untrustworthy, even if they did not meet this share threshold

This dataset can be used with the NORC-collected external passive web-tracking data to link the domains visited by consenting participants with the classification of domains used in this study.

Facebook Feed Intervention Experiment Participants

About

Table facebook_intervention_experiment_participants

Level Participant

Data Facebook activity data

Description Used to analyze the impact of platform experiments

SOMAR Page <https://socialmediaarchive.org/record/12>

Overview

Each row of the dataset corresponds to data on the platform activity of participating Facebook users' data. Each column contains a value, or set of values, that aggregates log data for this specific participant over a certain period of time.

Participants

The participants in this dataset include Facebook users recruited by Meta researchers to participate in the study who completed survey Waves 1 and 2.

Data categorization

This dataset contains metrics that summarize the attributes or behavior of these participants. This includes the following platform data types:

- User attributes
- User connections
- Time spent
- Content views

- Content engagement

[Go to the data types appendix](#) for descriptions of these categories.

Experimental design

This dataset is used to analyze the impact of the following platform experiments and their respective control groups, which participants were randomly assigned to:

- Chronological Feed
- Reshare Holdout
- Like-Minded Sources Demotion

Aggregation

Variables are computed using the following types of aggregations:

- Study period: duration of the experiment. Used to evaluate the impact of the platform experiment on participants' content engagement and views.
 - Start: September 23, 2020
 - End: December 22, 2020
- Pre-study period: the time period prior to the experiment. Used as a temporal benchmark for the participants' content engagement and views prior to the experiment.
 - Start: June 26, 2020
 - End: September 23, 2020
- Snapshot: metrics logged at a specific date (e.g., the day that the experiment began or ended).

Disclosures

Platform limitations

Given the discrepancy between factors such as devices used to access Facebook or how content is displayed across Facebook surfaces, there may be slight inaccuracies in the number of content views and engagement reported in this dataset. However, because these differences shouldn't vary meaningfully from day to day, we expect that any such potential discrepancy should be effectively consistent across the study period and across different types of users.

Reshares

Due to logging limitations, reshares of user posts to groups cannot be distinguished from reshares of user posts on their profiles. Since the latter is more common, all reshares of user posts are added into the **user** key for metrics by connection type, regardless of which of the two types of user posts they were. As a result, the reshare metrics broken down by connection type do not contain a **group** key.

Content engagement

Unless otherwise noted, post views and engagement metrics only include data for posts users viewed. Other Facebook surfaces, such as Marketplace, are excluded from those metrics, except for ones broken down by surface.

Data issues

We note the following known issues:

- **Content views** for the Chronological Feed experiment during the pre-study period:
 - Data is limited to the aggregation time period from August 7, 2020, to September 23, 2020.
 - The variable `view_count_chronological_feed_experiment_pre_study` calculates the number of times individual pieces of content were viewed by the participant in chronological order.
 - The variable `view_count_comparison_chronological_feed_experiment_pre_study` calculates the number of times individual pieces of content were viewed by the participant not in chronological order over this same period to allow for comparison.
- If a Page categorized as untrustworthy shared a post to a group not categorized as untrustworthy, the post is not categorized as untrustworthy.
- **Clicks** (`click_count*`):
 - Accuracy measures were not run on the source table for variables using the **clicks** metric. While the accuracy of the data is not guaranteed, Meta has robust testing and monitoring in place to assure the usability of data.

Filtering

Unless otherwise noted, post views and engagement metrics only include data corresponding to posts viewed in Feed. Other surfaces on Facebook, such as groups or News, are excluded from all post view metrics, except for metrics broken down by surface.

For purposes of this dataset, a view is counted whenever the post renders in the visible portion of their web browser or mobile device for more than 250 milliseconds.

Only interactions with organic content are included in these metrics. Sponsored content such as ads is not included.

Participants who withdrew their consent or who had deleted their Facebook account by the time this table was created are not included in the table.

Instagram Intervention Experiment Participants

About

Table	instagram_intervention_experiment_participants
Level	Participant
Data	Instagram activity data
Description	Used to analyze the impact of platform experiments

SOMAR Page <https://socialmediaarchive.org/record/10>

Overview

Each row of the dataset corresponds to data from a participant's Instagram user account. Each column contains a value, or set of values, which aggregates the data for this participant over a specified time period.

Participants

The participants in this dataset include Instagram users recruited by Meta researchers to participate in the study.

Invitations were sent to a random sample of Instagram monthly active users based in the United States, 18 years of age or older and eligible to receive general surveys as of August 17, 2020. Using reasonable efforts, Meta employees, advertisers and fake accounts were excluded from this sampling frame.

Data categorization

This dataset contains metrics that summarize the attributes or behavior of these participants. This includes the following platform data types:

- User attributes
- User connections
- Time spent

- Content views
- Content engagement

[Go to the data types appendix](#) for descriptions of these categories.

Experimental design

This dataset contains metrics used to analyze the impact of the following platform intervention experiments that participants were randomly assigned to:

- Chronological Feed (treatment group and control group)

Aggregation and filtering

Variables in this dataset are computed using the following types of aggregations:

- Study period: duration of the experiment.
 - Start: September 23, 2020
 - End: December 22, 2020
- Pre-study period: 90 days prior to the experiment.
 - Start: July 1, 2020
 - End: September 22, 2020
- Snapshot: metrics logged at a specific date, such as the day that the experiment began or ended.

Disclosures

Platform limitations

Given the discrepancy between factors such as devices used to access Instagram or how content is displayed across Instagram surfaces, there may be slight inaccuracies in the number of content views and engagement reported in this dataset. However, because these differences shouldn't vary meaningfully from day to day, we expect that any such potential discrepancy should be effectively consistent across the study period and across different types of users.

Missing values

Due to an error in the processing of the data, content views metrics exclude data for the following dates: October 8, 9, 10, 11 and 13, 2020.

Content engagement metrics are only available starting on July 17, 2020. To remain consistent, pre-study-period metrics are computed between July 17, 2020 and September 22, 2020.

Data issues

We note the following known issues:

- **Instagram chronological feed intervention** (`chronological_feed_intervention_active`):
 - This metric was implemented starting on September 30, 2020. Data from September 23, 2020, to September 29, 2020, is excluded.
- **Voting Information Center** metrics (`*day_count_voting_information_center`):
 - Some Voting Information Center content engagement and views metrics are reported as the number of days a participant initiated a relevant action.
- Content shared to Instagram from user's connections (`content_posted_from_connections_sept23_2020`):
 - This metric is estimated. Actual inventory size of content posted by connections is not logged on Instagram. This metric is computed as the total number of pieces of content created on September 23, 2020 (the date that the experiment began), from all of the accounts that the participant follows, plus content shared that includes hashtags that the participant follows. It serves as a pre-study baseline to compare the experiment's effects on participants in the treatment groups with high or low inventory.
- Breakdowns by **connection type** (`*by_connection_type`):
 - Connection type is calculated using data on the participant's follows that is measured once daily. There may be minor inconsistencies between accounts the participant is following throughout the day. Content created by accounts the participant was not following at the time it was measured are assumed to be unconnected accounts.
- Breakdowns by **content category** for the pre-study period: (`*count_by_content_category_pre_study`):
 - Values for the categories incivility and content with slur words used in these variables exclude data from the period between July 17, 2020, and August 3, 2020.

Filtering

Unless otherwise noted, content views and engagement metrics in this dataset only include data corresponding to the following Instagram surfaces: feed, Explore, profile, stories and hashtag search.

Metrics for posts and reshares only include content shared to Instagram feed.

For the purposes of this dataset, on Instagram a view is counted whenever a single pixel of the relevant content is displayed on a participant’s screen for any period of time.

Only interactions with organic content are included in this dataset; ads are not included.

Participants who withdrew their consent after participating in the platform intervention or who deleted their Instagram account by the time this dataset was created are not included.

Passive-Tracking Participants’ Daily Views of Facebook Posts with Civic News Domains

About

Table	passive_tracking_participant_daily_views_facebook_posts_civic_news_domains
Level	Domain
Data	Facebook activity data on participant views of posts with links
Description	Data on the civic news domains that participants in the platform interventions control group viewed in Facebook posts daily
SOMAR Page	https://socialmediaarchive.org/record/14

Overview

This dataset contains domain-level metrics from Facebook activity data for the subset of participants in the platform intervention experiment control group who consented to have their internet browsing behavior tracked by an external partner for the purpose of this study.

Each row corresponds to a set of civic news domains linked from Facebook posts that study participants viewed on that day. Columns contain a date, whether the measures are computed for domains in the top 90% or 99% of views, and aggregate measures of participants’ views on those posts.

Domains in this dataset consist of those:

- Classified as civic (defined as related to politics or social issues) and news by Meta internal classifiers
- Shared more than 100 times to Facebook by U.S. users during the study period
- Viewed by participants in posts in their Feed

[Go to the classifiers appendix](#) for more information on classifiers.

Participants

The participants in this dataset include Facebook users recruited by Meta researchers to participate in the study who were assigned to the platform interventions control group and who consented to have their internet browsing behavior tracked by an external partner for the purposes of the study.

Invitations were sent to a random sample of Facebook monthly active users based in the United States, 18 years of age or older and eligible to receive general surveys as of August 17, 2020. Using reasonable efforts, Meta employees, advertisers and fake accounts were excluded from this sampling frame.

Data categorization

This dataset contains metrics that summarize attributes and views of domains by participants. This includes the following platform data types:

- Content views
- Content attributes
- User attributes

[Go to the data types appendix](#) for descriptions of these categories.

Aggregation

Variables are computed using the following types of aggregations:

- Daily during the study period
 - Start: September 1, 2020
 - End: December 31, 2020
- Domain set: set of domains analyzed
 - civic news domains accounting for 99% of views by Facebook US adult users
 - civic news domains accounting for 90% of views by Facebook US adult users

- Ideology threshold: specific threshold used to predict whether Facebook US adult users are liberal or conservative, depending on the score assigned by the [US predicted ideology classifier](#).
 - Main thresholds used in the analysis: 0.35 and 0.65

Disclosures

None

Passive-Tracking Participants' Views of Facebook Posts with Civic News Domains

About

Table	<code>passive_tracking_participant_views_facebook_posts_civic_news_domains</code>
Level	Domain
Data	Facebook activity data on participant views of posts with links
Description	Data on the civic news domains that participants in the platform interventions control group viewed in Facebook posts
SOMAR Page	https://socialmediaarchive.org/record/41

Overview

This dataset contains domain-level metrics from Facebook activity data for the subset of participants in the platform interventions experiment control groups who consented to have their internet browsing behavior tracked by an external partner for the purpose of this study.

Each row corresponds to a set of civic news domains participants viewed in Facebook posts on that day. Columns contain a date, whether the measures are computed for domains in the top 90% or 99% of views, and values that are aggregate measures of participants' views on those posts.

Domains in this dataset consist of those:

- Classified as civic (defined as related to politics or social issues) and news by Meta internal classifiers
- Shared more than 100 times to Facebook by U.S. users during the study period
- Viewed by participants in posts in their Feed

[Go to the classifiers appendix](#) for more information on classifiers.

Participants

The participants in this dataset include Facebook users recruited by Meta researchers to participate in the study who were assigned to the platform interventions control group and who consented to have their internet browsing behavior tracked by an external partner for the purposes of the study.

Invitations were sent to a random sample of Facebook monthly active users based in the United States, 18 years of age or older and eligible to receive general surveys as of August 17, 2020. Using reasonable efforts, Meta employees, advertisers and fake accounts were excluded from this sampling frame.

Data categorization

This dataset contains metrics that summarize attributes and views of domains by participants. This includes the following platform data types:

- Content views
- Audience size
- Content attributes
- User attributes

[Go to the data types appendix](#) for descriptions of these categories.

Aggregation

Variables are computed using the following types of aggregations:

- Over the study period
 - Start: September 1, 2020
 - End: December 31, 2020
- Domain level: metrics at the news source level

Disclosures

None

Exposure to Facebook Posts with Civic News Domains

About

Table exposure_facebook_posts_with_civic_news_domains

Level Domain

Data Aggregated Facebook activity data for US adult monthly active users

Description Data on the civic news domains that users viewed in Facebook posts in their Feed

SOMAR Page <https://socialmediaarchive.org/record/31>

Overview

This dataset contains domain-level metrics from aggregated Facebook activity data for US adult monthly active users. These metrics measure **audience size** and **views** of posts with links to civic news domains that users viewed on their Facebook Feeds from September 1, 2020, to February 1, 2021.

Rows correspond to domains. Each domain is represented by four rows, one for each type of post owner (**user, Page, group, all**). Each column contains a value, or set of values, that aggregates platform log data for the specific domain over the study period.

Domains in this dataset consist of those:

- Classified as civic (defined as related to politics or social issues) and news by Meta internal classifiers
- Shared more than 100 times to Facebook by U.S. users during the study period
- Viewed by U.S. users in their Feed

[Go to the classifiers appendix](#) for more information on classifiers.

Active users

Aggregated platform activity data for purposes of this dataset includes that of Facebook users who were active on the platform between September 1, 2020, and February 1, 2021, whose predicted home country is the United States and stated age indicates they are 18 years or older.

Data categorization

This dataset contains metrics that summarize data from domains viewed during the study period. This includes the following platform data types:

- Content views
- Audience size
- Content attributes
- User attributes

[Go to the data types appendix](#) for descriptions of these categories.

Aggregation

Variables are computed using the following types of aggregations:

- During the study period
 - Start: September 1, 2020
 - End: February 1, 2021
- Domain level: metrics calculated at the news source level
- Post owner type: entity type that shared the post with the domain
 - All: posts from users, Pages and groups
 - User: posts from users to their profile
 - Page: posts from Pages
 - Group: posts from users to groups

Disclosures

Platform limitations

Exposed audience metrics are aggregated by domain and calculated by adding the daily exposed audience count for each post with a link to a URL within this domain, for every day over the study period.

- If a user viewed two different posts with a link to the same domain in a single day, they are counted in the exposed audience count twice for that day.
- If a user viewed the same post with a link to the domain but on different days, they are counted in that post's exposed audience count for both days.
- If a user viewed posts with links to two different URLs within the same domain in a single day, they are counted in the exposed audience count twice for that day.

Categories and classifiers

Domains categorized as untrustworthy

This dataset includes information about domains categorized as “untrustworthy sources” for the purposes of this study; it intends to capture domains that were repeatedly fact-checked as false by at least one independent fact-checking partner resulting in strikes under Meta’s Misinformation Repeat Offenders Policy. However, due to data limitations, the number of strikes is estimated using the number of URLs that received a false rating by one of the independent third-party fact-checking partners Meta works with through its fact-checking program. The untrustworthy domains category in this dataset was defined and used solely for the U.S. 2020 FIES and is not the same as how Meta defined misinformation Repeat Offenders in 2020 nor how [Meta defines it today](#).

The study uses a more expansive operationalization of an untrustworthy domain than was originally pre-registered, and will necessarily overestimate participant exposure to domains with two or more strikes. This is because in practice, third-party fact-checker ratings are an imperfect measure of misinformation strikes. Specifically, the number of individual pieces of content that are fact-checked as false does not necessarily equal the number of strikes an entity accrues under Meta’s Misinformation Repeat Offender Policy (see [Meta’s Transparency Center](#) for more details).

Exposure to Facebook Posts with Civic News URLs

About

Table	exposure_facebook_posts_with_civic_news_urls
Level	URL
Data	Aggregated Facebook activity data for US adult monthly active users
Description	Data on the civic news URLs that users viewed in Facebook posts in their Feed
SOMAR Page	https://socialmediaarchive.org/record/32

Overview

This dataset contains URL-level metrics from aggregated Facebook activity data for US adult monthly active users. These metrics measure **audience size** and **views** of posts with links to civic news URLs that users viewed on their Facebook Feeds from September 1, 2020, to February 1, 2021.

Rows correspond to URLs. Each URL is represented by four rows, one for each type of post owner (**user, Page, group, all**). Each column contains a value, or set of values, that aggregates platform log data for the specific URL over the study period.

URLs in this dataset consist of those:

- Classified as civic (defined as related to politics or social issues) and news by Meta internal classifiers
- Shared more than 100 times to Facebook by U.S. users during the study period
- Viewed by U.S. users in their Feed

[Go to the classifiers appendix](#) for more information on classifiers.

Active users

Aggregated platform activity data for purposes of this dataset includes that of Facebook users who were active on the platform between September 1, 2020, and February 1, 2021, whose predicted home country is the United States and stated age indicates they are 18 years or older.

Data categorization

This dataset contains metrics that summarize attributes and views of URLs by users. This includes the following platform data types:

- Content views
- Audience size
- Content attributes
- User attributes

[Go to the data types appendix](#) for descriptions of these categories.

Aggregation

Variables are computed using the following types of aggregations:

- During the study period
 - Start: September 1, 2020
 - End: February 1, 2021
- URL level: metrics calculated at the news story level
- Post owner type: entity type that shared the post with the URL
 - All: posts from users, Pages and groups
 - User: posts from users to their profile
 - Page: posts from Pages
 - Group: posts from users to groups

Disclosures

Platform limitations

Exposed audience metrics are aggregated by domain and calculated by adding the daily exposed audience count for each post with a link to a URL within this domain, for every day over the study period.

- If a user viewed two different posts with a link to the same domain in a single day, they are counted in the exposed audience count twice for that day.
- If a user viewed the same post with a link to the domain but on different days, they are counted in that post's exposed audience count for both days.
- If a user viewed posts with links to two different URLs within the same domain in a single day, they are counted in the exposed audience count twice for that day.

Categories and classifiers

Domains categorized as untrustworthy

This dataset includes information about domains categorized as “untrustworthy sources” for the purposes of this study; it intends to capture domains that were repeatedly fact-checked as false by at least one independent fact-checking partner resulting in strikes under Meta’s Misinformation Repeat Offenders Policy. However, due to data limitations, the number of strikes is estimated using the number of URLs that received a false rating by one of the independent third-party fact-checking partners Meta works with through its fact-checking program. The untrustworthy domains category in this dataset was defined and used solely for the U.S. 2020 FIES and is not the same as how Meta defined misinformation Repeat Offenders in 2020 nor how [Meta defines it today](#).

The study uses a more expansive operationalization of an untrustworthy domain than was originally pre-registered, and will necessarily overestimate participant exposure to domains with two or more strikes. This is because in practice, third-party fact-checker ratings are an imperfect measure of misinformation strikes. Specifically, the number of individual pieces of content that are fact-checked as false does not necessarily equal the number of strikes an entity accrues under Meta’s Misinformation Repeat Offender Policy (see [Meta’s Transparency Center](#) for more details).

URLs categorized as misinformation

For the purposes of this study, when a URL has received two or more ratings of partial or complete falsity from independent fact-checkers, the strictest rating (**false**) is used in this dataset. If any rating is **false**, then the value for this variable will be **true** (URL categorized as misinformation).

URLs classified as related to COVID-19

Meta’s COVID-19 classifier classifies individual posts and not the URLs shared in those posts. As a result, a URL’s classification can vary depending on the post in which it was shared. For the purposes of this study, to address this issue, we computed the number of times that each possible predicted topic was associated with a URL. If more than 50% of those posts were classified as related to COVID-19, the URL was assigned the COVID-19 topic.

URLs classified as news

Meta’s topic classifier classifies individual posts and not the URLs shared in those posts. As a result, a URL’s classification can vary depending on the post in which it was shared. For the purposes of this study, to address this issue, we computed the number of times that each possible predicted topic was associated with a URL, then assigned it the most frequently predicted topic.

Engagement with Facebook Posts with Civic News Domains

About this dataset

Table	engagement_facebook_posts_with_civic_news_domains
Level	Domain
Data	Aggregated Facebook activity data for US adult monthly active users
Description	Data on the civic news domains in Facebook posts that users engaged with in their Feed
SOMAR Page	https://socialmediaarchive.org/record/33

Overview

This dataset contains domain-level metrics from aggregated Facebook activity data for US adult monthly active users. These metrics measure **audience size** and **engagement** for posts with links to civic domains users engaged with from September 1, 2020, to February 1, 2021.

Rows correspond to domains. Each domain is represented by four rows, one for each type of post owner (**user, Page, group, all**). Each column contains a value, or set of values, that aggregates platform log data for the specific domain over the study period.

Domains in this dataset consist of those:

- Classified as civic (defined as related to politics or social issues) and news by Meta internal classifiers
- Shared more than 100 times to Facebook by U.S. users during the study period
- Engaged with by U.S. users in their Feed

[Go to the classifiers appendix](#) for more information on classifiers.

Active users

Aggregated platform activity data for purposes of this dataset includes that of Facebook users who were active on the platform between September 1, 2020, and February 1, 2021, whose predicted home country is the United States and stated age indicates they are 18 years or older.

Data categorization

This dataset contains metrics that summarize data for domains in posts users engaged with during the study period. This includes the following platform data types:

- Content engagement
- Audience size
- Content attributes
- User attributes

[Go to the data types appendix](#) for descriptions of these categories.

Aggregation

Variables are computed using the following types of aggregations:

- During the study period
 - Start: September 1, 2020
 - End: February 1, 2021
- Domain level: metrics calculated at the news source level
- Post owner type: entity type that shared the post with the domain
 - All: posts from users, Pages and groups
 - User: posts from users to their profile
 - Page: posts from Pages
 - Group: posts from users to groups

Disclosures

Platform limitations

Clicks are counted once per post and per day if a user clicked on any component of a post. This includes clicks on external URLs, to “See More” and expand the text of a post, on the post author’s name, to enlarge a picture, to reshare, as well as to add a like, reaction, comment or any other type of engagement with the post.

Reshares include those generated on any Facebook surface because logging limitations prevent us from identifying where on the platform a user was when they reshared a post. Due to a logging issue, reshares of user posts cannot be disaggregated into reshares on their profile and reshares to a group. We count them all as reshares on their profile. As a result, the value for this metric will be 0 for rows where the **post owner** is a **group**.

Engaged audience metrics are aggregated by URL and calculated by adding the daily engaged audience count for each post with a link to the URL, for every day over the study period.

- If an active user engaged with two different posts with a link to the same URL in a single day, they are counted in the engaged audience count twice for that day.
- If a user engaged with the same post with a link to the URL but on different days, they are counted in that post's engaged audience count for both days.
- If a user engaged in more than one way with a post with a link to the URL, they are counted in the engaged audience count once.

Categories and classifiers

Domains categorized as untrustworthy

This dataset includes information about domains categorized as “untrustworthy sources” for the purposes of this study; it intends to capture domains that were repeatedly fact-checked as false by at least one independent fact-checking partner resulting in strikes under Meta’s Misinformation Repeat Offenders Policy. However, due to data limitations, the number of strikes is estimated using the number of URLs that received a false rating by one of the independent third-party fact-checking partners Meta works with through its fact-checking program. The untrustworthy domains category in this dataset was defined and used solely for the U.S. 2020 FIES and is not the same as how Meta defined misinformation Repeat Offenders in 2020 nor how [Meta defines it today](#).

The study uses a more expansive operationalization of an untrustworthy domain than was originally pre-registered, and will necessarily overestimate participant exposure to domains with two or more strikes. This is because in practice, third-party fact-checker ratings are an imperfect measure of misinformation strikes. Specifically, the number of individual pieces of content that are fact-checked as false does not necessarily equal the number of strikes an entity accrues under Meta’s Misinformation Repeat Offender Policy (see [Meta’s Transparency Center](#) for more details).

Engagement with Facebook Posts with Civic News URLs

About

Table	engagement_facebook_posts_with_civic_news_urls
Level	URL
Data	Aggregated Facebook activity data for US adult monthly active users
Description	Data on the civic news URLs in Facebook posts that users engaged with in their Feed
SOMAR Page	https://socialmediaarchive.org/record/34

Overview

This dataset contains URL-level metrics from aggregated Facebook activity data for US adult monthly active users. These metrics measure **audience size** and **engagement** for posts with links to civic news URLs users viewed in their Feed from September 1, 2020, to February 1, 2021.

Rows correspond to URLs. Each URL is represented by four rows, one for each type of post owner (**user, Page, group, all**). Each column contains a value, or set of values, that aggregates platform log data for the specific URL over the study period.

URLs in this dataset consist of those:

- Classified as civic (defined as related to politics or social issues) and news by Meta internal classifiers
- Shared more than 100 times to Facebook by U.S. users during the study period
- Engaged with by U.S. users in their Feed

Refer to the [classifiers appendix](#) for more information on classifiers.

Active users

Aggregated platform activity data for purposes of this dataset includes that of Facebook users who were active on the platform between September 1, 2020, and February 1, 2021, whose predicted home country is the United States and stated age indicates they are 18 years or older.

Data categorization

This dataset contains metrics that summarize data for URLs in posts users engaged with during the study period. This includes the following platform data types:

- Content engagement
- Audience size
- User attributes

[Go to the data types appendix](#) for descriptions of these categories.

Aggregation

Variables are computed using the following types of aggregations:

- During the study period
 - Start: September 1, 2020
 - End: February 1, 2021
- URL level: metrics calculated at the news story level
- Post owner type: entity type that shared the post with the URL
 - All: posts from users, Pages and groups
 - User: posts from users to their profile
 - Page: posts from Pages
 - Group: posts from users to groups

Disclosures

Platform limitations

Clicks are counted once per post and per day if a user clicked on any component of a post. This includes clicks on external URLs, to “See More” and expand the text of a post, on the post author’s name, to enlarge a picture, to reshare, as well as to add a like, reaction, comment or any other type of engagement with the post.

Reshares include those generated on any Facebook surface because logging limitations prevent us from identifying where on the platform a user was when they reshared a post. Due to a logging issue, reshares of user posts cannot be disaggregated into reshares on their profile and reshares to a group. We count them all as reshares on their profile. As a result, the value for this metric will be 0 for rows where the **post owner** is a **group**.

Engaged audience metrics are aggregated by URL and calculated by adding the daily engaged audience count for each post with a link to the URL, for every day over the study period.

- If an active user engaged with two different posts with a link to the same URL in a single day, they are counted in the engaged audience count twice for that day.
- If a user engaged with the same post with a link to the URL but on different days, they are counted in that post's engaged audience count for both days.
- If a user engaged in more than one way with a post with a link to the URL, they are counted in the engaged audience count once.

Categories and classifiers

Untrustworthy domains

This dataset includes information about domains categorized as “untrustworthy sources” for the purposes of this study; it intends to capture domains that were repeatedly fact-checked as false by at least one independent fact-checking partner resulting in strikes under Meta’s Misinformation Repeat Offenders Policy. However, due to data limitations, the number of strikes is estimated using the number of URLs that received a false rating by one of the independent third-party fact-checking partners Meta works with through its fact-checking program. The untrustworthy domains category in this dataset was defined and used solely for the U.S. 2020 FIES and is not the same as how Meta defined misinformation Repeat Offenders in 2020 nor how [Meta defines it today](#).

The study uses a more expansive operationalization of an untrustworthy domain than was originally pre-registered, and will necessarily overestimate participant exposure to domains with two or more strikes. This is because in practice, third-party fact-checker ratings are an imperfect measure of misinformation strikes. Specifically, the number of individual pieces of content that are fact-checked as false does not necessarily equal the number of strikes an entity accrues under Meta’s Misinformation Repeat Offender Policy (see [Meta’s Transparency Center](#) for more details).

URLs categorized as misinformation

For the purposes of this study, when a URL has received two or more ratings of partial or complete falsity from independent fact-checkers, the strictest rating (**false**) is used in this dataset. If any rating is **false**, then the value for this variable will be **true** (URL categorized as misinformation).

URLs classified as related to COVID-19

Meta’s COVID-19 classifier classifies individual posts and not the URLs shared in those posts. As a result, a URL’s classification can vary depending on the post in which it was shared. For the purposes of this study, to address this issue, we computed the number of times that each possible predicted topic was associated with a URL. If more than 50% of those posts were classified as related to COVID-19, the URL was assigned the COVID-19 topic.

Potential Exposure to Facebook Posts with Civic News Domains

About this dataset

Table	potential_exposure_facebook_posts_with_civic_news_domains
Level	Domain
Data	Aggregated Facebook activity data for US adult monthly active users
Description	Data on the civic news domains in Facebook posts that users may have viewed in their Feed

SOMAR Page <https://socialmediaarchive.org/record/35>

Overview

This dataset contains domain-level metrics from aggregated Facebook activity data for US adult monthly active users. These metrics measure the potential audience of posts with links to political domains users could have viewed in their Feed from September 1, 2020, to February 1, 2021, because such a post was shared by one of their connections.

Rows correspond to domains. Each domain is represented by four rows, one for each type of post owner (**user**, **Page**, **group**, **all**). Each column contains a value, or set of values, that aggregates platform log data for the specific domain over the study period.

Domains in this dataset consist of those:

- Classified as civic (defined as related to politics or social issues) and news by Meta internal classifiers
- Shared more than 100 times to Facebook by U.S. users during the study period
- Potentially viewed by U.S. users in their Feed because such a post was shared by one of their US-based connections

Refer to the [classifiers appendix](#) for more information on classifiers.

Active users

Aggregated platform activity data for purposes of this dataset includes that of Facebook users who were active on the platform between September 1, 2020, and February 1, 2021, whose predicted home country is the United States and stated age indicates they are 18 years or older.

Data categorization

This dataset contains metrics that summarize data for URLs in posts users could have viewed during the study period. This includes the following platform data types:

- Audience size
- Content attributes
- User attributes

[Go to the data types appendix](#) for descriptions of these categories.

Aggregation

Variables are computed using the following types of aggregations:

- During the study period
 - Start: September 1, 2020
 - End: February 1, 2021
- Domain level: metrics calculated at the news source level
- Post owner type: entity type that shared the post with the domain
 - All: posts from users, Pages and groups
 - User: posts from users to their profile
 - Page: posts from Pages
 - Group: posts from users to groups

Disclosures

Platform limitations

Potential audience metrics are aggregated by domain and calculated by adding the daily potential audience count for each post with a link to a URL within this domain, for every day over the study period.

- If a user potentially viewed posts with links to two different URLs within the same domain in a single day, they are counted in the potential audience count twice for that day.
- If a user potentially viewed two different posts with a link to the same domain in a single day, they are counted in the potential audience count twice for that day.

- If a user potentially viewed the same post with a link to the domain but on different days, they are counted in that post’s potential audience count for both days.

Posts by Pages that had been deleted at the time that the data was pseudonymized (90 days after post creation) are excluded from this metric.

Untrustworthy domains

This dataset includes information about domains categorized as “untrustworthy sources” for the purposes of this study; it intends to capture domains that were repeatedly fact-checked as false by at least one independent fact-checking partner resulting in strikes under Meta’s Misinformation Repeat Offenders Policy. However, due to data limitations, the number of strikes is estimated using the number of URLs that received a false rating by one of the independent third-party fact-checking partners Meta works with through its fact-checking program. The untrustworthy domains category in this dataset was defined and used solely for the U.S. 2020 FIES and is not the same as how Meta defined misinformation Repeat Offenders in 2020 nor how [Meta defines it today](#).

The study uses a more expansive operationalization of an untrustworthy domain than was originally pre-registered, and will necessarily overestimate participant exposure to domains with two or more strikes. This is because in practice, third-party fact-checker ratings are an imperfect measure of misinformation strikes. Specifically, the number of individual pieces of content that are fact-checked as false does not necessarily equal the number of strikes an entity accrues under Meta’s Misinformation Repeat Offender Policy (see [Meta’s Transparency Center](#) for more details).

Potential Exposure to Facebook Posts with Civic News URLs

About this dataset

Table	potential_exposure_facebook_posts_with_civic_news_urls
Level	URL
Data	Aggregated Facebook activity data for US adult monthly active users
Description	Data on the civic news URLs in Facebook posts that users may have viewed in their Feed

SOMAR Page <https://socialmediaarchive.org/record/36>

Overview

This dataset contains URL-level metrics from aggregated Facebook activity data for US adult monthly active users. These metrics measure the potential audience of posts with links to political URLs that users could have viewed in their Feed from September 1, 2020, to February 1, 2021, because a post containing such a political URL was shared by one of their connections.

Rows correspond to URLs. Each URL is represented by four rows, one for each type of post owner (**user**, **Page**, **group**, **all**). Each column contains a value, or set of values, that aggregates platform log data for the specific domain over the study period.

URLs in this dataset consist of those:

- Classified as civic (defined as related to politics or social issues) and news by Meta internal classifiers
- Shared more than 100 times to Facebook by U.S. users during the study period
- Potentially viewed by U.S. users in their Feed because such a post was shared by one of their US-based connections

Active users

Aggregated platform activity data for purposes of this dataset includes that of Facebook users who were active on the platform between September 1, 2020, and February 1, 2021, whose predicted home country is the United States and stated age indicates they are 18 years or older.

Data categorization

This dataset contains metrics that summarize data for URLs in posts users could have viewed during the study period. This includes the following platform data types:

- Content engagement
- Audience size
- Content attributes
- User attributes

[Go to the data types appendix](#) for descriptions of these categories.

Aggregation

Variables are computed using the following types of aggregations:

- During the study period
 - Start: September 1, 2020
 - End: February 1, 2021
- URL level: metrics calculated at the news story level
- Post owner type: entity type that shared the post with the domain
 - All: posts from users, Pages and groups
 - User: posts from users to their profile
 - Page: posts from Pages
 - Group: posts from users to groups

Disclosures

Platform Limitations

Clicks are counted once per post and per day if a user clicked on any component of a post. This includes clicks on external URLs, to “See More” and expand the text of a post, on the post author’s name, to enlarge a picture, to reshare, as well as to add a like, reaction, comment or any other type of engagement with the post.

Reshares include those generated on any Facebook surface because logging limitations prevent us from identifying where on the platform a user was when they reshared a post. Due to a logging issue, reshares of user posts cannot be disaggregated into reshares on their profile and reshares to a group. We count them all as reshares on their profile. As a result, the value for this metric will be 0 for rows where the **post owner** is a **group**.

Engaged audience metrics are aggregated by URL and calculated by adding the daily engaged audience count for each post with a link to the URL, for every day over the study period.

- If an active user engaged with two different posts with a link to the same URL in a single day, they are counted in the engaged audience count twice for that day.
- If a user engaged with the same post with a link to the URL but on different days, they are counted in that post’s engaged audience count for both days.
- If a user engaged in more than one way with a post with a link to the URL, they are counted in the engaged audience count once.

Potential audience metrics are aggregated by domain and calculated by adding the daily potential audience count for each post with a link to a URL within this domain, for every day over the study period.

- If a user potentially viewed posts with links to two different URLs within the same domain in a single day, they are counted in the potential audience count twice for that day.
- If a user potentially viewed two different posts with a link to the same domain in a single day, they are counted in the potential audience count twice for that day.
- If a user potentially viewed the same post with a link to the domain but on different days, they are counted in that post’s potential audience count for both days.

Posts by Pages that had been deleted at the time that the data was pseudonymized (90 days after post creation) are excluded from this metric.

Categories and classifiers

Untrustworthy domains

This dataset includes information about domains categorized as “untrustworthy sources” for the purposes of this study; it intends to capture domains that were repeatedly fact-checked as false by at least one independent fact-checking partner resulting in strikes under Meta’s Misinformation Repeat Offenders Policy. However, due to data limitations, the number of strikes is estimated using the number of URLs that received a false rating by one of the independent third-party fact-checking partners Meta works with through its fact-checking program. The untrustworthy domains category in this dataset was defined and used solely for the U.S. 2020 FIES and is not the same as how Meta defined misinformation Repeat Offenders in 2020 nor how [Meta defines it today](#).

The study uses a more expansive operationalization of an untrustworthy domain than was originally pre-registered, and will necessarily overestimate participant exposure to domains with two or more strikes. This is because in practice, third-party fact-checker ratings are an imperfect measure of misinformation strikes. Specifically, the number of individual pieces of content that are fact-checked as false does not necessarily equal the number of strikes an entity accrues

under Meta’s Misinformation Repeat Offender Policy (see [Meta’s Transparency Center](#) for more details).

URLs categorized as misinformation

For the purposes of this study, when a URL has received two or more ratings of partial or complete falsity from independent fact-checkers, the strictest rating (**false**) is used in this dataset. If any rating is **false**, then the value for this variable will be **true** (URL categorized as misinformation).

URLs classified as related to COVID-19

Meta’s COVID-19 classifier classifies individual posts and not the URLs shared in those posts. As a result, a URL’s classification can vary depending on the post in which it was shared. For the purposes of this study, to address this issue, we computed the number of times that each possible predicted topic was associated with a URL. If more than 50% of those posts were classified as related to COVID-19, the URL was assigned the COVID-19 topic.

Coexposure to Facebook Posts with Civic News Domains

About this dataset

Table	coexposure_facebook_posts_with_civic_news_domains
Level	Domain (dyad)
Data	Aggregated Facebook activity data for US adult monthly active users
Description	Data on pairs of civic news domains shared in Facebook posts that users viewed in their Feed

SOMAR Page <https://socialmediaarchive.org/record/37>

Overview

This dataset contains domain-dyad-level metrics from aggregated Facebook activity data for US adult monthly active users.

These metrics measure the coexposed audience size and views of posts with links to civic news domains from September 1, 2020, to February 1, 2021. This is tracked for users who viewed posts with links to domain 1 and posts with links to domain 2.

Rows correspond to pairs of domains. Each pair of domains is represented by one row. Each column contains a value, or set of values, that aggregates platform log data for the specific pair of domains over the study period.

Domains in this dataset consist of those:

- Classified as civic (defined as related to politics or social issues) and news by Meta internal classifiers
- Shared more than 100 times to Facebook by U.S. users during the study period
- Viewed by U.S. users in their Feed

Due to the large size of this dataset, the table only includes domains that are part of the subset of nodes that represent the core of the network, as defined by all civic news domains that meet the criteria above. This subset of the dataset was identified using the disparity filter algorithm selected by the academic collaborators. Additional technical details on the definition and implementation of this method can be found in the Appendix to González-Bailón et al (2023) "Asymmetric Ideological Segregation in Exposure to News on Facebook."

Active users

Aggregated platform activity data for purposes of this dataset includes that of Facebook users who were active on the platform between September 1, 2020, and February 1, 2021, whose predicted home country is the United States and stated age indicates they are 18 years or older.

Data categorization

This dataset contains metrics that summarize attributes and views of domains by participants. This includes the following platform data types:

- Content views
- Audience size
- Content attributes
- User attributes
- Political interest

[Go to the data types appendix](#) for descriptions of these categories.

Aggregation

Variables are computed separately for each pair of domains and week. The data is stored as separate files, one for each week. The name of each file ends with the last date of the specific week for which it contains data, for example `_2020-09-07.csv`.

Disclosures

Platform Limitations

Coexposed audience count

Coexposed audience metrics are aggregated by pairs of domains and calculated by adding the daily coexposed audience count for each pair of posts, each of them containing a link to a URL within the domains in the pair, for every day over the study period. User views of a domain, prior to aggregating them to pairs of domains, could not be fully deduplicated.

- If a user viewed posts with links to two different URLs within the same domain in a single day, they are counted in the exposed audience count twice for that day.
- If a user viewed two different posts with a link to the same domain in a single day, they are counted in the exposed audience count twice for that day.
- If a user viewed the same post with a link to the domain but on different days, they are counted in that post's exposed audience count for both days.

Categories and Classifiers

Untrustworthy domains

This dataset includes information about domains categorized as “untrustworthy sources” for the purposes of this study; it intends to capture domains that were repeatedly fact-checked as false by at least one independent fact-checking partner resulting in strikes under Meta’s Misinformation Repeat Offenders Policy. However, due to data limitations, the number of strikes is estimated using the number of URLs that received a false rating by one of the independent third-party fact-checking partners Meta works with through its fact-checking program. The untrustworthy domains category in this dataset was defined and used solely for the U.S. 2020 FIES and is not the same as how Meta defined misinformation Repeat Offenders in 2020 nor how [Meta defines it today](#).

The study uses a more expansive operationalization of an untrustworthy domain than was originally pre-registered, and will necessarily overestimate participant exposure to domains with two or more strikes. This is because in practice, third-party fact-checker ratings are an imperfect measure of misinformation strikes. Specifically, the number of individual pieces of content that are fact-checked as false does not necessarily equal the number of strikes an entity accrues under Meta’s Misinformation Repeat Offender Policy (see [Meta’s Transparency Center](#) for more details).

Coexposure to Facebook Posts with Civic News URLs

About this dataset

Table	coexposure_facebook_posts_with_civic_news_urls
Level	URL (dyad)
Data	Aggregated Facebook activity data for US adult monthly active users
Description	Data on pairs of civic news URLs shared in Facebook posts that users viewed in their Feed

SOMAR Page <https://socialmediaarchive.org/record/38>

Overview

This dataset contains URL-dyad-level metrics from aggregated Facebook activity data for US adult monthly active users.

These metrics measure the audience size and content views of posts with links to civic news URLs from September 1, 2020, to February 1, 2021. This is tracked for participants who viewed posts with links to URL 1 and posts with links to URL 2.

Rows correspond to pairs of URLs. Each pair of URLs is represented by one row. Each column contains a value, or set of values, that aggregates platform log data for the specific pair of URLs over the study period.

URLs in this dataset consist of those:

- Classified as civic (defined as related to politics or social issues) and news by Meta internal classifiers
- Shared more than 100 times to Facebook by U.S. users during the study period
- Viewed by U.S. users in their Feed

Due to the large size of this dataset, the table only includes URLs that are part of the subset of nodes that represent the core of the network, as defined by all civic news URLs that meet the criteria above. This subset of the dataset was identified using the disparity filter algorithm selected by the academic collaborators. Additional technical details on the definition and implementation of this method can be found in the Appendix to González-Bailón et al (2023) "Asymmetric Ideological Segregation in Exposure to News on Facebook."

Active users

Aggregated platform activity data for purposes of this dataset includes that of Facebook users who were active on the platform between September 1, 2020 and February 1, 2021, whose predicted home country is the United States and stated age indicates they are 18 years or older.

Aggregation

Variables are computed separately for each pair of domains and week. The data is stored as separate files, organized by week. For each week, there may be one to five files, depending on the number of rows. The name of each file ends with the last date of the specific week for which it contains data, followed by “_partX,” where X is a numeral from 1 to 5, for example `_2020-09-07_part2.csv`.

Disclosures

Platform limitations

Coexposed audience count

Coexposed audience metrics are aggregated by pairs of URLs and calculated by adding the daily coexposed audience count for each pair of posts with links to the two URLs, for every day over the study period. User views of a URL, prior to aggregating them to pairs of domains, could not be fully deduplicated.

- If a user viewed two different posts with a link to the same URL in a single day, they are counted in the exposed audience count twice for that day.
- If a user viewed the same post with a link to the URL but on different days, they are counted in that post’s exposed audience count for both days.

Categories and classifiers

URLs categorized as misinformation

For the purposes of this study, when a URL has received two or more ratings of partial or complete falsity from independent fact-checkers, the strictest rating (**false**) is used in this dataset. If any rating is **false**, then the value for this variable will be **true** (URL categorized as misinformation).

URLs classified as related to COVID-19

Meta's COVID-19 classifier classifies individual posts and not the URLs shared in those posts. As a result, a URL's classification can vary depending on the post in which it was shared. For the purposes of this study, to address this issue, we computed the number of times that each possible predicted topic was associated with a URL. If more than 50% of those posts were classified as related to COVID-19, the URL was assigned the COVID-19 topic.

Daily Ideological Segregation of the Audience of Facebook Posts with Civic News Domains and URLs

About this dataset

Table	daily_ideological_segregation_audience_facebook_posts_with_civic_news_domains_and_urls
Level	Domain and URL
Data	Aggregated Facebook activity data for US adult monthly active users
Description	Data on the daily ideological segregation of the audience of posts with links to civic news
SOMAR Page	https://socialmediaarchive.org/record/39

Overview

This dataset contains domain- and URL-level metrics from aggregated Facebook activity data for US adult monthly active users.

These metrics measure the **ideological segregation index** and **favorability score** of the potential, exposed and engaged audience of posts with links to civic news domains and URLs daily from September 1, 2020, to February 1, 2021.

Rows correspond to a subset of domains or URLs that a user could have viewed in their Feed (because it was present in their inventory of eligible content), actually viewed in their Feed (exposure), or engaged with (likes, reactions, comments, reshares and clicks). Each column contains a value, or set of values, that aggregates platform log data for the subset of domains or URLs over a given date within the study period.

Domains and URLs in this dataset consist of those:

- Classified as civic (defined as related to politics or social issues) and news by Meta internal classifiers
- Shared more than 100 times to Facebook by U.S. users during the study period
- Potentially viewed in Feed, actually viewed in Feed, or engaged with in Feed.

Active users

Aggregated platform activity data for purposes of this dataset includes that of Facebook users who were active on the platform between September 1, 2020, and February 1, 2021, whose predicted home country is the United States and stated age indicates they are 18 years or older.

Data categorization

This dataset contains metrics that include the following platform data types:

- Ideological segregation index
- Favorability score
- Content attributes
- User attributes

[Go to the data types appendix](#) for descriptions of these categories.

Aggregation

Variables are computed using the following types of aggregations:

- Daily during the study period
 - Start: September 1, 2020
 - End: February 1, 2021
- Level
 - Domain: metrics calculated at the news source level
 - URL: metrics calculated at the news story and source level
- Post owner type: entity type that shared the post with the domain or URL
 - All: posts from users, Pages and groups
 - User: posts from users to their profile
 - Page: posts from Pages
 - Group: posts from users to groups
- Audience
 - Potential audience: users who potentially could have viewed a post because it was shared by one of their connections
 - Exposed audience: users who viewed a post
 - Engaged audience: users who engaged with a post
- Audience subset: the subset of users for which exposure is aggregated
 - Overall: all US adult users
 - Political interest
 - Top 1% engagement: US adult users in the top 1% in engagement with posts on Facebook classified as civic. Engagement includes comments, likes, reactions and reshares but excludes clicks in this dataset.

- Top 10% engagement: US adult users in the top 10% in engagement with posts on Facebook classified as civic. Engagement includes comments, likes, reactions and reshares but excludes clicks in this dataset.
 - Bottom engagement: US adult users in the bottom 90% in engagement with posts on Facebook classified as civic. Engagement includes comments, likes, reactions and reshares but excludes clicks in this dataset.
 - Top 1% of views: US adult users in the top 1% in views of posts on Facebook classified as civic.
 - Top 10% of views: US adult users in the top 10% in views of posts on Facebook classified as civic.
 - Bottom views: US adult users in the bottom 90% in views of posts on Facebook classified as civic.
- Input table:
 - Complete: all civic news domains and URLs shared by U.S. users during the study period
 - Filtered: only civic news domains and URLs shared more than 100 times by U.S. users during the study period
- Ideology threshold: specific threshold used to classify users as liberal or conservative, depending on their predicted political ideology
 - Main thresholds used in the analysis: 0.35 and 0.65
 - Binary threshold: 0.50
 - Stricter thresholds: 0.10 and 0.90

Disclosures

Platform limitations

Audience metrics

Exposed audience metrics report the sum of views.

Potential and engaged audience metrics are aggregated by URL or domain and calculated by adding the daily audience count for each post with a link to the URL or URLs within the domain, for every day over the study period.

- If a user potentially viewed or engaged with posts with links to two different URLs within the same domain in a single day, they are counted in the audience count twice for that day.

- If a user potentially viewed or engaged with two different posts with a link to the same domain or URL in a single day, they are counted in the audience count twice for that day.
- If a user potentially viewed or engaged with the same post with a link to the domain or URL but on different days, they are counted in that post's audience count for both days.

Users with potential exposures to posts by Pages that had been deleted by the time that the data was pseudonymized (90 days after post creation) are excluded from the potential audience metrics.

Categories and classifiers

URLs categorized as misinformation

For the purposes of this study, when a URL has received two or more ratings of partial or complete falsity from independent fact-checkers, the strictest rating (**false**) is used in this dataset. If any rating is **false**, then the value for this variable will be **true** (URL categorized as misinformation).

URLs classified as related to COVID-19

Meta's COVID-19 classifier classifies individual posts and not the URLs shared in those posts. As a result, a URL's classification can vary depending on the post in which it was shared. For the purposes of this study, to address this issue, we computed the number of times that each possible predicted topic was associated with a URL. If more than 50% of those posts were classified as related to COVID-19, the URL was assigned the COVID-19 topic.

Weekly Ideological Segregation of the Audience of Facebook Posts with Civic News Domains and URLs

About

Table	weekly_ideological_segregation_audience_facebook_posts_with_civic_news_domains_and_urls
Level	Domain and URL
Data	Aggregated Facebook activity data for US adult monthly active users
Description	Data on the weekly ideological segregation of the audience of posts with links to civic news
SOMAR Page	https://socialmediaarchive.org/record/40

Overview

This dataset contains domain and URL-level metrics from aggregated Facebook activity data for US adult monthly active users.

These metrics measure the **favorability score** of the potential, exposed and engaged audience of posts with links to civic news domains and URLs weekly from September 1, 2020, to February 1, 2021.

Rows correspond to domains or URLs whose favorability scores fall within the values defined by the **bin floor** and **bin ceiling** columns of the dataset in a specific week and were potentially viewed, actually viewed in a user's feed, or engaged with by U.S. users in their Feed. Each column contains a value, or set of values, that aggregates platform log data for the specific URL or domain weekly over the study period.

Domains and URLs in this dataset consist of those:

- Classified as civic (defined as related to politics or social issues) and news by Meta internal classifiers
- Shared more than 100 times to Facebook by U.S. users during the study period
- Potentially viewed, actually viewed, or engaged with by U.S. users in their Feed

Active users

Aggregated platform activity data for purposes of this dataset includes that of Facebook users who were active on the platform between September 1, 2020, and February 1, 2021, whose predicted home country is the United States and stated age indicates they are 18 years or older.

Data categorization

This dataset contains metrics that include the following platform data types:

- Favorability score
- Content attributes
- User attributes

[Go to the data types appendix](#) for descriptions of these categories.

Aggregation

Variables are computed using the following types of aggregations:

- During the study period
 - Start: September 1, 2020
 - End: February 1, 2021
- Weekly date stamps
 - Start: Monday September 7, 2020
 - End: Monday February 1, 2021
- Level:
 - Domain: metrics calculated at the news source level
 - URL: metrics calculated at the news story and source level

Disclosures

None

Ideological Alignment of Users' Facebook Networks

About

Table	ideological_alignment_users_facebook_networks
Level	US adult monthly active population
Data	Aggregated Facebook network data for US adult monthly active users
Description	Aggregated metrics for the proportion of Facebook friends, Pages, and groups that are categorized as like-minded or cross-cutting

SOMAR Page <https://socialmediaarchive.org/record/29>

Overview

This dataset contains aggregated metrics summarizing the ideological conformity of Facebook friends, Pages, and groups among US adult monthly active Facebook users. For each user, the proportion of Facebook friends, Pages they follow, and groups they belong to that are categorized as ideologically like-minded or cross-cutting was calculated. This table reports summary statistics about these proportions.

A Facebook friend, Page or group is defined as like-minded in relation to the user if:

- the user's ideology is above 0.50 (on a 0-1 ideology scale, where 0 is very liberal and 1 is very conservative) and the friend/Page/group's (audience) ideology is above 0.60, or
- the user's ideology is below 0.50 and the friend/Page/group's (audience) ideology is below 0.40

A Facebook friend, Page or group is defined as cross-cutting in relation to the user if:

- the user's ideology is above 0.50 and the friend/Page/group's (audience) ideology is below 0.40 or
- the user's ideology is below 0.50 and the friend/Page/group's (audience) ideology is above 0.60

This table consists of two columns. The first column indicates the name of a network metric, where each row corresponds to that metric (e.g., the proportion of Facebook friends who are classified as ideologically like-minded). The second column contains, for each metric, summary statistics about that metric for US adult monthly active Facebook users (namely, the mean

value, standard deviation, median, 2.5th percentile, and 97.5th percentile values for that metric, as well as the proportion of missing values). Values in the second column are stored as dictionaries with keys corresponding to the names of summary statistics (e.g., mean) and numeric values.

Active users

Aggregated platform activity data for purposes of this dataset includes that of Facebook users who logged into Facebook at least once in the 30 days preceding August 17, 2020, whose predicted home country is the United States and stated age indicates they are 18 years or older.

Data categorization

This dataset contains metrics that include the following platform data types:

- User connections
- User attributes

[Go to the data types appendix](#) for descriptions of these categories.

Aggregation

Variables are snapshot metrics logged as of September 23, 2020, for users who were monthly active users as of August 17, 2020.

Disclosures

None

Facebook User Attributes

About

Table facebook_user_attributes

Level US adult monthly active population

Data Aggregated Facebook attributes for US adult monthly active users

Description Aggregated user-level attribute metrics

SOMAR Page <https://socialmediaarchive.org/record/30>

Overview

This dataset contains descriptive statistics summarizing the user attributes of US adult monthly active Facebook users, including the predicted ideology; account age; count of the posts shared by a user's Facebook friends, Pages they follow, and groups they belong to; active days on the platform and Facebook friend count.

This table consists of two columns. The first column indicates the name of a network metric, where each row corresponds to that metric (e.g., predicted ideology). The second column contains, for each metric, summary statistics about that metric for US adult monthly active Facebook users (the mean value, standard deviation, median, 2.5th percentile, and 97.5th percentile values for that metric, as well as the proportion of missing values). Values in the second column are stored as dictionaries with keys corresponding to the names of summary statistics (e.g., mean) and numeric values.

Active users

Aggregated platform activity data for purposes of this dataset includes that of Facebook users who logged into Facebook at least once in the 30 days preceding August 17, 2020, whose predicted home country is the United States and stated age indicates they are 18 years or older.

Data categorization

This dataset contains metrics that summarize the following platform data types:

- User attributes
- User connections

[Go to the data types appendix](#) for descriptions of these categories.

Aggregation

Variables are snapshot metrics logged as of September 23, 2020, for users who were monthly active users as of August 17, 2020.

Disclosures

None

Exposure to and Engagement with Facebook Posts

About

Table	exposure_and_engagement_facebook_posts
Level	US adult monthly active population
Data	Aggregated Facebook interactions for US adult monthly active users
Description	Aggregated metrics about exposure to and interactions with like-minded and cross-cutting posts
SOMAR Page	https://socialmediaarchive.org/record/28

Overview

This dataset contains aggregated metrics summarizing US adult monthly active users' exposure to and engagement with posts on Facebook. Exposure and engagement metrics are broken down by predicted post topics (e.g., civic, news, religion and spirituality, etc.) and by like-minded or cross-cutting sources. Engagement metrics include post likes, reactions, comments, clicks and reshares.

A Facebook friend, Page or group is defined as like-minded in relation to the user if:

- the user's ideology is above 0.50 (on a 0-1 ideology scale, where 0 is very liberal and 1 is very conservative) and the friend/Page/group's (audience) ideology is above 0.60 or
- the user's ideology is below 0.50 and the friend/Page/group's (audience) ideology is below 0.40

A Facebook friend, Page or group is defined as cross-cutting in relation to the user if:

- the user's ideology is above 0.50 and the friend/Page/group's (audience) ideology is below 0.40 or
- the user's ideology is below 0.50 and the friend/Page/group's (audience) ideology is above 0.60

This table consists of two columns. The first column indicates the name of a network metric, where each row corresponds to that metric (e.g., the proportion of Facebook friends who are classified as ideologically like-minded). The second column indicates, for each metric, summary statistics about that metric for US adult monthly active Facebook users. Values in the second

column are stored as dictionaries with keys corresponding to the names of summary statistics (e.g., “mean”) and numeric values. For variables ending in “_histogram,” the second column is a dictionary where the key corresponds to a percentage bucket (0-1 in hundredth increments) and each value corresponds to the proportion of users in that bucket of exposure. For all other variables, the summary statistics consist of the mean value, standard deviation, median, 2.5th percentile, and 97.5th percentile values for that metric, as well as the proportion of missing values.

Active users

Aggregated platform activity data for purposes of this dataset includes that of Facebook users who logged into Facebook at least once in the 30 days preceding August 17, 2020, whose predicted home country is the United States and stated age indicates they are 18 years or older.

Data categorization

This dataset contains metrics that summarize the following platform data types:

- Content interactions
- User attributes
- User connections

[Go to the data types appendix](#) for descriptions of these categories.

Aggregation

Variables are computed for monthly active users as of August 17, 2020, using the following types of aggregations:

- Study period: duration of the Like-Minded Sources Demotion platform experiment.
 - Start: September 24, 2020
 - End: December 22, 2020
- Pre-study period: the time period prior to the Like-Minded Sources Demotion platform experiment.
 - Start: June 26, 2020
 - End: September 23, 2020

Disclosures

Data issues

We note the following known issues:

- **Clicks** (click_count*):
 - Accuracy measures were not run on the source table for variables using the **clicks** metric. While the accuracy of the data is not guaranteed, Meta has robust testing and monitoring in place to assure the usability of data.
- Due to logging limitations, reshares of user posts to groups cannot be distinguished from reshares of user posts on their profiles. Since the latter is more common, all reshares of user posts are added into the user key for metrics by connection type, regardless of which of the two types of user posts they were. As a result, the reshare metrics broken down by connection type do not contain a group key.

Filtering

Unless otherwise noted (i.e., for reshares as noted above), post views and engagement metrics only include data corresponding to posts viewed in Feed. Other surfaces on Facebook, such as groups or News, are excluded from all post view metrics.

Part 3: Data Dictionary

For each dataset, a data dictionary lists all variables included in that dataset. For each variable, it includes the variable name, description, type, group, map keys, aggregation methods and disclosures. The data dictionaries are included as separate files in the dataset's study page on ICPSR's Social Media Archive (SOMAR).

Appendix

Glossary of terms

A Glossary of terms is provided as a separate document available through the ICPSR SOMAR [U.S. 2020 Facebook and Instagram Election Study collection portal](#).

Appendix: Experimental design

Experimental studies

The experimental studies consisted of a series of platform interventions that changed consenting participants' experiences on Facebook or Instagram. Participants were assigned to a treatment or control group and were invited to participate in several survey waves. Initially, there were 5 survey waves planned, but following the January 6th United States Capitol attack, the research team added a sixth survey wave to study events including the presidential transition, the violence at the Capitol on January 6th, and the Presidential Inauguration.

The experimental studies consisted of two sets of experiments in which consenting participants were compensated:

- A series of [platform interventions](#) that changed participants' experiences on Facebook or Instagram
- A [deactivation experiment](#) that incentivized a subset of participants to deactivate their Facebook or Instagram accounts and measured the effects of deactivation; these data will be released after publication of the relevant paper.

Recruitment

Randomly selected users viewed a recruitment message at the top of their Instagram and/or Facebook Feed, asking them if they would like to share their opinion. Those clicking the Start Survey button were directed to an Institutional Review Board–approved consent form. The full text of all recruitment materials is [available in the Open Science Framework \(OSF\)](#).

Message shown to Facebook and Instagram users:



Platform interventions

The platform interventions started on September 24, 2020, and ended on December 23, 2020. The number of participants who were randomized into a study is rounded to the nearest 100 below. These counts exclude participants who withdrew from their study or deleted the Facebook or Instagram account connected to the study prior to delinking.

Control: Platform Interventions

Participants in the control group did not have their Facebook or Instagram experience changed in any way.

Number of participants:

Facebook: 19,600 Instagram: 15,300

Chronological Feed

A ranking test experiment that displayed all content on the participants' Facebook and Instagram Feed in chronologically reversed order.

Participants in treatment group:

Facebook: 8,700 Instagram: 10,800

Reshare Holdout

A platform experiment where content with posts reshared from other accounts was not displayed in participants' Facebook Feed.

Participants in treatment group:
Facebook: 8,800

Like-Minded Sources Demotion

A platform experiment that included demotion of posts in participants' Facebook Feed from sources that are ideologically aligned with the participants.

Participants in treatment group:
Facebook: 8,700

Appendix: Platform data types, classifiers and categories for purposes of these studies

Overview of platform data collected

The U.S. 2020 FIES uses a combination of aggregated platform activity data for US active users and research participants, and user-level platform activity data for research participants. Each study may use a different subset of the following data types, as detailed in the dataset descriptions. The definitions below may be specific to the U.S. 2020 FIES and may not be representative of how Meta defined similar metrics when the data was collected during the study period or defines or operationalizes similar metrics today.

Primary definitions

US user: an account owner whose predicted home country is the United States.

Account Center Accounts. Accounts Center account refers to either (1) an individual Facebook or Instagram account that has not been added to the same Accounts Center as another account, (2) the two or more Facebook and Instagram accounts that have been added to the same Accounts Center.

Adult user: a Facebook account owner whose **stated age** indicates they are 18 years of age or older, or an Instagram account owner whose **predicted age** indicates they are 18 years of age or older.

Active or monthly active user: Accounts (filtering out business accounts) that used Facebook or Instagram during a specific measurement period, such as within the previous 30 days (monthly). Usage refers to content exposure, engagement or creation.

Content: content published to Facebook or Instagram that the participant or user either created, viewed or engaged with. Unless otherwise noted, the scope of the data analyzed in this project includes only content created by Pages or US adult users or participants (when analyzing content production) or content viewed or engaged with by US adult users or participants (when analyzing content exposure and engagement).

Entity: a Facebook user, Page or group, or an Instagram user account.

Platform activity data types

Measures derived from different types of platform activity are included in replication datasets. Each study may use a different subset of platform data types. Examples of such data include:

- **Content interactions:** for organic content, includes views, engagement (likes, reactions, comments, reshares and clicks) and original content creation; for ads, includes impressions and reach.
- **Time spent:** the amount of time relative to a study-specific baseline a participant or user was actively engaging with Facebook or Instagram during, after or on a specific day of the study period.
- **Account connections:** A participants' or users' connections including friends and followers, Facebook Pages and Facebook groups, changes in those connections, and connection type of the source of content the participant or user engaged with, created or viewed.
- **User attributes:** can include age, gender, predicted ideology, predicted location, political interest, and activity measures.
- **Content and entity attributes:** attributes used to characterize the types of content and entities the participant or user engaged with, shared or viewed content from, such as civic content or civic Pages.

Each data type is further explained below.

Content interactions

Content interactions measure actions on content at the user or participant level. They can be broken down by user, content and entity attributes. There are many active and passive ways to engage with content on Facebook and/or Instagram. Not all of these were examined in the U.S. 2020 FIES. The replication datasets for the U.S. 2020 FIES include the following:

Content engagement: actions a participant or user initiated on organic content posted to Facebook or Instagram. Actions include likes, reactions, comments, reshares and clicks.

- **Comments:** the number of comments made on content posted to Facebook or Instagram, including replies to comments.
 - Comment: to leave a digital remark on a post, photo, video or story.
- **Reshares:** the number of times a piece of content was reshared. On Facebook, this includes any reshares of posts that participants viewed in their Feed or on other surfaces. On Instagram, this includes reshares of posts to a story.
 - Reshare: to share an item posted to Facebook or Instagram.
- **Likes:** the number of likes on content posted to Facebook or Instagram.
 - Like: to express a positive sentiment about a post, photo, chat message or video by selecting an option (an icon or the word “like”).
- **Facebook Reactions:** the number of reactions (other than likes, calculated separately in these datasets) made on posts to Facebook. Data is only available for this interaction on Facebook.
 - Reactions: an extension of the Like button on Facebook, that gives people more ways to express themselves and share their reaction to a post. The collection of Reactions (other than Like, represented by a thumbs-up emoticon) include Love, Haha, Wow, Care, Sad and Angry.
 - Love: represented by a heart emoticon.
 - Haha: represented by a laughing emoticon.
 - Wow: represented by a surprised emoticon.
 - Care: represented by an emoticon hugging a heart.
 - Sad: represented by a sad emoticon.
 - Angry: represented by a mad emoticon.
- **Clicks:** the number of times content was clicked or tapped on Facebook. Data is only available for this interaction on Facebook. Includes clicks on external URLs, to “See More” and expand the text of a post, on the post author’s name, to enlarge a picture, to reshare, as well as to add a like, reaction, comment or any other type of engagement with the post.

Content shared: the individual pieces of content the participant or user created on Facebook or Instagram.

- Share: “To share” refers to the act of publishing information via a platform (a feed, for example) to a specific audience

Views: the number of times content appeared on a participant or user’s screen on Facebook or Instagram. For purposes of this dataset, a view on Facebook is counted whenever the content renders in the visible portion of the web browser or mobile device for more than 250 milliseconds. For purposes of this dataset, a view on Instagram is counted whenever a single pixel of the relevant content is displayed for any period of time. Can be broken down by user, content and entity attributes.

Audience: for purposes of this study, the number of users or participants who viewed a piece of organic content at least once.

Impressions: specific to ads, it is the number of times an ad was displayed on screen.

Reach: specific to ads, it is the number of Accounts Center accounts that viewed an ad at least once. Reach is different from impressions, which may include multiple views of ads by the same Accounts Center accounts. This metric is estimated.

Time spent

For the purposes of this study, time spent is the amount of time a participant or user spent actively engaging with an applicable Meta platform relative to a study-specific baseline. Daily time spent metrics include time spent on Facebook and/or Instagram. Time spent metrics can be broken down by surfaces. Unless otherwise specified in the individual dataset, datasets use the following time-spent calculations.

Standardized time spent: This metric calculates the standardized difference in time spent compared to daily or monthly active users on the platform according to the following process:

1. It computes the difference in time spent between each participant or user and the average time spent by daily or monthly active users on that day, and divides it by the standard deviation of time spent by daily or monthly active users.
2. It aggregates each daily value into a single value for each participant or user by taking the daily average.

A value of -1 on this metric indicates that the participant or user’s average time spent is one standard deviation below the average time spent by daily or monthly active users; a value of 0 indicates they spent the same amount of time; and a value of 0.50, for example, indicates their time spent is half a standard deviation higher than daily or monthly active users.

Proportional time spent: This metric calculates the proportional difference in time spent compared to daily or monthly active users on the platform according to the following process:

1. It computes the proportional difference between each participant or user's time spent on a given day and the time spent by monthly active users on that day.
2. It averages these values over the entire study period for each participant or user.

A value of -1 on this metric indicates that the participant or user spent 100% less time on the platform compared to daily or monthly active users; a value of 0 indicates they spent the same amount of time; and a value of 0.50, for example, indicates they spent 50% more time than daily or monthly active users.

User connections

Includes user connections, such as friends and followers or Pages, groups and events they are connected to, and connection actions, such as following and unfollowing. This data can be used to analyze the number of connections a participant or user has and the connection type of the content they view or engage with on Facebook and Instagram.

Friends and followers

- **Facebook friends:** the accounts a participant or user is connected to on Facebook. To be someone's friend, one user sends a friend request that the other user must accept.
- **Facebook followers:** the accounts that are following a user's Facebook account or Facebook page.
- **Facebook follows:** the accounts a user is following on their Facebook account.
- **Instagram followers:** the accounts that are following a user's Instagram account.
- **Instagram follows:** the accounts that a user is following on their Instagram account.
- **Instagram mutual follows:** the accounts that a user is following on their Instagram account and that follow their account back.

Other Facebook connection types

- **Pages connections:** the Pages the participant or user likes, follows, unlikes, unfollows or is an admin of on Facebook.
- **Groups connections:** the groups the participant or user joins, leaves or is an admin of on Facebook.

- **Civic events connections:** the events classified as civic on Facebook that the participant or user created, was invited to or attended.

Connection actions

Changes in the accounts the participant or user is connected to on Facebook, including to follow, unfollow, block and unfriend.

- **Follow:** to opt in to seeing content or posts from a Feed, hashtag or social media account or profile.
- **Unfollow:** to opt out of seeing content from a social media account and receiving notifications about that content.
- **Block:** a record of one or more entities that is not allowed access to something and/or is not allowed to perform a certain action.
- **Unfriend:** to select to no longer be friends with someone on Facebook.

User attributes

Stated age and gender

Age derived from the date of birth of the user, as specified on the platform. Depending on the dataset, stated age is reported as Meta age brackets or survey panel age brackets.

- Meta age brackets:
 - 18-24
 - 25-34
 - 35-44
 - 45-54
 - 55-64
 - >65
- U.S. 2020 panel survey age brackets:
 - 18-29
 - 30-44
 - 45-65
 - >65
- Stated gender from survey responses or user profile:
 - Male
 - Female
 - Other

Predicted political ideology

The predicted political ideology of the user based on the [US ideology classifier](#).

Predicted user location

- For Facebook users, we predicted a primary location based on that user's information and activity on Facebook, including the city information reported on their Facebook profile as well as other device and connection information, if available.
- For Instagram users, we predicted a primary location based on that user's connection information such as IP address.

Political interest

For the purpose of this study, content views and engagement metrics in certain datasets are broken down by a participant's or user's estimated level of political interest.

In this study, political interest is determined based on engagement with Facebook posts classified by [Meta internal classifiers](#) as civic, defined as related to politics or social issues, during the 90 days prior to the study period.

High level of political interest:

- Participants or users whose engagement rate with content on Facebook classified as civic was in the top 1% or 10% of all US adult active Facebook users.
- Participants or users whose views of content on Facebook classified as civic were in the top 1% or 10% of all US adult active Facebook users.

Low level of political interest:

- Participants or users whose engagement rate with content on Facebook classified as civic was in the bottom 90% or 99% of all US adult active Facebook users.
- Participants or users whose views of content on Facebook classified as civic were in the bottom 90% or 99% of all US adult active Facebook users.

Content and entity attributes

Data on the characteristics of the entities and content the participant or user shared to, engaged with or viewed on Facebook and Instagram.

Surfaces

Where on Facebook and Instagram the content was viewed or shared. This data is used to break down content interaction metrics, such as engagement and views.

- Facebook surfaces include:
 - **Feed:** Main Facebook landing surface that displays posts from users, Pages and groups that the user is connected to, plus suggested posts and ads from businesses that may be interesting and relevant to the user.
 - **Pages:** A product that lets people, public figures, businesses, brands, organizations, and non-profits create a presence on Facebook and connect with the Facebook community.
 - **Groups feed:** A stream of content updates posted by members or followers of the Facebook Group communities a participant or user is part of. A Facebook Group is a virtual community where members can post, comment and message about a dedicated topic or shared interest.
 - **Group page:** A unique Facebook Group's page where participants or users can see content such as the group information and a stream of content updates posted by members or followers of that community.
 - **Other types:** consists of less common surfaces not included above such as events, Marketplace, and News tab.

- Instagram surfaces include:
 - **Feed:** Main Instagram landing surface that displays posts from accounts the user follows, plus suggested posts and ads from businesses.
 - **Story:** A feature that allows users to post photos and videos that are stitched together into one unit of content that is available for 24 hours (unless a user adds a story to their profile highlights). This surface is available at the top of the home tab, and sits right above the Instagram feed.
 - **Explore:** opens when a participant or user clicks on the search tab. It has a grid-like structure with images, videos and ads.
 - **Hashtag search:** opens when a participant or user enters a keyword in the search bar and selects hashtags.
 - **Profile:** opens when a participant or user clicks an account's username. Can view the account profile on this page.
 - **Other types:** consists of less common Instagram surfaces not included above.

Ideological segregation index

The ideological segregation index was created for the U.S. 2020 FIES. It provides a summary statistic of the study's information ecosystem. It measures the extent to which participants with a conservative or liberal predicted ideology score potentially viewed or users potentially viewed, viewed or engaged with the civic news content on Facebook.

This index was adapted from research on residential segregation¹. The equation used to measure segregation was initially developed in academic studies of residential segregation². For this study, instead of measuring segregation in where people live, we measure segregation in the news that participants may have viewed, did view, or engaged with on Facebook. The index is used to identify the similarities and differences in the content with links to civic news domains and URLs that conservative and liberal participants' or users' may have viewed, did view, or engaged with.

The index score approaches 1 as liberal and conservative participants or users are exposed to increasingly different civic news domains or URLs and 0 as their patterns of exposure to civic news domains or URLs increasingly align.

Favorability score index

The favorability score index was created specifically for the purpose of the U.S. 2020 FIES. It measures the political ideology of the audience of posts with links to URLs or domains classified as civic news. Depending on the dataset, the audience can refer to the potential audience, exposed audience or engaged audience. The index is computed as $(C+L)/(C-L)$, where C is the number of conservative users in the audience and L is the number of liberal users. The score is 1 if the audience included only conservative users, -1 if it included only liberals, and 0 if it included the same amount of liberals and conservatives.

Content and entity classifiers and categories

The studies use various types of data classifications and categorizations for content and entities, developed internally by Meta, adapted from external research or originally co-developed with the U.S. 2020 academic collaborators.

Internal Meta classifiers and categories include the following:

- Civic-related
- Topic

¹ Gentzkow, Matthew, and Jesse M. Shapiro. "Ideological segregation online and offline." *The Quarterly Journal of Economics* 126, no. 4 (2011): 1799–1839.

² White, Michael J. "Segregation and diversity measures in population distribution." *Population index* (1986): 198–221.

- News
- Local news
- COVID-19
- Predicted Ideology (US adult Facebook users)
- Entity Audience Ideology (Facebook Pages, Groups, and web domains)
- Violation categories
- Political figures
- Misinformation repeat offenders
- Political memes

Go to the [Meta classifiers and categorization methods appendix](#) for a list and descriptions.

Classifications and categorizations adapted from external research or originally co-developed with the study's academic team include the following:

- Content with slur words
- Incivility
- Civic events
- Untrustworthy sources
- Misinformation and Misleading information

Go to the [external classifiers and categorization methods appendix](#) for a complete list and descriptions.

Aggregation levels

Aggregation of platform activity data

The following is a list of potential levels of aggregation for datasets released as part of U.S. 2020 FIES. Each dataset will contain data aggregated by one or several of the options below. See dataset descriptions for additional details.

- **Full user population:** statistics that aggregate over active users by month, as defined for purposes of this study.
- **Full participant sample:** statistics that aggregated over all or subsets of participants in the platform intervention experiments.
- **Participant-level:** statistics reported for each individual participant.
- **Breakdown by participant or user attributes:** statistics for all participants or users who share the same user attributes, as defined in the data types.
- **Full entity population:** statistics that aggregated over all units of the same content or entity type, including Instagram and Facebook posts, comments, ads, Facebook Pages and groups, URLs and domains.

- **Breakdown by content or entity attributes:** statistics for all content or entities that share the same attributes, as defined in the data types.
- **Temporal levels:**
 - During the pre-study period: activity aggregated in the three months prior to the platform experiments.
 - Study period: activity aggregated during the three months the platform experiments were active.
 - Daily: daily trends in views and engagement with different types of content.
 - Weekly: statistics aggregated on a weekly basis during the experiments.

Appendix: Meta classifiers and categorization methods

The classifiers, concepts and categorization methods used in the U.S. 2020 research studies that were either developed at Meta or are defined under Facebook and Instagram's platform policies.

Civic classifier

- **Definition:** This classifier predicts whether a given post is related to *politics* (government, elections, politicians, activism, etc.) or *social issues* (major issues that affect a large group of people, such as the economy, inequality, racism, education, immigration, human rights, the environment, etc.).
- **How it was Used**
 - We use the classifications for Facebook posts of any type (such as links, photos, videos, text) and Instagram posts (photos or videos) and comments that were created, viewed or engaged with by U.S. users during the U.S. 2020 FIES.
 - We use the classifications for both English- and Spanish-language content.
- **Related content categories:**
 - *Civic entities:* For the purposes of the U.S. 2020 FIES, we categorize entities (Pages, groups, domains) as civic or non-civic based on how much civic content they produce on the platform.
 - *Political memes:* For purposes of the U.S. 2020 FIES, we heuristically define political memes as images containing text categorized as civic.
- **Performance:**
 - Based on a sample of approximately 10k labeled posts, the classifier has 83% precision and 82% recall on English-language Facebook content.
 - Based on a sample of approximately 17k labeled posts, the classifier has 81% precision and 85% recall on Spanish-language Facebook content.
 - Based on a sample of approximately 51k labeled posts, the classifier has 94% precision and 78% recall on English-language Instagram content.
 - Based on a sample of approximately 74k labeled posts, the classifier has 80% precision and 60% recall on Spanish-language Instagram content.

COVID-19 classifier

- **Definition:** This classifier predicts whether a given post is related to COVID-19, i.e. health concepts related to the virus, policy or public health guidance or the economic impact of the virus, as well as people's reactions to the situation.
- **How it was Used:**

- We use the classifier predictions for Facebook posts (any post type: links, photos, videos, text) and Instagram posts (photos or videos) that were created, viewed or engaged with by U.S. users during the U.S. 2020 FIES.
- We use the predictions for both English- and Spanish-language content.
- **Performance:**
 - Based on a sample of approximately 30k labeled posts, 94% precision and 94% recall for English-language Facebook posts.
 - Based on a sample of approximately 31k labeled posts, 94% precision and 95% recall for Spanish-language Facebook posts.
 - Based on a sample of approximately 57k labeled posts, 86% precision and 80% recall on English-language Instagram posts.
 - Based on a sample of approximately 58k labeled posts, 88% precision and 80% recall for Spanish-language Instagram posts.

Topic classifier

- **Definition:** This classifier predicts content categories along 26 different broad topics:
 - Full list of topics:
 - *Animals & Pets*
 - *Books & Literature*
 - *Business, Finance & Economics*
 - *Crime & Tragedy*
 - *Education & Learning*
 - *Fashion & Style*
 - *Children & Parenting*
 - *Fitness & Workouts*
 - *Food & Drink*
 - *Games, Puzzles & Play*
 - *Health & Medical*
 - *History & Philosophy*
 - *Holidays & Celebrations*
 - *Home & Garden*
 - *Music & Audio*
 - *Performing Arts*
 - *Politics*
 - *Relationships, Friends & Family*
 - *Religion & Spirituality*
 - *Science & Tech*
 - *Social Issues*
 - *Sports*
 - *Travel & Leisure Activities*
 - *TV & Movies*

- *Vehicles & Transportation*
 - *Visual Arts, Architecture & Crafts*
- **How it was Used:**
 - We use the classifications for text, video, photo, and link posts that were present in Facebook Feed and were created, viewed or engaged with by US active users during the U.S. 2020 FIES.
 - We use the predictions for both English- and Spanish-language content.
 - Posts can be classified as multiple topics.
- **Performance:**
 - Classification thresholds were chosen to yield 80% precision.
 - The recall across all topics at this threshold is 89% in English and 86% in Spanish.
 - The recall for individual topics at this threshold ranges from 49% to 99%, with a median value of 87%.

News classifier

- **Definition:** This binary classifier predicts whether content is about current events, timely information, and follows journalistic standards such as citing sources and having a byline. See [News Content on Facebook | Meta Business Help Center](#) for additional information.
- **How it was Used:**
 - We use the classifier predictions for Facebook posts with a link or video that were created, viewed or engaged with by US active users during U.S. 2020 FIES.
 - We use the predictions for both English- and Spanish-language content.
- **Related content categories:**
 - *News x topic subclassification:* After the news classifier is run, we use the topic classifier to assign predicted news posts to one or more subcategories from the topic classifier list above.
 - *Civic news domains:* Web domains whose URLs are frequently categorized as civic as defined in the “Civic classifier” section above and as news by the news classifier relative to other content on the platform.
 - *Civic news:* Posts categorized as civic as defined in the “Civic classifier” section above and as news by the news classifier.
- **Performance:**
 - Classification thresholds were chosen to yield 80% recall.
 - Based on a sample of approximately 52k labeled links, precision at this threshold is 90% for US English-language news links.
 - Based on a sample of approximately 36k labeled links, precision at this threshold is 42% for US Spanish-language news links.

US predicted ideology classifier

- **Definition:** This classifier predicts US adult active Facebook users' political ideology.
- **Methodology:**
 - The classifier predicts the stated ideology of US adult monthly active Facebook users based on their demographics, preferred language, location, and engagement with content, Pages, and groups. It was trained with a data set that used self-reported ideology from Facebook profile fields. The self-reported ideology was removed from profile fields in December 2022. This is an internal classifier and used for research and analytic purposes only.
 - The classifier outputs a numeric score ranging from 0 (indicating a user is predicted to be liberal) to 1 (indicating a user is predicted to be conservative).
 - The classifier has high coverage—it is able to place up to 95% of US adult monthly active Facebook users on this numeric scale. The classifier also has high week-over-week stability—on average, the weekly scores for individual users have a correlation of 0.96.
 - To categorize US adult monthly active Facebook users into distinct ideological groups, the continuous ideology scores must be discretized. The thresholds used in each study may vary depending on the lead authors' preferences regarding precision and recall tradeoffs, and/or the need for consistency with the thresholds applied to Page and Group audience ideology scores described in the “Related entity categories” section below. In most cases, however, the user-level scores are discretized as follows:
 - **3-class ideology:** Users with a score less than or equal to 0.40 are categorized as *liberal*. Those with a score greater than or equal to 0.60 are categorized as *conservative*. The remaining users with a score between 0.40 and 0.60 are categorized as *moderate*.
 - **2-class ideology:** Users whose ideology score is less than or equal to 0.5 are classified as *liberal*, else *conservative*.
- **Related entity categories:**
 - **Entity audience ideology scores:**
 - We use the user-level ideology scores to generate similar ideology scores for the audiences of Pages, groups, and web domains. We do so by computing a measure of the ideological composition of their audience: the average predicted ideology of Facebook users engaging with these entities in the previous 28 days.
 - This approach places Pages, groups, and web domains on the same 0-1 numeric ideology scale as users. As such, we use the same thresholds for categorizing these entities into ideological groups as those specified above, with one exception: we use the term “mixed” rather than “moderate” for the middle category in 3-class ideology, as the scores reflect the ideological composition of entities' audiences and not the ideology of the entities themselves.
- **How it was Used:**

- We use the ideological classifications of US monthly active Facebook users 18 years or older who were active during U.S. 2020 FIES.
- We use the audience ideology classifications for all Pages, groups, and domains producing content that US active Facebook users viewed or interacted with during U.S. 2020 FIES.
- **Performance:**
 - We evaluated how well the user-level inferences performed at predicting self-reported ideology by comparing our classifications to the survey responses of U.S. 2020 FIES panelists. Precision and recall for each ideological group are as follows:
 - 2-class classification
 - Self-reported liberals: 85% precision and 86% recall
 - Self-reported conservatives: 83% precision and 83% recall
 - 3-class classification
 - Self-reported liberals: 62% precision and 77% recall
 - Self-reported moderates: 53% precision and 30% recall
 - Self-reported conservatives: 59% precision and 74% recall
 - We note that the reason precision is lower across all groups and recall is lowest for self-reported moderates in the three-class classification task is that the model tends to predict self-reported moderates to actually be left- or right-leaning. This is consistent with external research finding self-reported moderates often have non-centrist ideological positions on various issues (Broockman, 2016).
 - We also evaluated how well the audience ideology approach might approximate the ideological affiliation of the entities themselves, if one exists. We compared the audience ideology scores we computed for the official Facebook Pages of US Members of Congress to a widely used external measure of those Congress members' ideology derived from their legislative voting history, DW-NOMINATE (Lewis et al, 2023), finding a 0.96 correlation ($N = 409$).

Content and entity categories

Facebook and Instagram policy violation categories

Posts taken down due to the post or account violating one or more of the following Facebook and Instagram enforcement policies:

- [Violence & Incitement](#)
- [Misinformation](#) (including harmful health misinformation)
- [Dangerous Individuals & Organizations](#)
- [Coordinating Harm](#)
 - Voter suppression and fraud
 - Other types of Coordinating Harm

- [Bullying and Harassment](#)
- [Spam](#)
 - Engagement abuse
 - Content spam
 - Inflating distribution
- [Hate Speech](#)
- [Coordinated Inauthentic Behavior](#)

Political figures

The list of political figures includes elected officials and candidates for office at the federal, state, and local level, sourced from [Ballotpedia](#), the [Center for Tech and Civic Life \(CTCL\)](#) and other sources. Content produced by Facebook Pages and Instagram users in this list will be categorized as content from political figures.

Misinformation Repeat Offenders

A misinformation 'strike' for the Misinformation Repeat Offender (MRO) program as of late 2020 was incurred on posts containing at least one piece of content that was rated **false** or **altered** by one of the independent third-party fact-checking partners Meta works with through its fact-checking program.

In 2020, Facebook Pages and web (sub)domains with two or more strikes in a 90-day period were classified as MRO. Facebook groups with three or more misinformation strikes in a 90-day period were classified as MRO.

An Instagram account was classified as MRO in 2020 if it had created two or more posts that were rated **false** or **altered** by a third-party fact checker in the past 90 days. An account could also be classified as MRO if it created three or more posts that were matched using matching algorithms to content rated **false** or **altered**.

The independent third party fact-checking partners Meta works with through its fact-checking program can rate content as **false**, **altered**, **partly false**, or **missing context**. Posts containing text, links, images or videos can also be matched to such content using matching algorithms. See [Rating options for fact-checkers](#) for a detailed description of how these rating categories are defined.

Meta works with independent third-party fact-checking partners that are certified by the nonpartisan International Fact-Checking Network. A full list of these partners is available [here](#). In the US, this list includes organizations such as Reuters, FactCheck.org and PolitiFact. All fact-checks are publicly available on the websites of these organizations and can be reviewed by any external source for accuracy.

External classifiers and categorization methods

This section describes the classifiers and categorization methods used in the U.S. 2020 Facebook and Instagram Election Study (FIES) that were either proposed by the academic team or adapted from published academic research. In the case of methods proposed by or co-developed with the academic team, a thorough performance evaluation by the U.S. 2020 FIES academic and Meta researchers was not conducted due to time constraints in applying the methods to content or entities within Facebook’s and Instagram’s data retention periods. Instead, we relied on existing performance metrics when available (referenced below in regard to each classifier). In addition, when possible, the U.S. 2020 FIES academic and Meta researchers did hand-label a small set of examples, and rough performance estimates based on these samples are included below.

Classifiers

Content with slur words classification method

- **Definition:**
 - This classification method is adapted from Siegel et al. (2021), which identifies content containing at least one term sourced from [Hatebase](#) and the [Racial Slur Database](#), and attempts to reduce false positives introduced by the inclusion of terms with ambiguous meaning. Siegel et al.’s classification method labels content as falling within any of eight different categories (listed below), as well as a joint category (which we call “content with slur words”) that encompasses all of them.
 - anti-Asian
 - anti-Black
 - anti-Immigrant
 - anti-Muslim
 - anti-Semitic
 - anti-Latinx
 - homophobic
 - misogynistic
 - This classification method aims to capture content that could be perceived as hateful but need not violate Facebook’s and Instagram’s Community Standards or may not be captured by Meta’s existing automated systems.
- **Methodology:**
 - The method developed by Siegel et al. (2021) consists of a dictionary method that was augmented using machine learning methods in order to reduce the false positive rate. First, it identifies any post or comment that contains at least one term associated with each of the eight categories described above, based on a list compiled by and sourced from Hatebase and the Racial Slur Database.

Second, it applies a text-based classifier trained on a random sample of tweets to attempt to reduce false positives from terms that have multiple meanings, not all of which may be considered a slur. For instance, a benign use of "sneakers", as opposed to its usage as an anti-Black slur.

- We note two limitations of this method, both of which are common to slur-based approaches to hate speech classification. First, the classification method only detects an inherently limited set of slurs. A great deal of hateful content is more nuanced, subtle, and/or complex, making it hard to automatically classify in any instance, but especially hard to do so using the detection of slurs (Bianchi et al, 2022). Second, in selecting this classification method, we chose to err on the side of false positives (recall) as opposed to false negatives (precision). As a result, some of the content classified as containing one or more slur words may capture ingroup discourse and banter (e.g., use of the n-word among Black communities and use of "bitch" among women), neither of which the false positive filter is designed to filter out. We therefore describe it here as "content with slur words" (a deviation from our preregistration). In general, we acknowledge that a slur-based approach may miss important context and produce both false positives and false negatives (Vidgen et al, 2021).
- **How it was Used:**
 - We apply the externally trained classification method to generate predictions for Facebook posts and comments, as well as Instagram captions and comments, that were created, viewed or engaged with by U.S. users during the U.S. 2020 FIES.
 - This classification method was applied to English-language content only.
 - For posts, the classification method is applied to the text of the post, the text contained in attached images, and a transcription of any attached videos.
 - For comments, the classification method is applied to the text of the comment only.
 - For reshared posts, a post is classified as containing one or more slur words if the original post or the reshared post (including the text that was added to the reshared post) was classified as containing one or more slur words.
- **Performance:**
 - Siegel et al. (2021) estimate this categorization method has 94% accuracy, 95% precision, and 90% recall for content with a slur.
 - In our data, over 90% of content predicted to contain a slur word using this method falls under the 'misogynistic' subcategory.

Incivility classifier

- **Definition:**
 - The classifier aims to capture uncivil content, defined as: "Features of discussion that convey an unnecessarily disrespectful tone toward the discussion forum's participants or its topics, which is including but not limited to:

- (1) Name-calling, mean-spirited or disparaging words directed at a person or group of people.
 - (2) Aspersion, mean-spirited or disparaging words directed at an idea, plan, policy, or behavior.
 - (3) Vulgarity, using profanity or language that would not be considered proper in professional discourse.
 - (4) Pejorative for speech, disparaging remarks about the way in which a person communicates.”
 - This approach follows the definitions from Coe, Kenski, & Rains (2014). Note that uncivil language as defined here may not necessarily be threatening or harmful and could be used to emphasize opinions, for example (Rossini, 2022).
- **Methodology:**
 - The classifier is a regularized logistic regression, using unigrams as features, trained on two datasets:
 - A random sample of 5,000 Reddit comments, collected by Davidson et al. (2020), annotated by three undergraduate students (inter-coder agreement = 91%)
 - A random sample of 4,000 tweets, annotated by crowd workers (intercoder agreement > 80%), along with a synthetic set of 16,000 tweets labeled by the Google Perspective API, collected by Theocharis et al. (2020)
 - The training dataset was complemented with 5 million labels generated using DistillBERT to improve its performance.
 - We applied the externally trained classifier to the Facebook and Instagram content categories described above, using text-based features (posts, captions, and comments) and Optical Character Recognition, a process that extracts text from images (for posts only).
 - For reshared posts, a post is classified as hateful if the original post was classified as uncivil or if the reshared post was classified as uncivil.
- **How it was Used:**
 - We use the externally trained classifier to generate predictions for Facebook posts and comments, as well as Instagram captions and comments, that were created, viewed or engaged with by U.S. users during the U.S. 2020 FIES.
 - This classifier was applied to English-language content only.
- **Performance:**
 - Predictions from this model were estimated by Davidson et al. (2020) to have 85% precision and 72% recall on the Reddit dataset (computed on a test set not used for training) and 89% precision and 70% recall on the Twitter dataset.
 - Based on human annotations conducted by the Meta research team and validated by a subset of the academic researchers on the U.S. 2020 FIES comprising a random sample of 100 public Facebook posts, 100 public Facebook comments, and 100 public Instagram comments, we estimate that the classifiers have:
 - Facebook comments: 86% precision

- Facebook posts: 83% precision
 - Instagram comments: 75% precision
- Based on a set of synthetic examples available in Jones (2015) from African-American Vernacular English, compared to their Modern Standard English variation, we found that the classifier yields similar predictions, which we take as evidence of adequate calibration.

Civic events

- **Definition:** This classifier aims to capture Facebook events that have a title or description similar to the titles or descriptions of events created by political figures.
- **Methodology:**
 - This classifier uses Naive Bayes with bigrams, trigrams, and 4-grams from the event title or description as features.
 - The positive labels in the training set consist of any events with at least 100 attendees that were (1) created by any Pages associated with government officials, elected officeholders, or candidates for elected office, except those associated with “government service” such as police or firefighters, OR (2) have a title or description containing the words “debate,” “protest,” “biden,” “trump,” “election(s),” “congressman,” “congresswoman,” “senate,” or “senator.”
 - The negative labels in the training set consist of events with at least 100 attendees that were created by a “government service” entity **and do not** contain the words “debate,” “protest,” “biden,” “trump,” “election(s),” “congressman,” “congresswoman,” “senate,” or “senator.”
- **How it was Used:**
 - We generate predictions for all Facebook events.
 - Available for English language, but could potentially surface some events with titles and descriptions that also contain Spanish
- **Performance:** Based on human annotations for a set of 150 events across three days in September 2020, we estimate that the classifier has:
 - 76.7% precision
 - 82.3% recall

Non-classifier-based content or entity categorization methods

Misinformation and Misleading information

For the purposes of this study, “Misinformation” refers only to content that is directly rated **false** by one of the independent third-party fact-checking partners Meta works with through its fact-checking program, or posts containing text, links, images or videos that are matched to such content using matching algorithms. Content rated as **altered**, **partly false**, or **missing**

context is categorized as “Misleading information” rather than “Misinformation.” See [Rating options for fact-checkers](#) for a detailed description of how these rating categories are defined.

Meta works with independent third-party fact-checking partners that are certified by the nonpartisan International Fact-Checking Network. A full list of these partners is [available here](#). In the US, this list includes organizations such as Reuters, FactCheck.org and PolitiFact. All fact-checks are publicly available on the websites of these organizations and can be reviewed by any external source for accuracy.

Untrustworthy Sources

Definition: For purposes of this research project, untrustworthy sources were defined in the pre-analysis plans based on the number of misinformation strikes they accrued under Meta’s Misinformation Repeat Offender Policy. In 2020, a strike may have been counted when an entity produced content rated false or altered by one of the independent third-party fact-checking partners Meta works with through its fact-checking program. For the Facebook component of this study, untrustworthy sources were defined to encompass Pages, groups, and domains that had accrued two or more misinformation strikes since the MRO program began on Facebook in 2018 (i.e., “lifetime” strikes). For the Instagram component of the study, untrustworthy sources were defined to encompass any public account with two or more strikes from content directly rated false or altered, or three or more strikes from content either directly rated false or altered or matching content rated false or altered since the MRO program began on Instagram in 2020. Note the untrustworthy sources category was defined and used solely for the U.S. 2020 FIES and is not the same as how Meta defined misinformation Repeat Offenders in 2020 nor how [Meta defines them today](#).

Operationalization: Due to data limitations encountered in the execution of this research project, we can only estimate the number of strikes accrued by an entity using third-party fact-checker ratings. Pages or groups that have posted content that received a **false** or **altered** rating by one of the independent third-party fact-checking partners in Meta’s fact-checking program at least twice, or domains with two or more URLs rated as **false** by a third-party fact-checker are considered untrustworthy sources. We exclude from this list of domains other social media platforms, hosting sites, and URL shorteners, even if they may host URLs that have been fact-checked as false.

The study uses a more expansive operationalization of an untrustworthy source than was originally pre-registered, and will necessarily overestimate participant exposure to these sources. This is because in practice, third-party fact-checker ratings are an imperfect measure of misinformation strikes. Specifically, the number of individual pieces of content that are fact-checked as false or altered does not necessarily equal the number of strikes an entity accrues under Meta’s Misinformation Repeat Offender Policy (see [Meta’s Transparency Center](#) for more details).

Appendix: Paper-to-dataset mapping

List of studies and associated tables

Paper title	Associated tables
<p>How do Social Media Feed Algorithms Affect Attitudes and Behavior in an Election Campaign?</p>	<p>Facebook Feed Intervention Experiment Participants</p> <p>Instagram Intervention Experiment Participants</p> <p>Metadata of Domains Shared on Facebook</p>
<p>Reshares on Social Media Amplify Political News but do not Detectably Affect Beliefs or Opinions</p>	<p>Facebook Feed Intervention Experiment Participants</p> <p>Metadata of Domains Shared on Facebook</p>
<p>Like-minded Sources on Facebook are Prevalent but not Polarizing</p>	<p>Facebook Feed Intervention Experiment Participants</p> <p>Instagram Intervention Experiment Participants</p> <p>Metadata of Domains Shared on Facebook</p> <p>Exposure to and Engagement with Facebook Posts</p> <p>Ideological Alignment of Users' Facebook Networks</p> <p>Facebook User Attributes</p>
<p>Asymmetric Ideological Segregation in Exposure to Political News on Facebook</p>	<p>Passive-Tracking Participants' Daily Views of Facebook Posts with Civic News Domains</p> <p>Passive-Tracking Participants' Views of Facebook Posts with Civic News Domains</p> <p>Exposure to Facebook Posts with Civic News Domains</p> <p>Exposure to Facebook Posts with Civic News URLs</p> <p>Engagement with Facebook Posts with Civic News Domains</p> <p>Engagement with Facebook Posts with Civic News URLs</p> <p>Potential exposure to Facebook Posts with Civic News Domains</p> <p>Potential exposure to Facebook Posts with Civic News URLs</p> <p>Coexposure to Facebook Posts with Civic News Domains</p> <p>Coexposure to Facebook Posts with Civic News URLs</p> <p>Daily Ideological Segregation of the Audience of Facebook Posts with Civic News Domains and URLs</p> <p>Weekly Ideological Segregation of the Audience of Facebook Posts with Civic News Domains and URLs</p>