Ad Library API Codebook

Erika Franklin Fowler, Mike Franz, Gary King, Greg Martin,
Zagreb Mukerjee, Nate Persily

Version 1.0
25 April 2019

This Ad Library API Codebook was created by via a collaboration between Facebook and Social Science One. The Ad Library API provides programmatic access to the Facebook Ad Library, a collection of all political advertisements run on Facebook and Instagram since May 2018 in the US, and for other dates in different countries. The codebook describes the scope, structure, and fields of these data.

All questions about this codebook, the API, and access, for those who do not yet have access, should be directed to help@socialscience.one. Approved researchers will receive an *access key*. For those who have already received access, direct questions to the *Social Science One community* Slack workspace.

Known issues with using the Ad Library for research can be found here. Social Science One and its counterparts at Facebook are working to keep these issues on the roadmap of the Ad Library team.

Requirements

All publications from these data must include this citation: Franklin Fowler, Erika; Franz, Mike; King, Gary; Martin, Greg; Mukerjee, Zagreb; Persily, Nate, 2019, "Facebook Ad Library", https://doi.org/10.7910/DVN/9OAMBW, Harvard Dataverse, V1.

Unit of Analysis

The unit of analysis is the "ad buy" — that a unique instance of an advertiser purchasing political advertising. One ad buy can differ from another in any of the following: page, funding entity, spending, targeting, a call to action button (if applicable) or creative ad content (image or video, etc), all of which are defined below.

Scope and Structure

The library includes all political ads run on Facebook and Instagram in the US, UK, India, and Brazil. The starting dates are as follows:

1. US - May 2018
2. Brazil - August 2018
3. UK - October 2018, and
4. India - February 2019.

A *political ad* is an ad that:

1. Is made by, on behalf of, or about a current or former candidate for public office, a political party, a political action committee, or advocates for the outcome of an election to public office; or
2. Relates to any election, referendum, or ballot initiative, including "get out the vote" or election information campaigns; or
3. Relates to any "national legislative issue of public importance" in any place where the ad is being run; or
4. Is legally regulated as political advertising.

"National legislative issues of public importance" are determined by Facebook in partnership with the non-partisan Comparative Agendas Project - for more information, see here. A list of these issues for the US can be found here. In other countries, there is no corresponding criterion. The relevant Facebook documentation is here.

The determination of which advertisements are "political" is carried out by Facebook through a mix of measures, relying on advertiser self-classification supplemented by human and algorithmic review. Facebook provides further information and a list of basic examples of political and non-political advertising here.

Advertisements are only deleted from the library if they are deemed, through the Facebook manual and algorithmic review process, to be apolitical. Terms of Service violations or page deletions will not cause ads to be removed from the archive - the ads will be taken down from Facebook, but this is not noted in the archive.

Ad Purchasing

1. Ads are purchased by an ad buyer, who must be an administrator for a Facebook Page (a Facebook website for a business, group, public person, topic etc). The page buyer must verify their identity with a document such as a US driver's license (the country-specific criteria are here), and answer a number of identity verification questions.

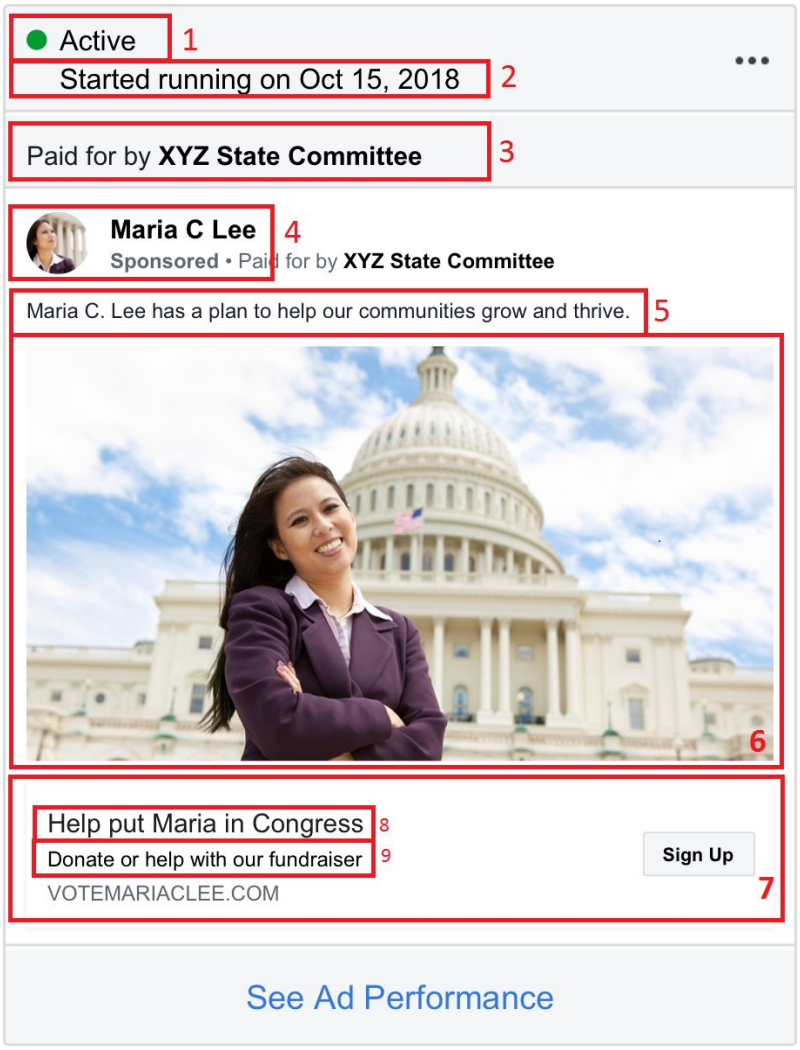
2. The buyer must also set up a "disclaimer", a statement of the form "Paid for by <X>", which accurately reflects the funding entity.
3. The actual advertisement can take a number of forms, most commonly a single image/video accompanied by text.
4. The ad can have a link, which can be a regular URL or have a "call to action", like "Sign Up" or "Donate Now"
5. The buyer can then select who they'd like to target with the advertisement. This targeting can be extremely granular on a variety of demographic, behavioral, and geographic dimensions. This targeting information is understandably of considerable interest to the advertisers competing with each other, and so is held closely. Thus, although it is within the scope of the Facebook-Social Science One collaboration and will eventually be made available to researchers, we presently do not make this detailed targeting information available in the API. We are looking into ways of making these data available to researchers in other ways.
6. The buyer sets a daily budget, which is treated as a guideline in the auction process below. The buyer's maximum budget is a hard cap on how much is spent. An advertiser can be paid by impression or link click.
7. The ad is then shown to users who match the targeting criteria, inside the newsfeed (desktop & mobile). (Political ads do not appear in the right column next to the Facebook interface -- the right column next to the newsfeed on the desktop).
8. Which ad is shown where, and when, is subject to an auctioning process. In summary, different Facebook ad buyers "bid" for space that is being shown to someone, with a winner selected based on budget and relevance to the user. More detail on this auction process can be found at the Facebook Help Center.
9. The ad purchaser can specify when the ad will stop being shown, or run it until a spending cap is hit.

For information about running ads on Facebook, please visit the following resources:

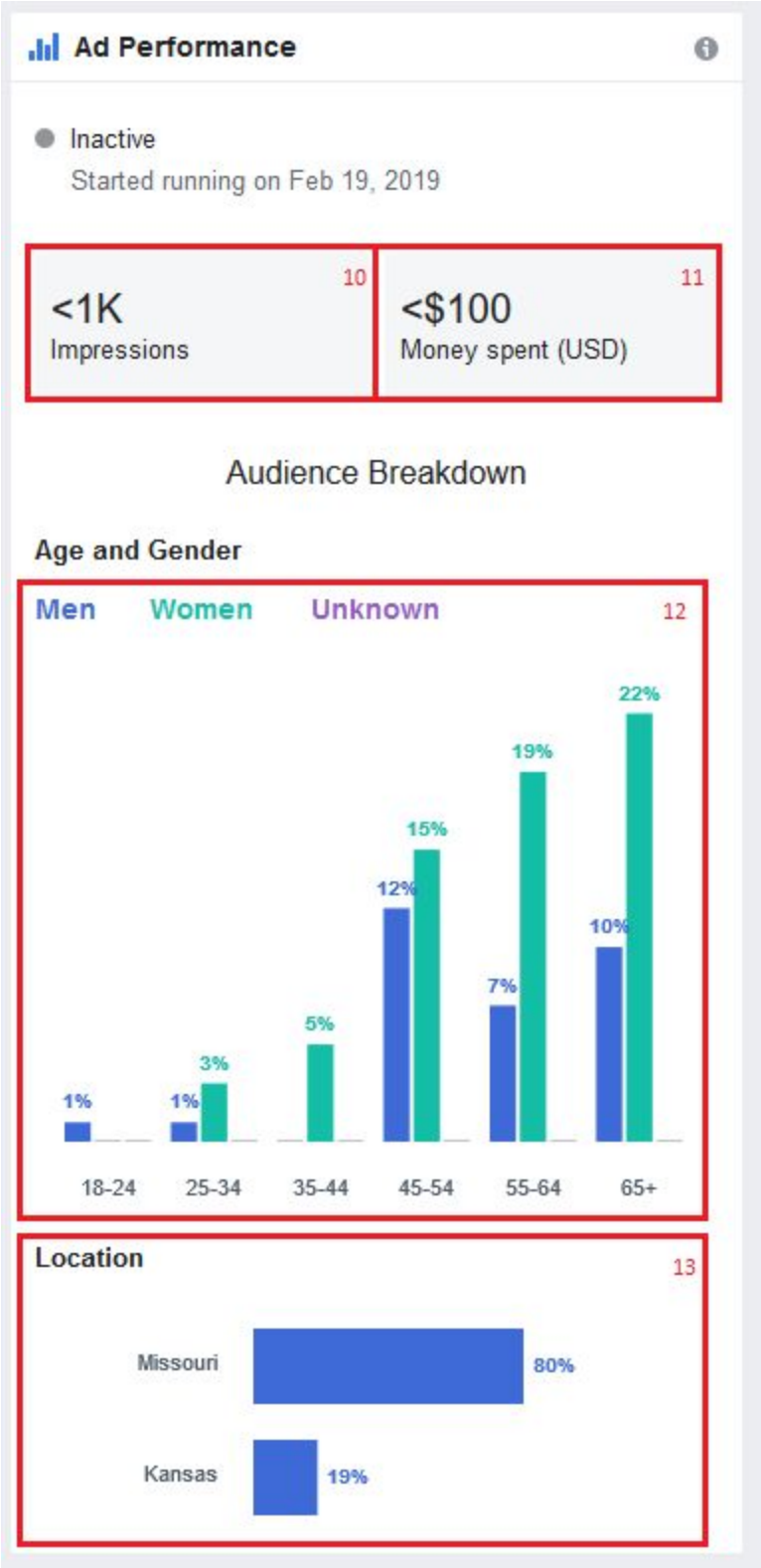
- Beginner's Guide for Advertisers can be found here
- Facebook Business: A collection of resources to help marketers learn about the Facebook advertising platform: https://www.facebook.com/business
- The advertiser Help Center can be found here, and contains more information about what it means to run an advertisement, where that advertisement goes, and what options are available to customize it.
- To view a course on running ads related to politics and issues, please visit here: https://our.internmc.facebook.com/intern/gmshub/content/?id=500957807044087

Variables

Consider an example advertisement from the web interface for the Ad Archive, before going into the particular pieces. We added the red boxes, with numbers corresponding to annotations below.



1. Active/Inactive status
2. Start Date: ad_delivery_start_time
3. Disclaimer/Funding: funding_entity
4. The page running ad: page_name
5. Ad creative body: ad_creative_body
6. Image/Video
7. Link & “Call to action”
8. Link caption: ad_creative_link_caption
9. Link description: ad_creative_link_description



10. Impressions Range: impressions
11. Spend Range: spend
12. Demographic breakdown: demographic_distribution
13. Location breakdown: region_distribution

Ad Content

ad_creation_time	The UTC date and time the ad was created (not when it was first run), with date and time separated by “T”. Example: 2019-01-24T19:02:04+0000 (+0000 is UTC offset)
ad_creative_body	The text accompanying the advertisement. Character limit is large and varies by ad format.
ad_creative_link_caption string	A string with the URL accompanying the ad - this is optional. This is referenced in the example advertisement above as part 8.
ad_creative_link_description	A string with the description text for link, if any. This is referenced in the example advertisement above as part 9.
ad_creative_link_title	A string with the link’s title, if any. This is referenced in the example advertisement above as part 7.
ad_delivery_start_time	The date and time when the ad started to deliver, in the same format as ad_creation_time. Chosen by the purchaser.
ad_delivery_stop_time	The time when purchaser wants to stop the campaign. This field can be left blank, which means an ad will run until stopped or until the budget runs out. In the same format as ad_creation_time.
ad_snapshot_url	A string with the URL link to the ad renderer that shows a snapshot of the ad, with images and videos uncompressed. This snapshot is saved by Facebook. While there is no bulk download feature available currently, users are able to download individual creatives

	for analysis, as long as they comply with the data storage terms outlined in the Facebook Terms of Service

Ad Performance

Each of these fields reflects the realized outcome of the advertisement - eg. how many impressions were actually received (within a range). Targeting information is not available.

demographic_distribution	The demographic distribution of the audience that the ad reached. This returns, in the JSON object (see How To Query, below), a list of tuples. Each tuple consists of (Age range, gender, percentage). For each combination of Age Range x Gender, this shows the percentage of impressions in that group. Age ranges: 18-24, 25-34, 35-44, 45-54, 55-64, 65+ Gender: “Male”, “Female”, “Unknown” (Unknown meaning the gender is not known to Facebook)
region_distribution	The region distribution of the audience that the ad reached. Tuples of (region, percentage), in the same format as demographic information. Regions are province/state level, depending on countries.
spend	A string showing the estimated total amount of money the ad cost to run. In units of currency (below). This is reported in ranges; <100, 100-499, 500-999, 1K-5K, 5K-10K, 10K- 50K, 50K-100K, 100K-200K, 200K-500K, >1M

currency	A string containing the currency that was used by the advertiser to pay for the ad.
impressions	A string containing the number of times the ad was shown on screen. This is reported in ranges: <1000, 1K-5K, 5K-10K, 10K-50K, 50K-100K, 100K-200K, 200K-500K, >1M

Ad Purchaser Information

funding_entity	A string containing a name of a person or entity funding the ad, submitted by the ad purchaser. More information can be found here and here.
page_id	A numeric string containing the ID of the Facebook Page which ran the ad. To navigate to the page with the respective page id a user can input the page id into www.facebook.com/[page_id]
page_name	A string containing the name of the Facebook Page which ran the ad. This string contains the name as it was at the time of ad purchase - later name changes do not propagate.

How To Query

Reading from the API will return archived ads-based keyword searches on text, image, audio from video and the call to action button. The API will also return archived ads for a given Facebook Page ID using the the following search parameter: search_page_ids= where one can enter up to 10 page IDs. Please note that keyword searches are not translated (eg searching Spanish language ads requires a Spanish language search string).

Examples

To get archived ads related to politics or issues of national importance that contain the word 'california' and reached audience in the US:

cURL

```
curl -G \-d "search_terms='california'" \-d "ad_type=POLITICAL_AND_ISSUE_ADS"
\-d    "ad_reached_countries=['US']"    \-d    "access_token=<ACCESS_TOKEN>"
\"https://graph.facebook.com/<API_VERSION>/ads_archive"
```

Response (this is a JSON object):

```
{
 "data": [
   {
     "page_id": "123",
     "page_name": "123",
                                                                  "ad_snapshot_url":
"https://www.facebook.com/ads/archive/render_ad/?id=123&access_token=<ACCESS_
TOKEN>"
   },
 ],
 "paging": {
   "cursors": {
     "before": "MAZDZD",
     "after": "MAZDZD"
   },
                                                                        "next":
"https://graph.facebook.com/v3.1/ads_archive?access_token=<ACCESS_TOKEN>&fiel
ds=page_id,page_name,ad_snapshot_url&search_terms='california'&ad_type=POLITI
CAL_AND_ISSUE_ADS&ad_reached_countries=['US']&limit=25&after=MAZDZD"
 }
}
```

Parameters

Parameter	Description
search_terms string	The search term of an ad archive search request. White space is treated as an "AND."
search_page_ids array<page ID>	Searches can be restricted to specific Facebook Page IDs (up to 10 page IDs separated by commas can be requested).
ad_active_status enum {ALL, ACTIVE, INACTIVE}	The current status of ads being searched. It can be active, inactive, or all.
ad_type enum {ALL, POLITICAL_AND_ISSUE_ADS, NEWS_ADS}	The type of ad. It can be all ads, political and issue ads, or ads for news related to politics or issues.
ad_reached_countries array<enum {US}>	The list of countries that the ad reached.
fields array, type any of those specified here.	List of data fields to include in the result. See "DATA" section below for a list of possible fields.

paging

For more details about pagination, see the Graph API guide.

Traversing Paged Results

When you make an API request to an object or relationship, you usually don't receive all of the results of that request in a single response. This is because some responses could contain thousands of objects so most responses are paginated by default.

Cursor-based Pagination

Cursor-based pagination is the most efficient method of paging and is the only one currently supported for the ad archive API. A cursor refers to a random string of characters which marks a specific item in a list of data. Unless this item is deleted, the cursor will always point to the same part of the list, but will be invalidated if an item is removed. Therefore, your app shouldn't store cursors and assume that they will be valid in the future.
When reading an edge that supports cursor pagination, you will see the following JSON response:
{"data": [... Endpoint data is here
],"paging": {"cursors": {"after": "MTAxNTExOTQ1MjAwNzI5NDE=","before": "NDMyNzQyODI3OTQw"},"previous":
"https://graph.facebook.com/me/albums?limit=25&before=NDMyNzQyODI3OTQw""next":
"https://graph.facebook.com/me/albums?limit=25&after=MTAxNTExOTQ1MjAwNzI5NDE="}}
A cursor-paginated edge supports the following parameters:

- before : This is the cursor that points to the start of the page of data that has been returned.
- after : This is the cursor that points to the end of the page of data that has been returned.
- limit : This is the maximum number of objects that *may* be returned. A query may return fewer than the value of limit due to filtering. Do not depend on the number of results being fewer than the limit value to indicate your query reached the end of the list of data, use the absence of next instead as described below. For example, if you set limit to 10 and 9 results are returned, there may be more data available, but one item was removed due to privacy filtering. Some edges may also have a maximum on the limit value for performance reasons. In all cases, the API returns the correct pagination links.
- next : The Graph API endpoint that will return the next page of data. If not included, this is the last page of data. Due to how pagination works with visibility, it is possible that a page may be empty but contain a 'next' paging link. Stop paging when the 'next' link no longer appears.
- previous : The Graph API endpoint that will return the previous page of data. If not included, this is the first page of data.

Don't store cursors. Cursors can quickly become invalid if items are added or deleted.

Facebook information sources:

Sources on Facebook information can be found here:
https://www.facebook.com/business/help
https://www.facebook.com/policies/ads
https://developers.facebook.com/docs/graph-api/reference/ads_archive/
https://developers.facebook.com/docs/graph-api/reference/archived-ad/
https://developers.facebook.com/docs/facebook-login/access-tokens
https://developers.facebook.com/docs/graph-api/using-graph-api
https://developers.facebook.com/docs/graph-api/reference/insights-range-value/