COPE-ID Database Documentation

The COPE-ID Database is made up of data mined from 10 social media platforms. Each data record is a post or comment about COVID-19. There are varying amounts of records collected from each social media platform and different identifying data elements per platform. Below is a list of the 10 sites COVID-19 related data was mined from along with the number of unique posts collected from the site.

- **4chan** – 15,039
- **8kun** – 709
- **Flickr**^ – 7,027**
- **Gab** – 171,288
- **Mastodon**^ – 77,054

- **Parler** – 80,742
- **Reddit** – 11,209,463
- **Tumblr**^ – 82,690
- **Twitter**^ – 3,054,857
- **YouTube**^ – 271,550

** Flickr's (photo-based platform) data file contains 24,931 unique photo ids, with 13,046 unique photo titles, 5,504 unique photo descriptions, and 7,027 unique photo tags.

To pull data from one of the given platforms an API key for the platform may be needed. The platforms that require an accepted application for an API key are marked with a caret^. To find a specific post, the unique identifier for the post/comment record and an API key (if applicable) are required. As an example, to query a batch of Twitter posts an API key is required. To find a specific Twitter post both the tweetId and API key are required.

Below are tables of each platform's data elements.

*Unique Identifier

4chan	
Element title	**Description**
commentNo*	Unique integer Id for the comment
inReplyTo	commentNo record that the given record is replying to, 0 if not a reply
createdAt	Timestamp for when comment was created

8kun	
Element title	**Description**
commentNo*	Unique integer Id for the comment
inReplyTo	commentNo record that the given record is replying to, 0 if not a reply
createdAt	Timestamp for when comment was created
commentURL	URL for the given record

Flickr

Element title	Description
photoId*	Unique integer Id for the photo
photoURL	URL for the given record

Gab

Element title	Description
postId*	Unique integer Id for the post
postURL	URL for the given record
inReplyTo	postId record that the given record is replying to, 0 if not a reply
createdAt	Timestamp for when post was created
replyCount	Integer count of reply posts to the record
reblogCount	Integer count of reblog posts to the record
favoriteCount	Integer count of favorites given to the record
userId	Unique integer Id for the user that created the post

Mastodon

Element title	Description
postId*	Unique integer Id for the post
postURL	URL for the given record
inReplyTo	postId record that the given record is replying to, 0 if not a reply
createdAt	Timestamp for when post was created
replyCount	Integer count of reply posts to the record
reblogCount	Integer count of reblog posts to the record
favoriteCount	Integer count of favorites given to the record
userId	Unique integer Id for the user that created the post

Parler

Element title	Description
postId*	Unique varchar Id for the post
parentPostId	postId record that the given record is replying to, 0 if not a reply
createdAt	Timestamp for when post was created
creator	Unique varchar Id for the user that created the post
numImpressions	Integer count of impressions for the record
numComments	Integer count of comment posts to the record
numUpvotes	Integer count of upvotes given the record
numReposts	Integer count of reposts to the record
postURL	URL for the given record
isArticle	Boolean value for whether post is an article
isSensitive	Boolean value for whether post contains sensitive content

Reddit

Element title	Description
CommentID*	Unique varchar Id for the comment
Time Created (in epoch)	Integer count of seconds since epoch that post was created
Score	Integer sum of upvotes and downvotes given to comment
SubredditID	Unique varchar Id for the subreddit the comment was posted under
LinkID	Unique varchar Id linking to comment
ParentID	postID/CommentID record that the given comment record is replying to

Tumblr

Element title	Description
postType	Text for type of media content
createdAt	Date-Time for when post was created
postURL	URL for the given record
noteCount	Integer count of all reblogs, likes, and replies given to the record
blogId*	Unique varchar Id for the post

Twitter	
Element title	**Description**
tweetId*	Unique integer Id for the post record

YouTube	
Element title	**Description**
commentId*	Unique varchar Id for the comment
commentThreadId	Unique varchar Id for the thread the comment belongs to
publishedAt	Timestamp for when the comment was posted
likeCount	Integer count of all likes given to the record