Generalized Data format (Raw)

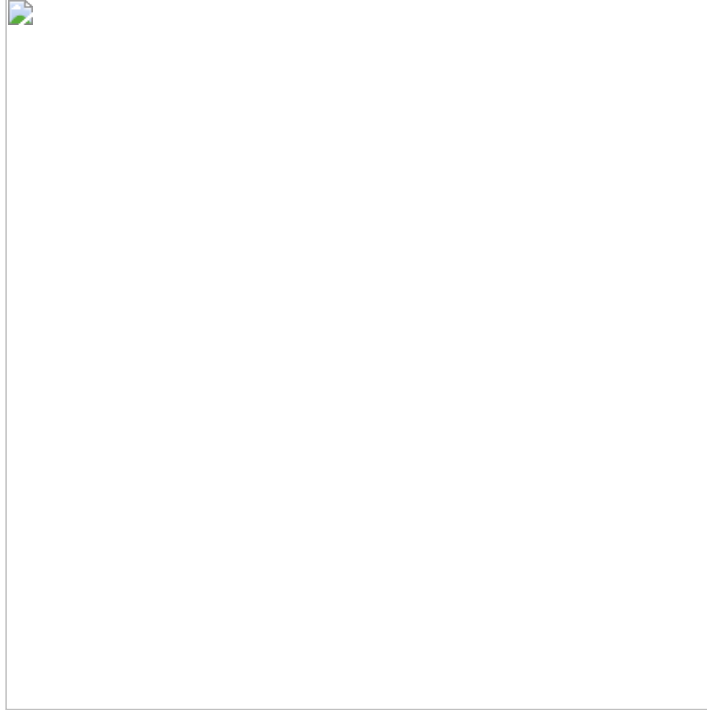| ⊙ Created	@January 12, 2023 2:10 PM
☰ Tags	

> 💫 format of data to collect from all platforms

▼ **MetaData for Tiktok:**

Tik_Tok Comments Keys:

Keys	Description
`cid`	The comment id, specified by the Tik_tok API
`text`	the text of the comment
`aweme_id`	No idea ?
`create_time`	the time the comment is created
`digg_count`	the likes count
`status`	??
`user`	The author of the comment
`reply_id`	??
`user_digged`	??
`reply_comment`	The replies of the comment
`text_extra`	??
`reply_comment_total`	The count of replies
`reply_to_reply_id`	??
`is_author_digged`	liked by the author
`stick_position`	??
`user_buried`	??
`label_list`	??
`author_pin`	pinned by the author
`no_show`	??
`collect_stat`	??
`trans_btn_style`	to adjust how the way the comment looks like
`share_info`	Sharing the comment metadata
`comment_language`	the language of the commen

A screenshot of output:

Tik_Tok User Keys:

Keys	Description
unique_id	The username of the account
id	Id used in the Tiktok system
nickname	The name of the user
sec_uid	Second id
private_account	bool var to check if the account private or not
verified	bool var to check if the account verfied or not
stats	UserStats {follower_count, following_count, heart_count, video_count, digg_count}
videos	The videos object

A screenshot of output:

```
unique_id='ii94n' id=6885445915738604549 nickname='طارق الغمري' sec_uid='MS4wLjABAAAArD6QGfaUUs2HucBPN7SEuknXezIj6rh8Jg4ceFrpJiWvofeeslD0aMynzcWRwWgL'
 private_account=False verified=False stats=UserStats(follower_count=683900, following_count=14, heart_count=8400000, video_count=92, digg_count=0)
 videos=<tiktokapipy.async_api.AsyncLightVideoIter object at 0x110de1430>
```

Tik_Tok Video Keys:

Keys	Description
ID	The id of the video
Stats	VideoStats: {digg_count, share_count, comment_count, play_count}
Create_time	The time the video was created
Desc	Description the author wrote under the video "hashtags"
Diversification_labels	?

Challenges	?
Video	VideoData: {id, height, width, duration, ratio, format, bitrate, cover, origin_cover, dynamic_cover, share_cover, reflow_cover, play_addr, download_addr }
Music	MusicData: {id, title, play_url, author_name, duration, original, album, cover_large, cover_medium, cover_thumb}
Image_post	?
Author	The username of the author
Comments	the comments object
creator	Unrelated to the tiktok data ..
tags	Unrelated to the tiktok data ..

A screenshot of output:

id=7189755680704826625 stats=VideoStats(digg_count=59700, share_count=1835, comment_count=1878, play_count=535900) create_time=datetime.datetime(2023, 1, 17, 22, 44, 34, tzinfo=datetime.timezone.utc) desc='فيك أملك لغة تذكري وانتي ،خواتك تذكر ،خرشيك لن نهايتها ما ترضي اله ما البداية #explore #fyp #viral #theoldchallenge #illu #monday #florida #العمري_طارق# حب# #برشلونة #مدريد_ريال ' diversification_labels=None challenges=[LightChallenge(title='ريال_مدريد'), LightChallenge(title='viral'), LightChallenge(title='برشلونة'), LightChallenge(title='fyp'), LightChallenge(title='العمري_طارق'), LightChallenge(title='florida'), LightChallenge(title='illu'), LightChallenge(title='حب'), LightChallenge(title='theoldchallenge'), LightChallenge(title='explore'), LightChallenge(title='monday')] video=VideoData(id=7189755680704826625, height=1024, width=576, duration=100, ratio='720p', format='mp4', bitrate=1110432, cover='https://p16-sign-sg.tiktokcdn.com/obj/tos-alisg-p-0037/0fd6f6b0588745928dbc9504dcd555e9_1673995477?x-expires=1674043200&x-signature=wEIqIJ9o9VkkxQOPGlXj39Wm7To%3D', origin_cover='https://p16-sign-sg.tiktokcdn.com/obj/tos-alisg-p-0037/9f1a00cfe06d4ccbbf5fa928d58d16cc_1673995475?x-expires=1674043200&x-signature=78IK%2BnLd8HKS%2F3Az7dOyNdFtCnA%3D', dynamic_cover='https://p16-sign-sg.tiktokcdn.com/obj/tos-alisg-p-0037/82aa100b98d24c01b3f1a71341fe3bdf_1673995476?x-expires=1674043200&x-signature=EkBt0ZOn%2FbBG6R9PywvCyBRPZ3w%3D', share_cover=[''], reflow_cover='', play_addr='https://v16-webapp.tiktok.com/e78e196005a79580c11cb3cdd5424d3c/63c7e63b/video/tos/alisg/tos-alisg-pve-0037c001/oqzHI5INxIoNCQAEJA4fKzIhU5zQLtAPrwtVRy/?a=1988&ch=0&cr=0&dr=0&lr=tiktok_m&cd=0%7C0%7C1%7C0&cv=1&br=2168&bt=1084&cs=0&ds=3&ft=4b~QyMZD8Zmo0cON864jVAzfPpWrKsdm&mime_type=video_mp4&qs=0&rc=NzU2TTxlODQ6aGY4NDU6ZUBpM203djc6Zm9saTMzODczNEAwYl8uLjNfX18xLzNjMmE0YSMxc2sycjRvNTBgLS1kMS1zcw%3D%3D&l=2023011806280688A76D607190AA04AB82&btag=80000', download_addr='https://v16-webapp.tiktok.com/e78e196005a79580c11cb3cdd5424d3c/63c7e63b/video/tos/alisg/tos-alisg-pve-0037c001/oqzHI5INxIoNCQAEJA4fKzIhU5zQLtAPrwtVRy/?a=1988&ch=0&cr=0&dr=0&lr=tiktok_m&cd=0%7C0%7C1%7C0&cv=1&br=2168&bt=1084&cs=0&ds=3&ft=4b~QyMZD8Zmo0cON864jVAzfPpWrKsdm&mime_type=video_mp4&qs=0&rc=NzU2TTxlODQ6aGY4NDU6ZUBpM203djc6Zm9saTMzODczNEAwYl8uLjNfX18xLzNjMmE0YSMxc2sycjRvNTBgLS1kMS1zcw%3D%3D&l=2023011806280688A76D607190AA04AB82&btag=80000') music=MusicData(id=7189755693384846081, title='الصوت الأصلي', play_url='https://sf16-ies-music-sg.tiktokcdn.com/obj/tiktok-obj/7189755704582966018.mp3', author_name='العُمري طارق', duration=100, original=False, album=None, cover_large='https://p16-sign-sg.tiktokcdn.com/aweme/1080x1080/tos-alisg-avt-0068/be9721e30e14f5dd223feef8d588d569.jpeg?x-expires=1674194400&x-signature=Vj1nqSokQqyCHFpkUWolz2zdP8o%3D', cover_medium='https://p16-sign-sg.tiktokcdn.com/aweme/720x720/tos-alisg-avt-0068/be9721e30e14f5dd223feef8d588d569.jpeg?x-expires=1674194400&x-signature=I29E8fvy%2FQSeqNs0dMVCthhmxuc%3D', cover_thumb='https://p16-sign-sg.tiktokcdn.com/aweme/100x100/tos-alisg-avt-0068/be9721e30e14f5dd223feef8d588d569.jpeg?x-expires=1674194400&x-signature=jyukjrqkMR9O3q9C5L5qUi1mMxk%3D') image_post=None author='ii94n' comments=[] creator=<tiktokapipy.async_api.AsyncLightUserGetter object at 0x10ae85340> tags=<tiktokapipy.async_api.AsyncLightChallengeIter object at 0x10ae8d550>

▼ **MetaData for Twitch:**

Twitch User Keys:

Keys	Description
uniqueId	Unique user identifier in twitch
verified	Flag that indicate if the user is verify or not
followersCount	Number of followers
followsCount	Number of followed users
username	name that is display to other user
description	bio set for the user
channelId	Unique identifier for user channel
videos	videos list
followers	list of followers
follows	list of follows

Twitch Video Keys:

Keys	Description
uniqueId	video identifiers
title	video title
viewersCount	number of viewers
game	game name
createdAt	the date when the video created
publishAt	the date when video is publish
length	the length of videos in seconds
tags	list of tags in video
commentCount	number of comment in the video

Twitch Comments Keys:

Keys	Description
comment	the comment text
author	the user that write the comment
date	the created date for the message
userBadges	the badges associated to the author

▼ **MetaData for YouTube:**

YouTube comment keys:

Keys	Description	collected
cid	comment id	x
text	the comment text	x
time	text → explain time elapsed since the comment was placed (i.e. "1 year ago"	x
author	first and last name of the comment author	x
channel	the channel of the author (URL)	x
votes	Number of likes on the comment	x
photo	URL to the photo of the author	x
heart	boolean → comment has a ……	x
reply	boolean → comment has a reply	x
conversation	replies"comments" to the comment	NY
dislike	number of dislikes on comment	NA (not shared publicly)

YouTube video keys:

Keys	Description	collected
id	the id that appears in the video url	x
title	video title	x

published time	text describing elapsed time. (i.e. 1 hour ago)	x
duration	video length	x
view counts	text describing how many views the video have	x
hashtags	hashtags in the description	NY
description	user defined description on the video	NY
likes	number of likes	NY
dislikes	number of dislikes	NA
comment counts	number of comments on the video	NY
category		NA

YouTube channel keys:

Keys	Description	collected
channel id	as seen in the url	NY
channel avatar	url of the channel user avatar	NY
channel name	name displayed on the channel	NY
subscribers count	number of subscribers	NY
videos count	number of videos in the channel	NY
description	user defined description of the channel	NY
join data	text of channel creation date	NY
views count	number of views of all videos	NY
location (opt)	user defined location	NY
twitter link (opt)	link to user defined twitter account	NY
tiktok link (opt)	//	NY
instagram link (opt)	//	NY
soundcloud link (opt)	//	NY
channels	list of urls of other channels the channel is subscribing to	NY
playlists	list of playlist of videos	NY - TBD
videos	list of urls of videos uploaded by the user	NY
community	list of community posts	NY - TBD
shorts	list of short videos urls	NY - TBD

Keys	Description
channel id	as seen in the url
channel avatar	url of the channel user avatar
channel name	name displayed on the channel
subscribers count	number of subscribers
videos count	number of videos in the channel
description	user defined description of the channel

join date	text of channel creation date
views count	number of views of all videos

Data format for all

Comment (Main focus)

platform_name	username	mention	video_url	text	timestamp	likes_count

json samples:

```
{

    "video_url":
    "plaform_id": "xxxxxxxxxx", //{youtube: 1 , twitch: 2, tiktok: 3}
    "message_id": "xxxxxxxxxx",
    "message": "actual message goes here",
    "message_type": "text_message", //
    "timestamp": 1613761152565924, //format is unix
    "time_in_seconds": 1234.56,
    "time_text": "20:34",
    "author": {
        "id": "UCxxxxxxxxxxxxxxxxxxxxxxx", //platform given id if exists
        "name": "username_of_sender", //
    },
    "mention": {
        "id": "UCxxxxxxxxxxxxxxxxxxxxxxx", //platform given id if exists
        "name": "username_of_sender", //
    },
    "stats": {
        "likes_count": 123
    },

    //"label": key to be added
}
```

Video *(Future)*

platform_name	username (uploader)	url			

User *(Future)*

platform_name	username		

Indexing

- to avoid redundancy with numbering, depends on the models tbd later