Analysing Toxicity in Formula 1 Fandom - Computational Analysis of Communications Final

Author: Leon Knorr

Matr-Nr: 1902854

Introduction

Formula 1 is the highest class of international racing for open-wheel single-seater formula racing cars and is generally considered the most competitive, fastest and hardest class of motor racing. Since it's first season in 1950, Formula 1 is visiting a diverse list of many different countries, where the best drivers in the world are racing against each other in teams of two drivers to determine the best driver and the best team on the Formula 1 grid (Lowrey, 2019). These events are visited by thousands of Fans, with millions more following them on television and social media. With the 2021 season being one of the closest and most entertaining seasons in the history of Formula 1, where Red Bulls Max Verstappen beat Mercedes driver Lewis Hamilton in the grand finale of the season under controversial circumstances after a full season of controversy, drama and intense on track battles and with the release of Netflix Drive To Survive, Formula 1s popularity is growing rapidly. But, reports of Toxic and abusive Fan behavior at events and in comment sections on social media are accumulating, and casts an ugly shadow over Formula 1s latest successes (Woodhouse, 2022). As the reports over toxic and abusive fan behaviours in social media and at live events are rising, Formula 1 as well as Fans and drivers are taking a stand against toxicity in the Formula 1 community. However, an independent and scientific analysis of this topic is missing and therefore the accusations are sort of hanging in the air without a solid scientific foundation. Therefore, in order to tackle this problem research into the toxicity of Formula 1 fandom is a necassety to gain valuable insights into understanding the problem, where it originates from and to build a foundation for future measures to make attending Formula 1 events as well as the media around it a safer and more enjoyable experience. To take the first step into this direction, this thesis will analyse Youtube comments of the Formula 1 channel in order to determine:

- If the Formula 1 fandom is toxic
- Are there specific groups that are more toxic then others?
- Is the toxicity a "self-made" problem of Formula 1 and where is the toxicity originating from?

Fundamentals

In this chapter the necessary fundamental knowledge is presented.

Formula 1

Formula 1 is the worlds most prestigous motor racing competition, as well as the world's most popular annual sporting series (Lowrey, 2019). It marks the highest class of international open-wheel single-seater formula racing. The first Formula 1 competition was held in 1950, since then the competiton for the world drivers championship (wdc) which determines the worlds best driver and the world constructors championship (wcc) which determines the best team, is held annualy and is sanctioned by the Fédération Internationale de l'Automobile (FIA). During the competition (also called a season), Formula 1 visits a variety of different countries and racing tracks, each event (Grands Prix) is attended by thousands of people with millions watching from home ("Formula One," 2023). All rights of the Formula 1 brand and the competition itself is owned by Formula One World Championship Limited, which is a corporation, that provides media distribution and promotion services, besides that, it controls the contracts, distribtution, and commercial management of rights and licenses of formula 1 (*Formula One World Championship Ltd - Company Profile and News - Bloomberg Markets*, n.d.). The term Formula 1 is used to describe the corporation, as well as the competition, as they can't exist without each other.

What is Fandom

According to Cornel Sandvoss Fandom is a community of people that are regularly, consuming a given popular narrative or text with great emotional involvement ("What Is Toxic Fandom?" n.d.). The members of the community are called fans, which is a short form of "fanatic" (Arouh, 2020). In other words, a fandom is a community of people that are fanatic about a popular narrative or text such as a tv series, movie franchise or sports.

Becoming a fan starts with the adoption of a fan identity about a fan object, thus fandom can be a powerful of defining the self. The fan object can be anything that people can be fanatic about, this may be a simple object such as trains or a virtual asset such as a movie franchise. Therefore, by taking part in a fandom, people are expressing themselfs through an identity they've chosen for themselfs. As a result, fans may lead to see the fan object as an extension of themselfs and thus react personally threatened if the fan object is facing a threat such as accusations etc ("What Is Toxic Fandom?" n.d.). In addition to creating a strong part of their own identity, fans feel more connected or socialised through their fandom, as studies indicate, that even if fans don't interact with other members of a fan community, they still perceive themselfs as part of that community. Because of that, fans not only become personally invested in their fandom, they become socially invested as well ("What Is Toxic Fandom?" n.d.).

As a result of the strong connection fans build up to their fan object, the time-frame in which this self identity has been chosen is also playing a role. As an example, many people build a fandom in their childhood about a tv series, franchise or sport, this often leads to them feeling entitled to having their fan

object preserved as they deem acceptable. This behaviour is also called fan entitlement. A good example for this behaviour are the news movies and series in the Lord of the Rings and Star Wars franchises, as most fan communities of these franchises have been outraged about the new characters and story lines, where many people claimed that this "ruined their childhood" ("What Is Toxic Fandom?" n.d.).

From an economic point of view, fandom and fan cultures are seen as the ideal costumers. They are eager to get their hands on the newest products and they are stable with re-occuring purchases, since intense consumption is considered a part of the fan identity (Arouh, 2020).

Defining Toxic Fan behaviour

Youtube API

Concept

The Dataset

```python
from dotenv import dotenv_values
import googleapiclient.discovery
import pandas as pd


api_keys = dotenv_values("keys.env")
api_service_name = "youtube"
api_version = "v3"
api_key = api_keys["YOUTUBE_API_KEY"]
max_results = 1000
youtube_api = googleapiclient.discovery.build(api_service_name, api_version, developerKey =

Formula1_official_channel = youtube_api.channels().list(part='snippet' ,forUsername='Formula
videos_after_2020 = youtube_api.search().list(channelId=Formula1_official_channel["id"],
        maxResults=max_results,
        publishedAfter="2020-01-01T00:00:00Z",
        part='id').execute()
video_ids_after_2020 = [item['id']['videoId'] for item in videos_after_2020['items']]
while len(video_ids_after_2020) < max_results and "nextPageToken" in videos_after_2020.keys(
        videos_after_2020 = youtube_api.search().list(channelId=Formula1_official_channel["i
        maxResults=max_results,
        publishedAfter="2020-01-01T00:00:00Z",
        part='id',
        pageToken=videos_after_2020["nextPageToken"]).execute()
        video_ids_after_2020 = video_ids_after_2020 + [item['id']['videoId'] for item in vid

df_list = []
for video_id in video_ids_after_2020:
    video_data = youtube_api.videos().list(part='snippet, statistics', id=video_id).execute(
```

```python
    snippet = video_data['items'][0]['snippet']
    statistics = video_data['items'][0]['statistics']
    df_list.append(
    {
        "video_id":video_id,
        "title": snippet['title'],
        "description": snippet['description'],
        "channel": snippet['channelTitle'],
        "published_at": snippet['publishedAt'],
        "tags": snippet['tags'] if "tags" in snippet.keys() else None,
        "like_count": statistics['likeCount'],
        "favorite_count": statistics['favoriteCount'],
        "comment_count": statistics['commentCount'] if "commentCount" in statistics.keys() e
    })

videos = pd.DataFrame(df_list)
videos

video_ids_after_2020 = videos.video_id.to_list()
video_ids_after_2020

df_list_comments = []
for video_id in video_ids_after_2020:
    if videos.loc[videos['video_id'] == video_id].comment_count.iloc[0] == 0:
        continue
    top_level_comments = youtube_api.commentThreads().list(part="snippet",
        maxResults=50,
        order="relevance",
        videoId=video_id).execute()['items']
    for top_level_comment in top_level_comments:
        replies = youtube_api.comments().list(part="snippet",
            maxResults=50,
            parentId=top_level_comment['snippet']['topLevelComment']['id']).execute()['items
        df_list_comments.append(
        {
            "video_id": video_id,
            "id": top_level_comment['snippet']['topLevelComment']['id'],
            "text": top_level_comment['snippet']['topLevelComment']['snippet']['textDisplay
            "user": top_level_comment['snippet']['topLevelComment']['snippet']['authorChanne
            "like_count": top_level_comment['snippet']['topLevelComment']['snippet']['likeCo
            "published_at": top_level_comment['snippet']['topLevelComment']['snippet']['publ
            "reply_count": top_level_comment['snippet']['totalReplyCount']
        })
        for reply in replies:
            df_list_comments.append(
            {
                "video_id": video_id,
```

```python
                "id": reply['id'],
                "text": reply['snippet']['textDisplay'],
                "user": reply['snippet']['authorChannelId']['value'],
                "like_count": reply['snippet']['likeCount'],
                "published_at": reply['snippet']['publishedAt'],
                "reply_count": 0
        })

comment_df: pd.DataFrame = pd.DataFrame(df_list_comments)
comment_df

videos.to_pickle("datasets/video_data.pkl")
comment_df.to_pickle("datasets/comment_data.pkl")

videos: pd.DataFrame = pd.read_pickle("datasets/video_data.pkl")
comment_df: pd.DataFrame = pd.read_pickle("datasets/comment_data.pkl")

videos
```

Dataset limitations

Dictionary Analysis

Othrus Lexicon for Toxicity

```python
comment_df
```

```python
with open("dictionaries/toxic_words.txt") as toxic_words_file:
    set_of_toxic_words: set = set([word.strip() for word in toxic_words_file.readlines()])
set_of_toxic_words
```

```python
import numpy as np
from collections import Counter
toxic_word_counter: Counter = Counter()
toxic_word_count: list = []
for row in comment_df.text:
    toxic_words_in_comment: set = set(row.split(" ")).intersection(set_of_toxic_words)
    toxic_word_counter.update(toxic_words_in_comment)
    toxic_word_count.append(len(toxic_words_in_comment))
comment_df["toxic_word_count"] = toxic_word_count
```

```python
comment_df.loc[comment_df["toxic_word_count"] > 0]
```

```python
toxic_word_counter
```

Grievance Dictionary

Ethnic Slurs

```python
from os import listdir
import os.path
```

```
dict_files: list = list(filter(lambda f: f[-4:] == ".csv" ,listdir("dictionaries/ethnic_slur
dict_df: pd.DataFrame = pd.DataFrame()
for file in dict_files:
    part = pd.read_csv(os.path.join("dictionaries/ethnic_slurs", file))
    dict_df = pd.concat([part, dict_df])
dict_df.reset_index(inplace=True, drop=True)
dict_df
```

Transformer Classifiers

Results

Bibliography

Arouh, M. (2020). Toxic Fans: Distinctions and Ambivalence. *Ex-Centric Narratives: Journal of Anglophone Literature, Culture and Media*, *4*, 67–82. https://doi.org/10.26262/exna.v0i4.7917

Formula One. (2023). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Formula_One&oldid=1132488425

Formula One World Championship Ltd - Company Profile and News - Bloomberg Markets. (n.d.). Retrieved January 9, 2023, from https://www.bloomberg.com/profile/company/1935454Z:LN

Lowrey, T. (2019). *About F1 | Formula One World Championship Limited*. https://corp.formula1.com/about-f1/

What Is Toxic Fandom? (n.d.). In *Verywell Mind*. Retrieved January 7, 2023, from https://www.verywellmind.com/what-is-toxic-fandom-5214499

Woodhouse, J. (2022). "Scary" how toxic Formula 1 community became during 2021. In *PlanetF1*. https://www.planetf1.com/news/toxic-formula-1-community-2021/