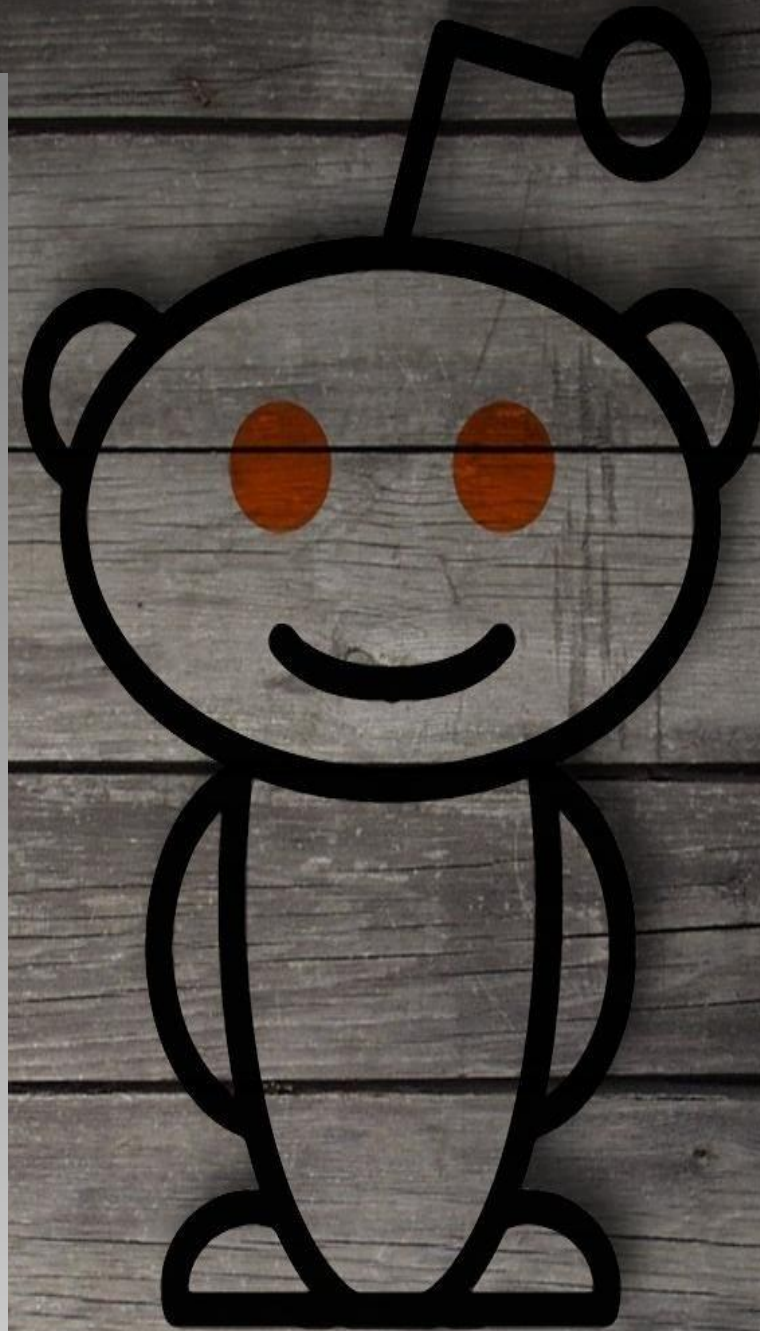# The **Bright** and **Dark** Sides of Reddit.

(1324 words)

FEBRUARY 21, 2020.

INDIANA UNIVERSITY (BLOOMINGTON)

Authored by: VARSHA RAVI VARMA
IU ID: 2000751388
Sem: Spring 2022

reddit

Introduction

> *"Reddit is an American social news aggregation, web content rating and discussion website. Registered members submit content to the sites such as links, text posts, images and videos, which are then voted up or down by other members."* - Wikipedia

Founded in 2005, Reddit is a network of communities that allows people to explore their interests, hobbies, and passions. It is home to thousands of communities, endless conversation, and genuine human connections. There are millions of people worldwide who post, vote, and comment in communities that are based around their interests [1]. If we talk about interests, they can include either positive, informative, motivational content such as news, sports, fitness, or negative, infuriating, and upsetting content like rape, murder, suicide, underage sexual content, etc. It doesn't matter which domain you're interested in. Reddit has a community for you. Posts are categorized by topic into boards called 'communities' or 'subreddits', each representing a different area of interest [2]. The subreddit 'r/cute', for instance, contains animal cutesy content. Posts are arranged according to their vote counts (highest to lowest) under these subreddits.

Several domains such as positive, cheerful, motivating, etc. will be considered as the bright side of Reddit in this study, while negative, disturbing, explicit, hateful, etc. will be considered as the dark side. Some social media platforms attempt to censor the explicit content displayed on their websites, but reddit displays its bold side by allowing all wild and explicit content. It is safe to say that the limitations on users have increased in recent years, but the platform still allows them to post sensitive content without restriction unlike other platforms like Facebook, Instagram, and Pinterest which have strict policies on content [3]. The purpose of this study is to evaluate the bright side and dark side of Reddit, which includes the following:

- o How are the two sides of Reddit being used?
- o Is Reddit mostly used for the right side of the platform?
- o The type of content this platform offers in terms of bright and dark content.
- o A user's opinion on these types of content.

Research question

The research question for this study can be phrased as follows:

As two sides of Reddit, bright (positive) and dark (negative) contents, how much is each being interpreted and used by the users?

Method

Data

Data source: Reddit [4]

Data: Posts and its comments section

As the data was gathered manually from Reddit, the legitimacy of the data can be affirmed. In this study, we collect and analyze the 10 most popular posts within each subreddit. To understand how the users of Reddit are using the platform in the present, we have used the latest data generated, which is not older than 1 year. We are unable to examine all the bright and dark content available on Reddit due to the lack of clean and augmented data. As a result, only a few of the most common and important topics have been considered for this study. In particular, they are:

- Bright side:
    - Fitness
    - Interesting facts
    - Sports
    - Animals
    - Nature

- Dark side:
    - Rape
    - Murder
    - Abuse
    - Suicide
    - Human trafficking

Study material consists of the top ten most upvoted posts from each subreddit of the above-mentioned domains. Using content analysis, data from Reddit's comments on posts was analyzed. With a growing number of Reddit users, a large user database, and the persuasive nature of Reddit, it is only natural that it is extensively used by people who leave their comments over all kinds of content. These responses allow us to analyze how the content is being interpreted by users, as well as how it is being used by them. These responses form the basis of the analysis.
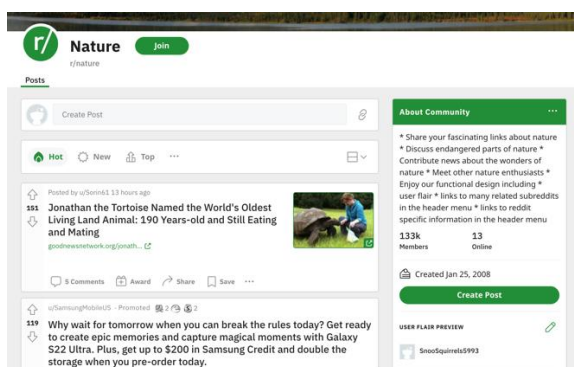

Fig 1: View of a subreddit for nature.

https://www.reddit.com/r/nature/

Fig 2: View of the comments section of a post under a subreddit for suicide.

https://www.reddit.com/r/SuicideWatch/comments/sx7oln/i_failed_a_suicide_attempt_a_couple_days_ago_and/

Analysis

The comments of the users indicate whether the user views the post as bright content or not. As part of the snowball approach, we started with a few seed subreddits, analyzed the comments. A daily crawl of their posts in the new category was obtained for each of these subreddits. As corresponding to each post, the comments were categorized into positive (bright) and negative (dark) and along with that, the number of upwards and downwards it had achieved thus far. In reading and analyzing the comments listed under the subreddit posts, it is safe to conclude that there is a mixture of positivity and negativity in all sorts of posts, regardless of which side it appears on. Most of the comments on the bright-sided posts are positive, with a few negative ones sprinkled in. There is a great deal of sarcasm in the dark posts, along with anger, frustration, and disgust.

Based on the data presented in the table below, we can perform a quantitative analysis to determine the number of users interested in the bright and dark domains. In comparison with the other nine subreddits that we took into consideration, the fitness subreddit posts had extremely positive feedback and also was the second most active subreddit (9.1 million members). One of the top ten most popular subreddits is the interesting facts subreddit, which contains contents from around the world that are labeled as interesting. Even though sports are one of the most important and popular topics, the subreddit only has around 3500 members. Suicide subreddit has three to four times the members as animals or nature subreddits. The table below shows that subreddits dedicated to fitness, interesting facts, and suicide are among the most popular among users.

	Domain	Subreddit	Member count
1.	Fitness	r/fitness	9,100,000
2.	Interesting facts	r/interestingasfuck	9,200,000
3.	Sports	r/sport	3,500
4.	Animals	r/animals	78,700
5.	Nature	r/nature	133,000
6.	Rape	r/rape	36,200
7.	Murder	r/murder	6,100
8.	Abuse	r/emotionalabuse	21,400
9.	Suicide	r/suicidewatch	334,000
10.	Human trafficking	r/humantrafficking	5,300

Table 1: Subreddit data popularity data for the 10 selected domains

The below graphs obtained from Reddit shows the usage of both bright and dark contents:
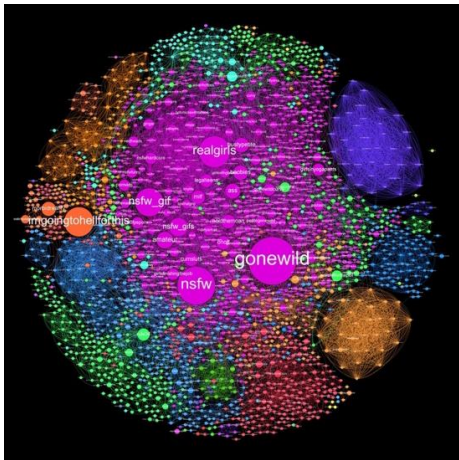

Fig 3: The types of dark content such as NSFW (not safe for work) and wild explicit contents and its magnitude.

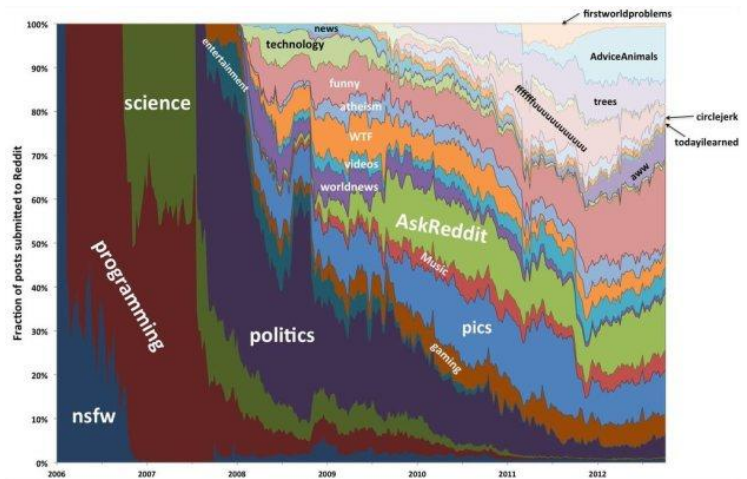https://www.reddit.com/r/coolguides/comments/rky90l/map_of_all_the_nsfw_subs/

Fig 4: The most used subreddits during the early years of platform release.

https://www.reddit.com/r/dataisbeautiful/comments/21u8wj/breakdown_of_reddit_content_by_subreddit_20062013/

In recent years, Reddit's censorship policies have become less explicit as the platform has improved. In response to a ban, Reddit's dark side was formatted and most of the explicit wild data was purged. Before the ban, the dark and light sides of Reddit were equally popular, but now dark content has declined steeply in usage.

Results

Whenever we analyze Reddit, we can either see the bright side or the dark side. Some users might interpret a post or domain as negative, while others may see it as positive. Based on the data, we can observe that users have spread the concept of self-love even on dark content such as suicide. On the other hand, we can also find negative and demotivating comments associated with lively domains such as animals and nature. We can only take into account the data and determine which side of Reddit do the comments primarily belong to. In the case of a majority of negative comments, the post is most likely to fall under the dark side of Reddit and vice versa. Based on this analysis, we can determine which side is being contributed to more frequently by the users. Since the implementation of the ban, there is no doubt that a minimum 70% of the dark content has been removed, which results in an imbalance of the statistics between the two sides of the debate.

Below are graphs illustrating the effectiveness of the ban and how much data was removed.

- Harassment: 15,870 (31.4%)
- Ban evasion: 11,692 (23.2%)
- Minor sexualization: 10,781 (21.4%)
- Violent content: 5,990 (11.9%)
- Involuntary porn: 2,544 (5.0%)
- Controlled goods: 1,639 (3.2%)
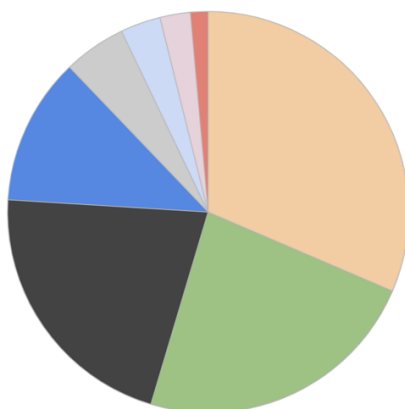- Private information: 1,234 (2.4%)
- Impersonation: 717 (1.4%)

Fig 5: Shows stats of the top dark domain contents available on Reddit before the ban.

https://www.redditinc.com/policies/transparency-report-2019

Vote manipulation, 9,412, 1% Involuntary porn, 7,158, <1%   Prohibited goods, 5,701, <1%
Violent content, 10,653, 1%   Minor sexualization, 21,946, 2%
Harassment, 17,268, 1%   Impersonation, 1,033, <1%
Personally identifiable
information (PII), 4,012, <1%
Hateful content, 67,620, 5%

Ban evasion, 1,267,214, 90%
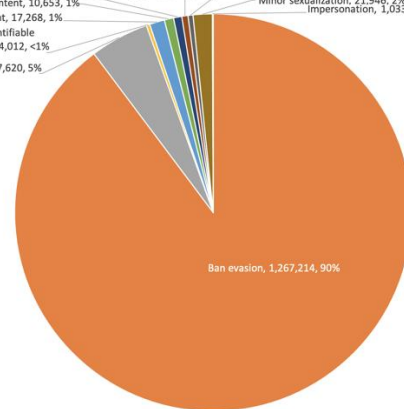
Fig 6: Shows stats of the top dark domain contents available on Reddit after the ban.

https://www.redditinc.com/policies/transparency-report-2020-1

In this way, the bright content dominates by a clear margin. Reddit is trying to be characterized as a bright, positive platform.

Conclusions and Limitations

A study of the past years' Reddit content would make it difficult to determine whether the dark side or the light side of Reddit is more popular and more widely used. Since this is a study related to the current situation in which the ban on the dark side of Reddit has been implemented, it is fairly easy to conclude that the bright side of Reddit has been fully embraced. Since we all want the internet to be a source of positive content rather than dark negative ones, it is difficult to argue against the ban on explicit uncensored content. In conclusion, the world primarily uses the bright side of reddit at the present, which contains positive, good, happy, motivating, and informative content.

Limitations:

o   The study lacks the access to the authentic dark side of Reddit since it has been banned now. Only stats and text information of it remains.
o   Reddit keeps changing its policies very often which makes it hard to study a wide time frame.

References

Introduction:

[1] Jamnik, Matthew R., and David J. Lane. "The use of Reddit as an inexpensive source for high-quality data." *Practical Assessment, Research, and Evaluation* 22.1 (2017): 5.

[2] Anderson, Katie Elson. "Ask me anything: what is Reddit?." *Library Hi Tech News* (2015).

[3] Fiesler, Casey, et al. "Reddit rules! characterizing an ecosystem of governance." *Twelfth International AAAI Conference on Web and Social Media*. 2018.

Data:

[4] https://www.reddit.com/

Analysis:

Tadesse, Michael M., et al. "Detection of depression-related posts in reddit social media forum." *IEEE Access* 7 (2019): 44883-44893.

Schrading, Nicolas, et al. "An analysis of domestic abuse discourse on reddit." *Proceedings of the 2015 conference on empirical methods in natural language processing*. 2015.

Mohan, Shruthi, et al. "The impact of toxic language on the health of reddit communities." *Canadian Conference on Artificial Intelligence*. Springer, Cham, 2017.

Copland, Simon. "Reddit quarantined: Can changing platform affordances reduce hateful material online?." *Internet Policy Review* 9.4 (2020): 1-26.

De Choudhury, Munmun, and Sushovan De. "Mental health discourse on reddit: Self-disclosure, social support, and anonymity." *Eighth international AAAI conference on weblogs and social media*. 2014.

Transparency report 2020:
https://www.redditinc.com/policies/transparency-report-2020-1

Transparency report 2021:
https://www.redditinc.com/policies/transparency-report-2021-2