Investigating Human Values in Online Communities

Anonymous EMNLP submission

Abstract

Human values play a vital role as an analytical tool in social sciences, enabling the study of diverse dimensions within society as a whole and among individual communities. This paper addresses the limitations of traditional survey-based studies of human values by proposing a computational application of Schwartz's values framework to Reddit, a platform organized into distinct online communities. After ensuring the reliability of automated value extraction tools for Reddit content, we automatically annotate six million posts across 10,000 subreddits with Schwartz values. Our analysis unveils both previously recorded and novel insights into the values prevalent within various online communities. For instance, when examining subreddits with differing opinions on controversial topics, we discover higher universalism values in the Vegan subreddit compared to Carnivores. Additionally, our study of geographically specific subreddits highlights the correlation between traditional values and conservative U.S. states.

Val	Subreddits
AC	r/study, r/resumes, r/PhD
BE	r/AmItheAsshole, r/motherinlawsfromhell, r/family
CO	r/Anger, r/confession, r/pettyrevenge,
HE	r/happyrelationships, r/sex, r/depression
PO	r/FinancialPlanning, r/smallbusiness, r/GamblingAddiction
SE	r/foreignpolicy, r/warinukraine, r/CombatFootage
SD	r/antidepressants r/AnorexiaRecovery, r/sugarfree, r/Allergies
ST	r/ShouldIbuythisgame, r/socialanxiety, r/MakeNewFriendsHere, r/FreeCompliments
TR	r/PrayerRequests, r/Christians, r/Anglicanism
UN	r/mixedrace, r/IsraelPalestine, r/conservation, r/LGBTnews

Table 1: Subreddits with the highest expression of each one of the ten Schwartz values. AC=achievement, BE=benevolence, CO=conformity, HE=hedonism, PO=power, SE=security, SD=self-direction, ST=stimulation, TR=tradition, UN=universalism.

1 Introduction

Values have been a useful analysis tool for social sciences scholars (Ponizovskiy et al., 2020; Boyd et al., 2021; Schwartz, 1994). Applied to communities or individuals, they are used to study political affiliations, cultural integration, human disagreement, economic growth, and human development, among others. Many frameworks exist for studying human values, including Rockeach values (Rokeach, 1973), Hofstede's cultural dimensions (Hofstede, 1984), and Moral Foundations Theory (Graham et al., 2013). However, studies using these value frameworks often struggle with sample sizes and rely on self-reported surveys to calculate the values of communities (Weld et al., 2023), leading to concerns about the generalisation of the results to populations beyond the ones studied (Gerlach and Eriksson, 2021; Bašnáková et al., 2016; Fontaine et al., 2008).

In contrast, social media (SM) platforms provide unadulterated access to vast and diverse expressions of human thoughts and opinions in the form of posts, discussions, and comments. This rich data source is invaluable for investigating various aspects of human society and has the potential to overcome many challenges faced by social science scholars (Newell et al., 2016; Agrawal et al., 2023; Zomick et al., 2019; Turcan and McKeown, 2019).

In this work, we computationally apply a value framework on Reddit at scale to supply scholars with a complementary tool when studying online behaviour and communities. Specifically, we follow Schwartz's Theory of Human Values (Schwartz, 1994) due to its wide adoption as a value framework in the social sciences as well as in Natural Language Processing. Reddit is conceptually based on named communities, or, **subreddits**,

Value	Description
Power	Social status and prestige, control or dominance over people and resources
Achievement	Personal success through demonstrating competence according to social standards.
Hedonism	Pleasure and sensuous gratification for oneself.
Stimulation	Excitement, novelty, and challenge in life.
Self-direction	Independent thought and action-choosing, creating, exploring.
Universalism	Understanding, appreciation, tolerance, and protection for the welfare of all people and for nature.
Benevolence	Preservation and enhancement of the welfare of people with whom one is in frequent personal contact.
Tradition	Respect, commitment, and acceptance of the customs and ideas that traditional culture or religion provide.
Conformity	Restraint of actions, inclinations, and impulses likely to upset or harm others and violate social expectations or norms.
Security	Safety, harmony, and stability of society, of relationships, and of self.

Table 2: The ten Schwartz values and their meaning; descriptions from Schwartz (1994)

and studying the values they exhibit is particularly interesting for studying online behaviour as the discussions are already segregated by topic or subcommunities. For instance, by examining the `r/teenagers` subreddit, we can gain insights into adolescent perspectives, while the `r/ConservativeValues` allows us to study the impact of conservative values on worldviews.

To extract values from Reddit, we train a supervised value extraction model to classify the presence of Schwartz values in text and apply that to Reddit posts. We then evaluate the model and validate its effectiveness, addressing gaps in existing research. Subsequently, we use this model to infer the expressed values of 6 million user posts across 10,000 most popular subreddits on Reddit.

Through such large-scale computational value analysis, we are able to analyse digital trace and behaviour data at scale, complementing social science studies by highlighting patterns as well as outliers needing further study. As an instance of this, through our analysis, we find that people contributing to `r/vegan` demonstrate higher values in universalism, as substantiated by previous studies. We also demonstrate the prevalence of high traditional values in conservative US states. In sum, our **contributions** are:

- We propose a flexible and extendable method of conducting large-scale value analysis to complement work done in the social sciences to study online communities;
- We conduct values analyses at a large scale and find both previously recorded phenomena as well as novel insights that can complement traditional methods in understanding societal behaviours;
- We release a dataset of 10k subreddits with their corresponding Schwartz values for future analysis.[1]

Although our findings offer unique insights into the values of individuals from various cultures, we emphasize that the focus of our paper is the analysis of values present in *online communities*. That is, the subject of our study is the communities themselves and how they operate, rather than the broader cultural backgrounds of their members, which we do not measure directly. Moreover, our methods are designed to complement, not replace, traditional survey-based approaches for studying values, by providing additional perspectives from digital trace data.

2 Related Work and Background

2.1 Schwartz's Values Framework

Values represent a crucial aspect of human nature. According to Schwartz (2012), *"A person's value priority or hierarchy profoundly affects his or her attitudes, beliefs, and traits, making it one core component of personality."*

Schwartz (1994) defines values as (1) concepts or beliefs, that (2) pertain to desirable end states or behaviors, (3) transcend specific situations, (4) guide selection or evaluation of behavior and events, and (5) are ordered by relative importance. Based on this, they outline ten basic human values: *Security, Conformity, Tradition, Benevolence, Universalism, Self-direction, Stimulation, Hedonism, Achievement, Power*. We provide descriptions of each of the values in Table 2. The values

[1] `https://github.com/[anonymised]/[anonymised]`

were originally applied to measure differences in values across cultures (Schwartz, 1994). The framework is suggested as a tool to study populations rather than individual people, hence it serves as a suitable tool for the analysis of online communities. We leverage the fact that a person's identity and values are often reflected in the linguistic choices they make (Jaffe, 2009; Norton, 1997) to analyse values embedded in text.

2.2 Values and Natural Language Processing

Recently, there have been a number of studies exploring values and morals using Natural Language Processing and language models. The closest to our study is work done by van der Meer et al. (2023). They train a value extraction model based on datasets from Qiu et al. (2022) and Kiesel et al. (2022). Using the model, they analyse values at an individual user level to understand disagreement in online discussions. We, on the other hand, perform an analysis at the community level in order to understand the values of communities rather than individuals. In analysing personal values and language on social media, Ponizovskiy et al. (2020) released a Schwartz value dictionary. Boyd et al. (2021) conduct a study on assessing values from self-reported questionnaires and posts on Facebook, concluding that free response language data shows greater promise for value extraction. There have also been studies exploring morals and norms in text. Trager et al. (2022) release the Moral Foundations Reddit Corpus with 16k Reddit comments and Roy et al. (2021) similarly study morality in political tweets.

2.3 Studying Online Communities

Much work has been done to explore Reddit and its user base, focusing on diverse aspects. Reddit users' personalities have been studied (Gjurković and Šnajder, 2018) as well as their mental health (Zomick et al., 2019; Turcan and McKeown, 2019). Previous work has also studied Reddit by focusing on events affecting community dynamics (Newell et al., 2016; Agrawal et al., 2023). For instance, Chandrasekharan et al. (2022) examines the effect of content moderation on two controversial subreddits. Soliman et al. (2019) studies the characteristics and differences between left-leaning and right-leaning political subreddits. Closer to our work, Weld et al. (2023) craft a taxonomy of community values and associate them with specific subreddits. However, their methodology involves collecting values through self-reporting questionnaires of limited size and scale, a limitation we address in this paper.

	Avg	std	Percentile			Total
			25	50	75	
p.p.s	579.5	206.0	396	570	754	5736192
w.p.p	77.6	197.7	14	24	67	445404141

Table 3: Statistics of the Reddit dataset we analyse. p.p.s is posts per subreddit, and w.p.p is word per post.

3 Method

This section details the collection and processing of the Reddit dataset we use for our experiments, and our approach to training and evaluating the Schwartz values extractor model.

3.1 Data Collection

We download an image of Reddit posts authored between January and August 2022[2] using Pushshift's[3] API. We filter out posts with fewer than ten words or with less than ten upvotes to reduce possible noise from low-quality posts. We then aggregate the posts by subreddit and filter out subreddits that are NSFW, have fewer than 5,000 subscribers, or have fewer than 250 remaining posts. Finally, we down-sample large subreddits to 1,000 random posts and remove non-English content. This process results in a dataset $\mathcal{D}$ of 10k unique subreddits. See Table 3 for dataset statistics.

3.2 Value Extraction Model

For extracting Schwartz values from text, we follow van der Meer et al. (2023) in training a neural Schwartz Values Extractor classifier, updating the RoBERTa model they use to a DeBERTa model (He et al., 2021) due its improved performance and speed. As a single post can potentially express several values simultaneously, the model is trained in a *multi-label setting*, predicting a vector of 10 independent probabilities for each input. We fine-tune the classifier on the concatenation of two supervised Schwartz values datasets, ValueNet (Qiu et al., 2022) and ValueArg (Kiesel et al., 2022). Given a labelled triplet (s, v, y), with s being a string of text from the post, v the name of one of the ten Schwartz values (Ta-

[2]We were restricted to a relatively narrow date range due to our system limitations. An image of a single month is over 300GB in size.

[3]https://pushshift.io/

ble 2), and $y \in \{0, 1\}$,[4] we construct the input $x =$ [CLS] v [SEP] s [SEP] and train the model to predict $p(y|x)$. See App. D.1 for additional training details. At test time, we use the model to predict the probabilities for all Schwartz values for a given input instance, resulting in a vector $\boldsymbol{u} \in [0, 1]^{10}$. Specifically, given a new string s, we construct $\boldsymbol{u} = [p(y^1|x^1), ..., p(y^{10}|x^{10})]$ by replacing v^k in the construction of $x^k =$ [CLS] v^k [SEP] s [SEP] with each one of the ten possible Schwartz values. That is, each entry in the vector is the independent probability that s expresses the value k, supporting a multi-label approach

We then assign each subreddit $\mathcal{S} \in \mathcal{D}$ a single vector of Schwartz probabilities $\boldsymbol{u}_{\mathcal{S}}$. Given $\mathcal{S}$ with posts $(t_i, b_i)|i \in \mathcal{S}$, where t_i is the post title and b_i is the post body, we construct s_i by concatenating t_i with b_i. We then predict $\boldsymbol{u}_i$ from s_i using the above-mentioned process. Finally, we calculate a single probability vector for the subreddit by averaging over the vectors predicted for each post:

$$\hat{\boldsymbol{u}}_{\mathcal{S}} = \frac{1}{|\mathcal{S}|} \sum_{i \in \mathcal{S}} \boldsymbol{u}_i$$

The value extraction model may be more sensitive to some values and less to others, rendering a direct comparisons of any two values within the same subreddit unreliable. For example, the model may be more sensitive to the value of *security* than, say, *benevolence*, meaning that a higher predicted probability for *security* for some subreddit does not necessarily imply the subreddit has a higher *security* signal than *benevolence*. Therefore, instead of using the $\hat{\boldsymbol{u}}_{\mathcal{S}}$ vectors directly, we first normalise them by computing the differences of each vector entry from its baseline. We calculate the baseline signal of each Schwartz value across $\mathcal{D}$ by pointwise averaging the probability vectors:

$$\hat{\boldsymbol{u}} = \frac{1}{|\mathcal{D}|} \sum_{\mathcal{S} \in \mathcal{D}} \hat{\boldsymbol{u}}_{\mathcal{S}}$$

Finally, we define $\boldsymbol{u}_{\mathcal{S}} = \hat{\boldsymbol{u}}_{\mathcal{S}} - \hat{\boldsymbol{u}}$.

3.3 Evaluation of Value Extraction model

An important limitation of van der Meer et al. (2023) is the lack of a direct evaluation of the value extraction model. The need for this becomes particularly pronounced considering the large domain shift of applying the model—which, similarly to us, was trained on debating data—to analyze posts and comments on platforms like Reddit. Therefore, we conduct a thorough evaluation to test the model's capabilities of extracting values from Reddit posts.

Ideally, we would want to assess the model's performance using a large, annotated dataset of randomly sampled Reddit posts. However, this approach is impractical because randomly selected posts are unlikely to contain any Schwartz values. Consequently, annotating these posts would yield a dataset with only a few positive examples. To address this challenge, we evaluate the Value Extraction model directly. Specifically, we use our Value Extraction model to label 10,000 posts for values. For each value v, we sample 10 posts where the model predicted the presence of v with high confidence. We then task three authors of the paper to determine whether the posts indeed exhibit the Schwartz value v, i.e., whether the model is correct in its prediction. To ensure that the annotators are not biased, we add to these posts a set of 10 randomly selected posts, where the value v is unlikely to be expressed, serving as negative samples. Repeating this process for all values, each annotator annotated 200 posts. Refer to App. A for annotation guidelines and dataset samples.

Inner-annotator agreement was good for such a subjective task—the average Cohen's Kappa was 0.59 among the three annotators. Impressively, in 91% of the cases, the annotator's majority vote was aligned with the Value Extractor model's prediction. That is, the model was correct in its prediction for 91% of the posts. Half of the incorrect model predictions are errors in misclassifying the value "*Self-direction*", a value that the annotators noted to be highly ambiguous. See Table 9 in App. B for examples of model misclassification.

4 Experiments

The section serves two primary goals: first, to validate our approach by comparing the values extracted using our method to existing knowledge, and second, to demonstrate its extensibility to new topics that have not been empirically investigated in social sciences (to the best of our knowledge). We start by conducting a preliminary qualitative evaluation of our approach in §4.1. In §4.2, we correlate values extracted using our method to values gained through traditional approaches like surveys and questionnaires. In §4.3, we investigate subred-

[4] As van der Meer et al. (2023), we collapse the two non-neutral labels (-1 and 1) in ValueNet into a single positive class, indicating that the value v plays a role in s.

dits with differing opinions on controversial topics. Finally, we systematically evaluate whether similar online communities share similar values in §4.4.

4.1 Qualitative Analysis

We start by assessing how well our values extraction model performs on our Reddit dataset.

Strong signals Intuitively, certain Schwartz values are expected to be distinctly present within specific online communities. For instance, the value of *tradition* is expected to be significantly present in religion-related communities. Therefore, we find subreddits with particularly strong signals of individual Schwartz values. For each value, we sort the subreddits' values probability vectors $\boldsymbol{u}$ by their entry corresponding to the particular value. Table 1 lists a sample of the top subreddits for each value (see Table 6 in App. B for a complete list of the top 20 subreddits per value). Many of the subreddits collected for each value seem to be semantically related to it (e.g., `r/PhD` with *achievement*, `r/conservation` with *Universalism*). The results are encouraging, demonstrating the effectiveness of our approach and its potential in real-world settings. Nevertheless, some subreddits are less clearly connected to their assigned value (e.g., `r/Allergies` and *self-direction*), connections intriguing to explore.

Value magnitude We hypothesise that online communities differ not only in the set of values they express but also in the total **magnitude** of expressed values. Online communities pertaining to more objective topics (e.g., linear algebra) should express fewer Schwartz values than communities dedicated to subjective or controversial topics (e.g., politics). To test this, for each subreddit $\mathcal{S} \in \mathcal{D}$ we calculate $\text{mag}(\mathcal{S}) = |\boldsymbol{u}_{\mathcal{S}}|_2$, the total magnitude of values expressed in the subreddit. The 20 subreddits with the highest and lowest value magnitudes are detailed in Table 8, located in App. B. Subreddits with particularly high magnitudes are those communities that focus on sharing personal and emotional narratives. Examples of these include `r/gratitude`, `r/confessions`, and `r/venting`. On the other hand, subreddits with notably low magnitudes are generally more objective and neutral in tone. These include `r/identifythisfont`, `r/mushroomID` and `r/analog`. Some subreddits that exhibit low value magnitudes are dedicated to sharing photos taken by community members, e.g., `r/portraits`.

AC	BE	CO	HE	PO	SE	SD	ST	TR	UN
-0.11	-0.31	-0.05	-0.28	0.00	0.03	0.23	-0.12	0.02	0.06

Table 4: Spearman correlation between the Schwartz values obtained from a questionnaire and the values extracted from Reddit.

4.2 Correlation with Surveys

We next compare the Schwartz values extracted from Reddit with those obtained from traditional methods such as questionnaires. While our approach only provides a proxy into the values held by individual community members and cannot be generalized beyond the online communities we study, we believe this analysis can still yield valuable insights.

US States First, we use our method to investigate the premise that conservative US states are more traditional than liberal states. Figure 1 displays the *tradition* value of the 50 states subreddits, colour-coded by the recent US election results – "red" states clearly tend to have higher *tradition* values than "blue" states. Political surveys also support these findings. A Spearman's R of 0.3 (p-value of 0.03) is achieved by correlating *tradition* values extracted using our method with the results of a survey (Center, 2014) reporting the level of conservative ideologies in the different states.

Country-level values Next, we extend this correlation experiment to additional Schwartz values and countries other than the USA. Schwartz and Cieciuch (2022) used questionnaires to analyse the Schwartz values associated with 49 countries and populations,[5] which we correlate with values extracted from subreddits related to these countries using our classifier. We start by manually identifying a set of subreddits relevant to each country studied in Schwartz and Cieciuch (2022).[6] We successfully found at least one subreddit for 41 out of the 49 countries. We also determined which language is the most likely to be used online by citizens of the country (and not, for example, by tourists).[7] See Table 5 in App. B for a comprehensive list of countries and their associated languages and subreddits. Next, we collected posts written in

[5] We will use the term "countries" for simplicity.

[6] We searched for country related subreddits online and filter based on their public description and the posts on their first page, disregarding subreddits that tourists or non-residents majorly use.

[7] We acknowledge that this is a simplification, as there are many countries where more than one language is commonly spoken.

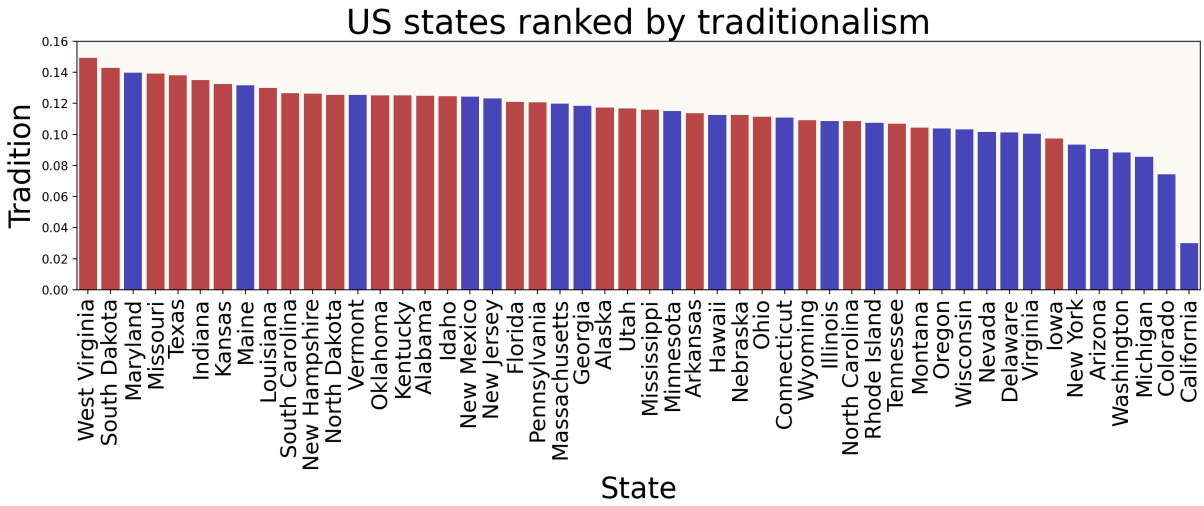


Figure 1: US states sorted by their *tradition* values extracted from Reddit, colour coded based on the recent US election results (Blue – democratic majority. Red – republican majority).

the country's language[8] from the country's subreddits. We then randomly sampled 1,000 posts per country, excluding countries with fewer than 250 posts. Listing 2 lists the 31 countries that passed this filter. Finally, we used Google Translate API to translate the posts into English[9] and applied our approach to extract Schwartz values for each country.

The above-mentioned process generates a country-value matrix of dimensions 31×10 of Schwartz values obtained from Reddit. We compare this matrix to the 49×10 the country-value described in Schwartz and Cieciuch (2022) by removing rows corresponding to countries that do not exist in Listing 2. We then correlate the columns of the two now-identical matrices using Spearman's ρ. Table 4 reports the results. In line with previous studies (Nasif et al., 1991), we find no correlation between the values extracted from SM and values extracted from questionnaires. This lack of correlation may arise from the distinct demographics of Reddit users compared to questionnaire respondents or the unique dynamics of online communities. These results back our position that our method does not replace traditional approaches; rather, the two methods complement each other.

4.3 Controversial Topics

Different online communities have differing viewpoints on widely discussed issues based on their values. To test whether we can also observe this in subreddits, we extend the analysis to **controversial topics**, based on Wikipedia's list of controversial topics.[10] For a selected topic, we identify subreddits with differing viewpoints on the topic, e.g., `r/Feminism` and `r/MensRights`.

Besides the topics covered in the literature, we extend our analysis to two additional topics (monarchy and values regarding communism vs capitalism). An overview of the results on all controversial topics can be found in Figure 2. An overview of closely related subreddits (in terms of values) to each of the subreddits discussed here can be found in Appendix C.

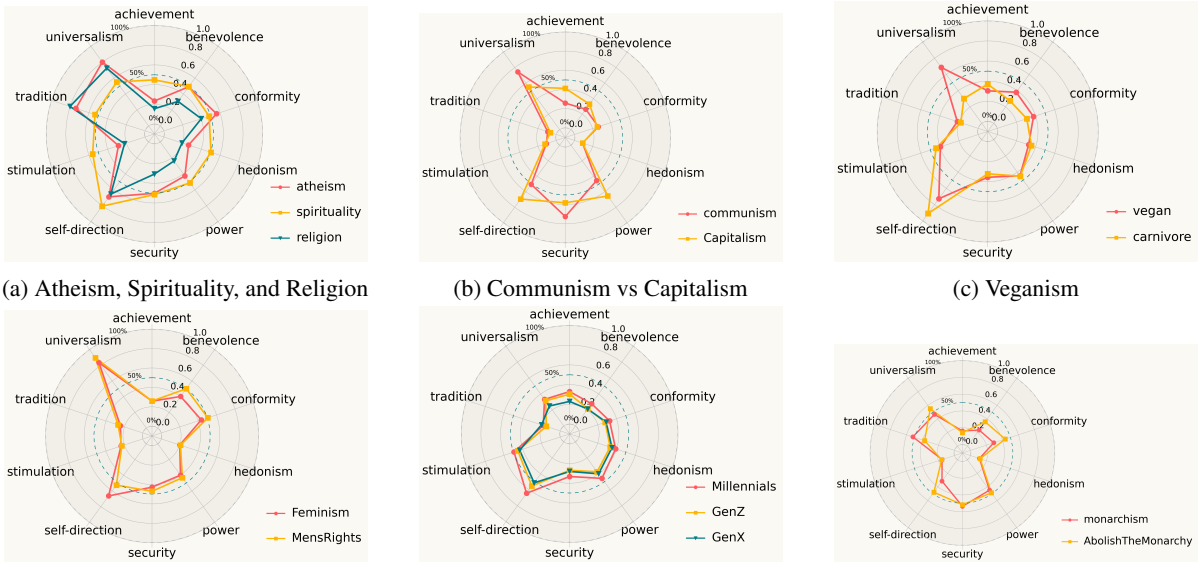
Feminism Prior work has examined the value systems of feminists and non-feminists (Zucker and Bay-Cheng, 2010). Significant differences emerged in the values *universalism*, *conformity*, and *tradition*. Further, feminists demonstrated significantly higher scores in the value *self-direction* than non-feminists. In our investigation involving the Reddit communities of `r/Feminism` and `r/MensRights` (assumed to be comparable to the nonfeminist group in the previous study), we also find the highest difference between the two subredits in the *self-direction* value, see Figure 2d.

Religion The topic of values and religiousness has been studied across multiple large-scale studies using Schwartz values, aiming to establish the connection between religion and values. Saroglou et al.

[8] We used `lingua` (`https://github.com/pemistahl/lingua-py`) to detect the language of each post.

[9] We manually scanned a sample of the translations and deemed them as having satisfactory quality.

[10] `https://en.wikipedia.org/wiki/Wikipedia:List_of_controversial_issues`



(a) Atheism, Spirituality, and Religion (b) Communism vs Capitalism (c) Veganism

(d) Feminism (e) Generations (f) Monarchism

Figure 2: Radar plots displaying the ten Schwartz values of subreddits dedicated to controversial topics.

(2004) present a meta-analysis across 15 countries, finding that religious people prefer conservative values, e.g., *tradition*, and dislike values related to openness (*stimulation*, *self-direction*). We see similar trends in our comparison of the subreddits `r/atheism`, `r/spirtuality`, and `r/religion` (Figure 2a). Here, we compare the contributors to the subreddit `r/religion` to religious people in previous studies as those find values to be consistent across different religious denominations (Christians, Jews, and Muslims). Similarly, lower scores for *power*, *hedonism*, and *achievement* align with previous findings.

Veganism In our comparison of the values for the subreddits `r/vegan` and `r/carnivore`, the largest difference can be observed in the value of *universalism*, and *self-direction*. Holler et al. (2021) review studies related to the values of vegetarians/vegans and omnivores, and find that vegetarians were found to have a stronger relation to *universalism*, however, they also found them to have a greater emphasis on *self-direction*, which does not align with our findings for the Reddit communities in Figure 2c. We attribute this misalignment to the difference between omnivores and carnivores conceptually, especially in the context of Reddit, where people might be more vocal in expressing their opinions. I.e., the subreddit might represent a more anti-veganism viewpoint than the omnivores selected for the studies cited.

Generations Lyons et al. (2007) study the differences in values across generations. Similarly, we compare the subreddit communities `r/Millenials`, `r/GenZ`, and `r/GenX` in Figure 2e. While the values of the different generations' subreddits are highly aligned, small differences can be observed, such as the higher scores in self-direction for Millenials in the *self-direction* value.

Communism vs capitalism Using our proposed method, we can extend the analysis of community values to demographics not previously studied. In specific, we describe the difference in values for different economic ideologies, i.e., communism vs capitalism in Figure 2b. While Schwartz and Bardi (1997) describe the effect of communism on Eastern Europe (e.g., high importance to conservatism and hierarchy values), they did not include the values held by the people supporting the ideology on a theoretical level, as it is possible to observe on Reddit. Here, we find that contributors to the `r/communism` subreddit hold strong values of *universalism* and *security*, whereas contributors to the `r/Capitalism` subreddit hold higher values in *self-direction* and *power*.

Monarchism We further present the values of people supporting and opposing monarchism in Figure 2f, as another example of a controversial topic that has yet to be studied in terms of values. We find people contributing to the subreddit opposing monarchies, `r/AbolishTheMonarchy`, hold higher values in *universalism*, *benevolence*,

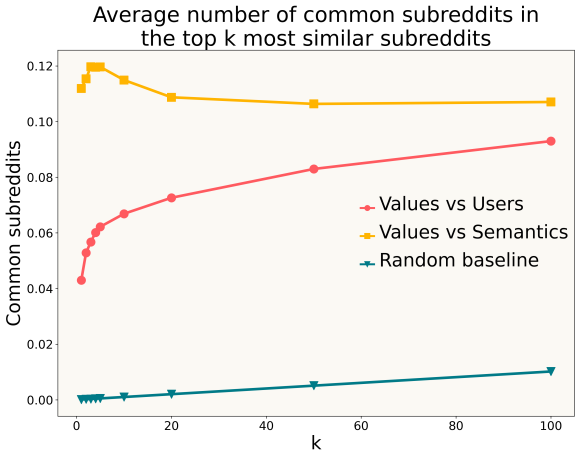


Figure 3: The average number of shared subreddits amongst the k most similar subreddits.

and *self-direction*, i.e., values that indicate progress and openness as values overall. The contributors to the `r/monarchism` subreddit value *tradition*, which is a core conservative value and aligns intuitively with support for monarchies.

4.4 Shared Community Values

In the previous sections, we made an implicit assumption: similar subreddit communities share similar values. We systematically investigate this intuition by defining three measures of similarity between subreddits:

Value similarly. We define $\sigma_{\text{val}}(\mathcal{S}_1, \mathcal{S}_2) = \cos(\boldsymbol{u}_{\mathcal{S}1}, \boldsymbol{u}_{\mathcal{S}2})$, that is, the cosine similarity between the value vectors of the subreddits.

Semantic similarity. We scrape from the page of each subreddit its **public description** (see Figure 5 in App. B), a natural language description of the subreddit. We then embed the descriptions using a sentence transformer[11] to construct a semantic embedding vector $\boldsymbol{e}_{\mathcal{S}}$ for each subreddit $\mathcal{S}$[12]. We now define $\sigma_{\text{sem}}(\mathcal{S}_1, \mathcal{S}_2) = \cos(\boldsymbol{e}_{\mathcal{S}1}, \boldsymbol{e}_{\mathcal{S}2})$.

Community similarity For each pair of subreddits, we define their community similarity to be the *overlap coefficient*[13] between the users of the two subreddits. That is,

$$\sigma_{\text{com}}(\mathcal{S}_1, \mathcal{S}_2) = \frac{|U(\mathcal{S}_1) \cap U(\mathcal{S}_2)|}{\min(|U(\mathcal{S}_1)|, |U(\mathcal{S}_2)|)}$$

Where $U(\mathcal{S})$ is the set of the subreddit's members.[14]

To answer the question of if similar subreddits share similar values, we correlate σ_{val} with σ_{sem} and σ_{com} as follows. First, for each subreddit $\mathcal{S}$ we find $\overline{\mathcal{S}}_{\text{sem}} = \text{argmax}_{\mathcal{S}'}\,[\sigma_{\text{sem}}(\mathcal{S}, \mathcal{S}')]$ and $\overline{\mathcal{S}}_{\text{com}} = \text{argmax}_{\mathcal{S}'}\,[\sigma_{\text{com}}(\mathcal{S}, \mathcal{S}')]$. If similar subreddits share a similar set of values, we should expect $\mathbb{E}_{\mathcal{S}}[\sigma_{\text{val}}(\mathcal{S}, \overline{\mathcal{S}}_{\text{sem}})]$ and $\mathbb{E}_{\mathcal{S}}[\sigma_{\text{val}}(\mathcal{S}, \overline{\mathcal{S}}_{\text{com}})]$ to be significantly larger than $\mathbb{E}\,[\sigma_{\text{val}}(\cdot, \cdot)]$. That is, the Schwartz values of two similar subreddits should be, in expectation, significantly closer to each other than the Schwartz values of two random subreddits. Our results, computed using empirical expectations over our dataset, confirm this hypothesis. Significant differences exist between the expected Schwartz values of similar and random subreddits. The expected values are both 0.560 for semantic and community similarity versus 0.009 for random subreddits. The z-test scores are 117.8 and 117.6, respectively, indicating a high level of statistical significance.

Next, we expand $\overline{\mathcal{S}}_{\text{sem}}$ by defining $\overline{\mathcal{S}}^k_{\text{sem}}$ to be the set of the k subreddits with the largest semantic similarity to $\mathcal{S}$. Equivalently, we define $\overline{\mathcal{S}}^k_{\text{val}}$ and $\overline{\mathcal{S}}^k_{\text{com}}$ in the same manner. For each $k \in [1...100]$, we calculate the overlap ratio of $\overline{\mathcal{S}}^k_{\text{sem}}$ and $\overline{\mathcal{S}}^k_{\text{val}}$, as well as $\overline{\mathcal{S}}^k_{\text{com}}$ and $\overline{\mathcal{S}}^k_{\text{val}}$. Figure 3 demonstrates that the level of overlap is significantly greater than a random baseline, again corroborating the hypothesis that similar subreddits tend to convey a similar set of values.

5 Conclusion

In this paper, we apply Schwartz's Theory of Human Values to Reddit at scale, thus circumventing the limitations of traditional survey methods and offering scholars a complementary tool for studying online communities. We train a supervised value extraction model and evaluate it to close existing gaps in the literature. We then conduct an extensive analysis of six million posts across 10,000 popular subreddits. Through our analysis, we both confirm previous findings from the social sciences, as well as uncover novel insights into the values expressed within various online communities, shedding light on existing patterns and outliers that warrant further investigation.

[11] `https://www.sbert.net/`. We use the `all-mpnet-base-v2` pretrained model.

[12] A limitation of this approach is that, occasionally, the public description is rather vague. However, from our experience, this rarely occurs.

[13] `https://en.wikipedia.org/wiki/Overlap_coefficient`

[14] As Reddit does not make this information publicly available, we estimate $U(\mathcal{S})$ by defining it as the set of users that had posted or commented in the subreddit within our dataset.

Limitations

The core limitations of our work stem from the value extraction model. While we have validated the labels produced by the model (§3.3), the model is not foolproof. Further, the training procedure of the value extraction model we adopted from van der Meer et al. (2023) adds some limitations. As outlined in §3.2, the model is trained to predict the *presence* of Schwartz values within a given text. However, it does not account for the text's *polarity* towards the value. Hence, the model is trained to put both positive and negative expressions of a value into the same category of being exhibitive of the value. An important direction for future work is incorporating the polarity of the value expression into the value extraction model.

Another limitation of our approach is the consolidation of the values of posts in a subreddit into a single vector. While this is necessary for understanding the values of communities, it discounts the values the individual posts of users that constitute these communities. This also opens up much room for future research, particularly in studying the internal dynamics of online communities and understanding the role that values play at the individual post level.

Lastly, while our approach can be leveraged to identify phenomena worthy of further investigation, it lacks the means to explain these observations fully. Values are an abstract concept that is challenging to quantify and analyse, with the values frameworks themselves drawing criticism (Jackson, 2020). We argue that interdisciplinary collaboration with experts in psychology and sociology is essential to understand these phenomena and their implications properly. Such collaboration will not only enrich our understanding of online communities but also contribute to developing more robust and nuanced machine learning models in the future, given the inclusion of such text in the training data.

Broader Impact and Ethical Considerations

We believe our work goes beyond the methodological toolkit of social scientists. The accessibility, large scale, and relatively high quality of Reddit data have positioned it as a valuable resource for training data for diverse NLP tasks (Overbay et al., 2023; Blombach et al., 2020; Huryn et al., 2022). It opens the door to training models that are better attuned to the specific needs of various communities, potentially protected ones. Recognizing the values of online communities, often utilized as training data for LM development, is crucial for making informed decisions about incorporating the data and contributing to the creation of models that genuinely reflect the diverse perspectives within these communities (Arora et al., 2023). These also percolate into downstream applications of LMs (Jakesch et al., 2023) and have a broader impact on their use.

As for ethical considerations, we made specific efforts to ensure that our work does not impair the privacy and anonymity of Reddit users (Sugiura et al., 2017). We refrain from attributing values to individual users and instead study communities as a whole by aggregating and condensing individual data points into a single vector. Nevertheless, many online communities were created to serve as a safe space for vulnerable individuals, where they share highly sensitive and private information. Therefore, research on Reddit and other online communities should make utmost efforts to respect these spaces and handle their data with care.

References

Pratik Agrawal, Tolga Buz, and Gerard de Melo. 2023. Wallstreetbets beyond gamestop, yolos, and the moon: The unique traits of reddit's finance communities. In *AMCIS 2022*.

Arnav Arora, Lucie-aimée Kaffee, and Isabelle Augenstein. 2023. Probing pre-trained language models for cross-cultural differences in values. In *Proceedings of the First Workshop on Cross-Cultural Considerations in NLP (C3NLP)*, pages 114–130, Dubrovnik, Croatia. Association for Computational Linguistics.

Jana Bašnáková, Ivan Brezina, and Radomír Masaryk. 2016. Dimensions of culture: The case of slovakia as an outlier in hofstede's research. *Ceskoslovenska Psychologie*, 60(1).

Andreas Blombach, Natalie Dykes, Philipp Heinrich, Besim Kabashi, and Thomas Proisl. 2020. A corpus of German Reddit exchanges (GeRedE). In *Proceedings of the 12th Language Resources and Evaluation Conference*, pages 6310–6316, Marseille, France. European Language Resources Association.

Ryan Boyd, Steven Wilson, James Pennebaker, Michal Kosinski, David Stillwell, and Rada Mihalcea. 2021. Values in words: Using language to evaluate and understand personal values. *Proceedings of the International AAAI Conference on Web and Social Media*, 9(1):31–40.

Pew Research Center. 2014. Political ideology by state.

Eshwar Chandrasekharan, Shagun Jhaver, Amy Bruckman, and Eric Gilbert. 2022. Quarantined! examining the effects of a community-wide moderation intervention on reddit. *ACM Trans. Comput.-Hum. Interact.*, 29(4).

Johnny R. J. Fontaine, Ype H. Poortinga, Luc Delbeke, and Shalom H. Schwartz. 2008. Structural equivalence of the values domain across cultures: Distinguishing sampling fluctuations from meaningful variation. *Journal of Cross-Cultural Psychology*, 39(4):345–365.

Philipp Gerlach and Kimmo Eriksson. 2021. Measuring cultural dimensions: external validity and internal consistency of hofstede's vsm 2013 scales. *Frontiers in Psychology*, 12:662604.

Matej Gjurković and Jan Šnajder. 2018. Reddit: A gold mine for personality prediction. In *Proceedings of the Second Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media*, pages 87–97, New Orleans, Louisiana, USA. Association for Computational Linguistics.

Jesse Graham, Jonathan Haidt, Sena Koleva, Matt Motyl, Ravi Iyer, Sean P. Wojcik, and Peter H. Ditto. 2013. Chapter two - moral foundations theory: The pragmatic validity of moral pluralism. In Patricia Devine and Ashby Plant, editors, *Advances in Experimental Social Psychology*, volume 47 of *Advances in Experimental Social Psychology*, pages 55–130. Academic Press.

Pengcheng He, Jianfeng Gao, and Weizhu Chen. 2021. Debertav3: Improving deberta using electra-style pretraining with gradient-disentangled embedding sharing.

Geert Hofstede. 1984. *Culture's consequences: International differences in work-related values*, volume 5. sage.

Sophie Holler, Holger Cramer, Daniela Liebscher, Michael Jeitler, Dania Schumann, Vijayendra Murthy, Andreas Michalsen, and Christian S Kessler. 2021. Differences between omnivores and vegetarians in personality profiles, values, and empathy: a systematic review. *Frontiers in psychology*, 12:579700.

Daniil Huryn, William M. Hutsell, and Jinho D. Choi. 2022. Automatic generation of large-scale multi-turn dialogues from Reddit. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 3360–3373, Gyeongju, Republic of Korea. International Committee on Computational Linguistics.

Terence Jackson. 2020. The legacy of geert hofstede.

Alexander Jaffe. 2009. Stance: Sociolinguistic perspectives. In *Stance: Sociolinguistic Perspectives*.

Maurice Jakesch, Advait Bhat, Daniel Buschek, Lior Zalmanson, and Mor Naaman. 2023. Co-writing with opinionated language models affects users' views. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, New York, NY, USA. Association for Computing Machinery.

Johannes Kiesel, Milad Alshomary, Nicolas Handke, Xiaoni Cai, Henning Wachsmuth, and Benno Stein. 2022. Identifying the human values behind arguments. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 4459–4471, Dublin, Ireland. Association for Computational Linguistics.

Sean T Lyons, Linda Duxbury, and Christopher Higgins. 2007. An empirical assessment of generational differences in basic human values. *Psychological reports*, 101(2):339–352.

Ercan G Nasif, Hamad Al-Daeaj, Bahman Ebrahimi, and Mary S Thibodeaux. 1991. Methodological problems in cross-cultural research: An updated review. *MIR: Management International Review*, pages 79–91.

Edward Newell, David Jurgens, Haji Saleem, Hardik Vala, Jad Sassine, Caitrin Armstrong, and Derek Ruths. 2016. User migration in online social networks: A case study on reddit during a period of community unrest. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 10, pages 279–288.

Bonny Norton. 1997. Language, identity, and the ownership of english. *TESOL Quarterly*, 31(3):409–429.

Keighley Overbay, Jaewoo Ahn, Fatemeh Pesaran zadeh, Joonsuk Park, and Gunhee Kim. 2023. mRedditSum: A multimodal abstractive summarization dataset of Reddit threads with images. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 4117–4132, Singapore. Association for Computational Linguistics.

Vladimir Ponizovskiy, Murat Ardag, Lusine Grigoryan, Ryan Boyd, Henrik Dobewall, and Peter Holtz. 2020. Development and validation of the personal values dictionary: A theory-driven tool for investigating references to basic human values in text. *European Journal of Personality*, 34(5):885–902.

Liang Qiu, Yizhou Zhao, Jinchao Li, Pan Lu, Baolin Peng, Jianfeng Gao, and Song-Chun Zhu. 2022. Valuenet: A new dataset for human value driven dialogue system. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 11183–11191.

Milton Rokeach. 1973. *The Nature of Human Values.* The Nature of Human Values. Free Press, New York, NY, US.

Shamik Roy, Maria Leonor Pacheco, and Dan Goldwasser. 2021. Identifying morality frames in political tweets using relational learning. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 9939–9958, Online

and Punta Cana, Dominican Republic. Association for Computational Linguistics.

Vassilis Saroglou, Vanessa Delpierre, and Rebecca Dernelle. 2004. Values and religiosity: A meta-analysis of studies using schwartz's model. *Personality and individual differences*, 37(4):721–734.

Shalom H. Schwartz. 1994. Are there universal aspects in the structure and contents of human values? *Journal of Social Issues*, 50(4):19–45.

Shalom H. Schwartz. 2012. An overview of the schwartz theory of basic values. *Online Readings in Psychology and Culture*, 2:11.

Shalom H Schwartz and Anat Bardi. 1997. Influences of adaptation to communist rule on value priorities in eastern europe. *Political psychology*, 18(2):385–410.

Shalom H Schwartz and Jan Cieciuch. 2022. Measuring the refined theory of individual values in 49 cultural groups: psychometrics of the revised portrait value questionnaire. *Assessment*, 29(5):1005–1019.

Ahmed Soliman, Jan Hafer, and Florian Lemmerich. 2019. A characterization of political communities on reddit. In *Proceedings of the 30th ACM Conference on Hypertext and Social Media*, HT '19, page 259–263, New York, NY, USA. Association for Computing Machinery.

Lisa Sugiura, Rosemary Wiles, and Catherine Pope. 2017. Ethical challenges in online research: Public/private perceptions. *Research Ethics*, 13(3-4):184–199.

Jackson Trager, Alireza S. Ziabari, Aida Mostafazadeh Davani, Preni Golazizian, Farzan Karimi-Malekabadi, Ali Omrani, Zhihe Li, Brendan Kennedy, Nils Karl Reimer, Melissa Reyes, Kelsey Cheng, Mellow Wei, Christina Merrifield, Arta Khosravi, Evans Alvarez, and Morteza Dehghani. 2022. The moral foundations reddit corpus.

Elsbeth Turcan and Kathy McKeown. 2019. Dreaddit: A Reddit dataset for stress analysis in social media. In *Proceedings of the Tenth International Workshop on Health Text Mining and Information Analysis (LOUHI 2019)*, pages 97–107, Hong Kong. Association for Computational Linguistics.

Michiel van der Meer, Piek Vossen, Catholijn Jonker, and Pradeep Murukannaiah. 2023. Do differences in values influence disagreements in online discussions? In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 15986–16008, Singapore. Association for Computational Linguistics.

Galen Weld, Amy X. Zhang, and Tim Althoff. 2023. Making online communities 'better': A taxonomy of community values on reddit.

Jonathan Zomick, Sarah Ita Levitan, and Mark Serper. 2019. Linguistic analysis of schizophrenia in Reddit posts. In *Proceedings of the Sixth Workshop on Computational Linguistics and Clinical Psychology*, pages 74–83, Minneapolis, Minnesota. Association for Computational Linguistics.

Alyssa N Zucker and Laina Y Bay-Cheng. 2010. Minding the gap between feminist identity and attitudes: The behavioral and ideological divide between feminists and non-labelers. *Journal of personality*, 78(6):1895–1924.

A Annotation Guidelines

```
For each row in the table, start by
reading the sentence in column A.
Next, review the Schwartz value
listed in column B and ensure you
understand its meaning, using the
figure at the top of the table as a
reference. Then, determine whether
the sentence expresses the given
value. It is important to note that
our focus is solely on whether the
value is expressed, not on how it
is expressed. This means that even
if the sentence violates the value,
it still expresses it. For
instance, the sentence ``I never
went to church'' expresses the
value of ``Tradition,'' despite
contradicting it.


This task is subjective, and in
many cases, there is no single
correct answer. If you are
uncertain about a particular
instance, respond to the best of
your ability.
```

Listing 1: Guidelines for the task of annotating the dataset used to evaluate the Value Extractor model.

As described in §3.3, Three authors of this paper annotated the evaluation set. Specifically, for each sample (s_i, v_i), where s_i is a sentence, and v_i is a Schwartz value, the annotators were tasked with determining whether s_i exhibits the value v_i. Figure 4 contains a print screen of the annotation task, whereas Listing 1 specifies the annotation guidelines.

B Additional Material

```
China, Estonia, Portugal, Colombia,
Italy, Finland, USA, Philippines,
Turkey, Serbia, Slovakia, Czech
Republic, South Africa, Peru,
Poland, Spain, Germany, France,
Australia, Costa Rica, Romania,
India, New Zealand, Greece,
Croatia, Brazil, Iceland, Sweden,
United Kingdom, Indonesia,
```

Listing 2: List of countries used in the correlation experiment in §4.2.

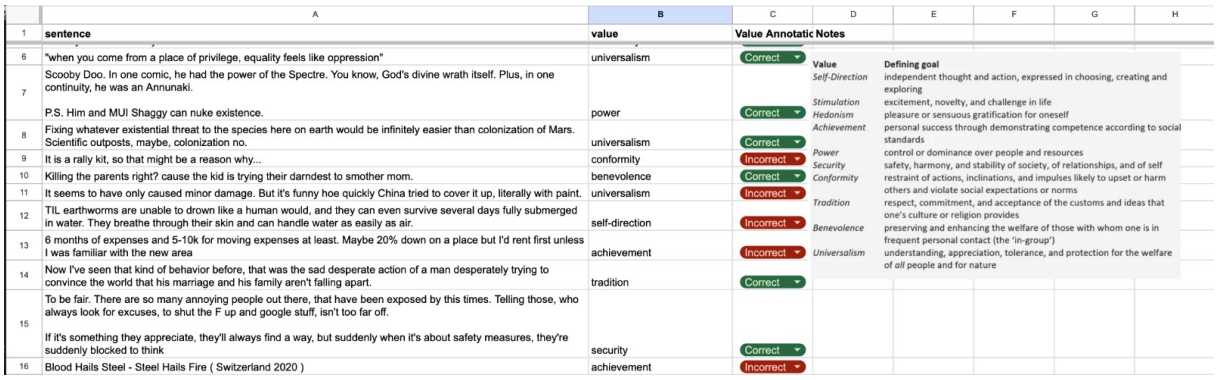
C Controversial Topics – Extended

Additionally to the analysis presented in Section 4.3, Table 10 displays the ten most similar subreddits in terms of values. We use value similarity as described in Section 4.4 to determine the closest subreddits to each of the controversial topic subreddits.

D Reproducibility

D.1 Training the Schwartz Values Extractor

We trained the Value extractor model for ten epochs with early stopping, using a learning rate of $5 \cdot 1e^{-5}$, batch size of 32, Adamw optimiser with default parameters and linear learning rate scheduler. We trained the model on a single TitanRTX GPU for 80 minutes.



	A	B	C	D	E	F	G	H
1	sentence	value	Value Annotatic	Notes				
6	"when you come from a place of privilege, equality feels like oppression"	universalism	Correct					
7	Scooby Doo. In one comic, he had the power of the Spectre. You know, God's divine wrath itself. Plus, in one continuity, he was an Annunaki. P.S. Him and MUI Shaggy can nuke existence.	power	Correct					
8	Fixing whatever existential threat to the species here on earth would be infinitely easier than colonization of Mars. Scientific outposts, maybe, colonization no.	universalism	Correct					
9	It is a rally kit, so that might be a reason why...	conformity	Incorrect					
10	Killing the parents right? cause the kid is trying their darndest to smother mom.	benevolence	Correct					
11	It seems to have only caused minor damage. But it's funny hoe quickly China tried to cover it up, literally with paint.	universalism	Incorrect					
12	TIL earthworms are unable to drown like a human would, and they can even survive several days fully submerged in water. They breathe through their skin and can handle water as easily as air.	self-direction	Incorrect					
13	6 months of expenses and 5-10k for moving expenses at least. Maybe 20% down on a place but I'd rent first unless I was familiar with the new area	achievement	Incorrect					
14	Now I've seen that kind of behavior before, that was the sad desperate action of a man desperately trying to convince the world that his marriage and his family aren't falling apart.	tradition	Correct					
15	To be fair. There are so many annoying people out there, that have been exposed by this times. Telling those, who always look for excuses, to shut the F up and google stuff, isn't too far off. If it's something they appreciate, they'll always find a way, but suddenly when it's about safety measures, they're suddenly blocked to think	security	Correct					
16	Blood Hails Steel - Steel Hails Fire (Switzerland 2020)	achievement	Incorrect					

Value	Defining goal
Self-Direction	independent thought and action, expressed in choosing, creating and exploring
Stimulation	excitement, novelty, and challenge in life
Hedonism	pleasure or sensuous gratification for oneself
Achievement	personal success through demonstrating competence according to social standards
Power	control or dominance over people and resources
Security	safety, harmony, and stability of society, of relationships, and of self
Conformity	restraint of actions, inclinations, and impulses likely to upset or harm others and violate social expectations or norms
Tradition	respect, commitment, and acceptance of the customs and ideas that one's culture or religion provides
Benevolence	preserving and enhancing the welfare of those with whom one is in frequent personal contact (the 'in-group')
Universalism	understanding, appreciation, tolerance, and protection for the welfare of *all* people and for nature

Figure 4: Print screen of the annotation task. The three annotators were tasked with determining whether the sentence in column A expresses the Schwartz value in column B.

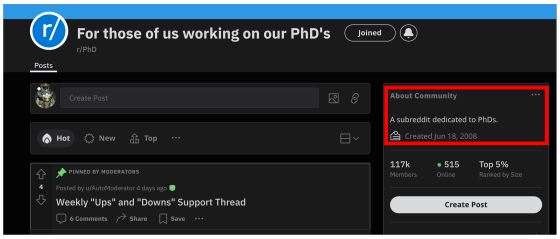


Figure 5: Location of the "Public description" attribute on a subreddit page.

Country	Language	Subreddits
Australia	English	r/australia, r/Australia_, r/australian, r/sydney, r/melbourne
Brazil	Portuguese	r/brasil, r/brasilivre, r/GTAorBrazil, r/brasilia, r/saopaulo
China	Chinese	r/China, r/China_irl, r/real_China_irl
Colombia	Spanish	r/Colombia, r/ColombiaReddit, r/Bogota
Costa Rica	Spanish	r/costarica, r/Ticos
Croatia	Croatian	r/croatia, r/zagreb
Czech Republic	Czech	r/czech
Ecuador	Spanish	r/ecuador
Estonia	Estonian	r/Eesti
Faroe Islands	All	r/Faroeislands
Finland	Finnish	r/Suomi
France	French	r/france, r/paris
Georgia	Georgian	r/Sakartvelo, r/tbilisi
Germany	German	r/de, r/deuchland, r/Munich, r/berlin
Ghana	English	r/ghana
Greece	Greek	r/greece
Hong Kong	Chinese	r/HongKong
Iceland	Icelandic	r/Iceland
India	All	r/india, r/IndiaSpeaks, r/unitedstatesofindia, r/askindia
Indonesia	Indonesian	r/indonesia
Israel	Hebrew	r/Israel, r/telaviv, r/jerusalem
Italy	Italian	r/italy, r/Italia
Japan	Japanese	r/japan, r/tokyo, r/newsokur
New Zealand	English	r/newzealand, r/auckland
Oman	Arabic	r/Oman
Peru	Spanish	r/PERU
Philippines	English	r/Philippines
Poland	Polish	r/Polska
Portugal	Portuguese	r/portugal, r/lisboa, r/porto
Romania	Romanian	r/Romania, r/bucuresti, r/iasi, r/cluj, r/timisoara
Serbia	Serbian	r/serbia
Slovakia	Slovak	r/Slovakia
South Africa	All	r/southafrica, r/capetown, r/johannesburg
South Korea	Korean	r/hanguk
Spain	Spanish	r/Espana, r/spain, r/Madrid, r/Barcelona
Sweden	Swedish	r/Sverige, r/sweden, r/stockholm, r/svenskpolitik
Turkey	Turkish	r/Turkey, r/istanbul, r/ankara, r/Ismir
Ukraine	Ukrainian	r/ukraine, r/Ukraine_UA
United Kingdom	All	r/unitedkingdom, r/london, r/manchester, r/CasualUK, r/unitedkingdom, r/askuk
USA	All	All states subreddits
Vietnam	Vietnamese	r/VietNam

Table 5: Countries, languages and subreddits used in the correlation study.

Value	Subreddits
achievement	r/GED, r/CompTIA, r/AWSCertifications, r/EngineeringResumes, r/study, r/AzureCertification, r/GetStudying, r/cissp, r/findapath, r/learnprogramming, r/WGU_CompSci, r/pmp, r/resumes, r/ADHD_Programmers, r/dataanalysis, r/language_exchange, r/PassNclex, r/PhD, r/C25K, r/technicalwriting
benevolence	r/TIHI, r/AITAH, r/AmItheAsshole, r/AmItheButtface, r/cheating_stories, r/motherinlawsfromhell, r/regretfulparents, r/JUSTNOFAMILY, r/SingleParents, r/CaregiverSupport, r/Fosterparents, r/inlaws, r/amiwrong, r/Infidelity, r/family, r/AmITheDevil, r/Custody, r/stepparents, r/JUSTNOMIL, r/JustNoSO
conformity	r/TIHI, r/AITAH, r/AmItheAsshole, r/AmItheButtface, r/Anger, r/spiderversedailymemes, r/cheating_stories, r/AmITheDevil, r/motherinlawsfromhell, r/HistoryAnimemes, r/amiwrong, r/Infidelity, r/confession, r/toxicparents, r/pettyrevenge, r/JUSTNOFAMILY, r/entitledparents, r/inlaws, r/CPTSDFightMode, r/JUSTNOMIL
hedonism	r/gratitude, r/TIHI, r/happy, r/DeadBedrooms, r/happyrelationships, r/TwoXSex, r/sex, r/depressed, r/Vancouver4Friends, r/SEXONDRUGS, r/drivinganxiety, r/depression, r/loneliness, r/lonely, r/love, r/patientgamers, r/CongratsLikeImFive, r/sad, r/StopGaming, r/HLCommunity
power	r/TIHI, r/FinancialPlanning, r/sidehustle, r/GamblingAddiction, r/PFJerk, r/DaveRamsey, r/personalfinance, r/Fire, r/CRedit, r/UKPersonalFinance, r/freelance, r/eupersonalfinance, r/smallbusiness, r/irishpersonalfinance, r/gambling, r/FIREUK, r/problemgambling, r/ynab, r/dividends, r/shoppingaddiction
security	r/CredibleDefense, r/foreignpolicy, r/syriancivilwar, r/EndlessWar, r/AfghanConflict, r/ww3, r/warinukraine, r/CombatFootage, r/geopolitics, r/GunsAreCool, r/TIHI, r/war, r/policeporn, r/UkraineConflict, r/LessCredibleDefence, r/UkrainianConflict, r/UkraineRussiaReport, r/UkraineInvasionVideos, r/WarCollege, r/RussiaUkraineWar2022
self-direction	r/antidepressants, r/GERD, r/prozac, r/depressionregimens, r/AnorexiaRecovery, r/Effexor, r/sugarfree, r/panicdisorder, r/intuitiveeating, r/decaf, r/Gastritis, r/Asthma, r/EatingDisorders, r/Hashimotos, r/insomnia, r/Gastroparesis, r/lexapro, r/Hypothyroidism, r/Allergies, r/fuckeatingdisorders
stimulation	r/TIHI, r/gratitude, r/drivinganxiety, r/Vancouver4Friends, r/CongratsLikeImFive, r/C25K, r/happy, r/depressed, r/depression, r/loneliness, r/lonely, r/penpals, r/ShouldIbuythisgame, r/toastme, r/TeensMeetTeens, r/socialanxiety, r/sad, r/MakeNewFriendsHere, r/Needafriend, r/FreeCompliments
tradition	r/PrayerTeam_amen, r/Christian, r/PrayerRequests, r/Christians, r/TrueChristian, r/OrthodoxChristianity, r/Christianity, r/Bible, r/Episcopalian, r/AskAPriest, r/Catholicism, r/NoFapChristians, r/Anglicanism, r/OpenChristian, r/TIHI, r/GayChristians, r/Christianmarriage, r/exchristian, r/SaintRampalJi, r/RadicalChristianity
universalism	r/mixedrace, r/AskFeminists, r/IsraelPalestine, r/IsraelCrimes, r/LeftWingMaleAdvocates, r/israelexposed, r/megafaunarewilding, r/conservation, r/LGBTnews, r/AgainstHateSubreddits, r/evolution, r/Palestine, r/MensRights, r/TIHI, r/aznidentity, r/ChristianUniversalism, r/InsaneProtestors, r/Indigenous, r/PurplePillDebate, r/transgender

Table 6: Subreddits with the highest signal for each one of the ten Schwartz values.

Value	Subreddits
Achievement	r/study, r/resumes, r/PhD
Benevolence	r/AmItheAsshole, r/motherinlawsfromhell, r/family
Conformity	r/Anger, r/confession, r/pettyrevenge,
Hedonism	r/happyrelationships, r/sex, r/depression
Power	r/FinancialPlanning, r/smallbusiness, r/GamblingAddiction
Security	r/foreignpolicy, r/warinukraine, r/CombatFootage
Self-direction	r/antidepressants r/AnorexiaRecovery, r/sugarfree, r/Allergies
Stimulation	r/ShouldIbuythisgame, r/socialanxiety, r/MakeNewFriendsHere, r/FreeCompliments
Tradition	r/PrayerRequests, r/Christians, r/Anglicanism
Universalism	r/mixedrace, r/IsraelPalestine, r/conservation, r/LGBTnews

Table 7: Subreddits with the highest signal for each of the ten Schwartz values.

Magnitude	Subreddits
Maximal	r/TIHI, r/gratitude, r/AITAH, r/AmItheAsshole, r/AmItheButtface, r/Anger, r/depressed, r/JUSTNOFAMILY, r/SuicideWatch, r/depression, r/JustNoSO, r/toxicparents, r/HistoryAnimemes, r/motherinlawsfromhell, r/CPTSDFightMode, r/drivinganxiety, r/confessions, r/venting, r/amiwrong, r/confession
Minimal	r/itookapicture, r/whatisthiscar, r/CityPorn, r/MoissaniteBST, r/identifythisfont, r/Whatisthis, r/CulinaryPlating, r/ShroomID, r/OldSchoolCelebs, r/ArtPorn, r/analog, r/whatsthisrock, r/whatanime, r/ExposurePorn, r/EarthPorn, r/mushroomID, r/whatisthisthing, r/Moviesinthemaking, r/badscificovers, r/portraits

Table 8: Subreddits with the highest and lowest total magnitude of Schwartz values, calculated according to $mag(\mathcal{S}) = |\boldsymbol{u}_{\mathcal{S}}|_2$.

Statement	Value
Has anyone else been seeing an uptick in patients who say they can't afford to buy food?	*Self-direction*
"Move to a lower cost of living area. Increase your income. Househack."	*Power*
Lenzilumab significantly improved survival without invasive mechanical ventilation in hospitalised patients with COVID-19, with a safety profile similar to that of placebo.	*Self-direction*
Plus, that shit is going to bankrupt us with the increased health costs related to our obesity epidemic.	*Self-direction*
Even low levels of air pollution can damage health, study finds	*Self-direction*
i hate those strange guys who peek into my phone from the back	*Conformity*

Table 9: Examples of posts that were misclassified by the Value Extractor model.

subreddit	closest subreddits
r/vegan	r/intersex, r/transgenderUK, r/TransSpace, r/Transmedical, r/ABCDesis, r/LGBTindia, r/DeepThoughts, r/nonduality, r/exvegans, r/BlockedAndReported
r/carnivore	r/carnivorediet, r/Psoriasis, r/PsoriaticArthritis, r/rheumatoid, r/acne, r/Hashimotos, r/Testosterone, r/PlasticSurgery, r/lupus, r/kratom
r/communism	r/InformedTankie, r/CommunismWorldwide, r/CPUSA, r/socialism, r/StupidpolEurope, r/LatinAmerica, r/sendinthetanks, r/ROI, r/Africa, r/myanmar
r/Capitalism	r/georgism, r/Unions, r/SandersForPresident, r/labor, r/NewDealAmerica, r/BernieSanders, r/WorkersStrikeBack, r/ndp, r/theydidthemath, r/MayDayStrike
r/monarchism	r/leftistvexillology, r/UsefulCharts, r/RoughRomanMemes, r/IndiaPlace, r/pureasoiaf, r/MedievalHistory, r/shittyhalolore, r/MemriTVmemes, r/ancientrome, r/darkwingsdankmemes
r/AbolishTheMonarchy	r/Sham_Sharma_Show, r/forwardsfromgrandma, r/WordAvalanches, r/Malaphors, r/Lal_Salaam, r/Metal, r/BanVideoGames, r/QanonKaren, r/PoliticalHumor, r/insanepeoplefacebook
r/Millennials	r/GenZ, r/PlusSizedAndPregnant, r/ThailandTourism, r/indiasocial, r/Liverpool, r/TallGirls, r/Psychedelics, r/Chandigarh, r/Cardiff, r/ChronicIllness
r/GenZ	r/Millennials, r/indiasocial, r/TallGirls, r/Kibbe, r/Zillennials, r/aggretsuko, r/precure, r/nerdfighters, r/KoeNoKatachi, r/comiccon
r/GenX	r/lesbianfashionadvice, r/CosplayNation, r/Xennials, r/bigboobproblems, r/femininebоys, r/PlusSize, r/distantsocializing, r/CasualPH, r/bald, r/TallGirls
r/Feminism	r/antiwoke, r/MensLib, r/ControversialOpinions, r/EnoughIDWspam, r/prochoice, r/prolife, r/fourthwavewomen, r/TransSpace, r/TrueUnpopularOpinion, r/AntiVegan
r/MensRights	r/antifeminists, r/AntiHateCommunities, r/AsABlackMan, r/LeftWingMaleAdvocates, r/aznidentity, r/EnoughPCMSpam, r/FragileWhiteRedditor, r/BlatantMisogyny, r/TheLeftCantMeme, r/ForwardsFromKlandma
r/atheism	r/exmuslim, r/antitheistcheesecake, r/extomatoes, r/Antitheism, r/atheismindia, r/mormon, r/progressive_islam, r/Judaism, r/RadicalChristianity, r/religiousfruitcake
r/spirituality	r/SpiritualAwakening, r/awakened, r/pureretention, r/Stoicism, r/Shamanism, r/starseeds, r/Mediums, r/Soulnexus, r/Semenretention, r/SASSWitches
r/religion	r/AskAChristian, r/ChristianUniversalism, r/RadicalChristianity, r/bahai, r/hinduism, r/shia, r/Bible, r/progressive_islam, r/extomatoes, r/mormon

Table 10: For each subreddit in the controversial topics analysis, the 10 most similar subreddits in terms of values.