Received 24 April 2024, accepted 3 July 2024, date of publication 18 July 2024, date of current version 2 August 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3430495

RESEARCH ARTICLE

From Polarization to Pro-Sociality: Measuring Beneficence in Controversial Online Conversations

MATTEO CINELLI[1], (Member, IEEE), GABRIELE ETTA[1], (Member, IEEE), VALENTINA FRANZONI[2,3], (Senior Member, IEEE), MAURIZIO MANCINI[1], AND RADOSLAW NIEWIADOMSKI[4]

[1]Department of Computer Science, Sapienza University of Rome, 00185 Rome, Italy
[2]Department of Mathematics and Computer Science, University of Perugia, 06123 Perugia, Italy
[3]Department of Computer Science, Hong Kong Baptist University, Hong Kong
[4]Department of Computer Science, Bioengineering, Robotics and Systems Engineering (DIBRIS), University of Genova, 16145 Genova, Italy

Corresponding author: Valentina Franzoni (valentina.franzoni@unipg.it)

This work was supported by the EmoRe Laboratory on Affective Computing, Department of Computer Science, University of Perugia, Italy.

ABSTRACT This study presents a novel computational approach to quantifying beneficence, defined as a pro-social attitude that positively influences others, in the polarized context of online debates on controversial topics. Starting from a dataset of conversations on Facebook pages on controversial and polarized topics such as vaccination, we used semantic proximity measures to analyze the linguistic landscape, such as confidence, normalized Google distance, and pointwise mutual information. We built an undirected weighted co-occurrence network in which two users are connected if they both comment on the same post. We analyzed polarization trends toward the semantics of beneficence from the point of view of comments, users, and the neighborhood. We found that the formation of echo chambers on the vaccination topic did not correspond to echo chambers on beneficence, with both groups of pro-vax and anti-vax users exhibiting similar levels of beneficence in their discourse. These findings highlight the challenges of bridging the communicative gaps in communities around controversial topics that form echo chambers, showing that opposing parties can share similar beneficence levels. Future research should explore the dynamics of opinion formation and the role of beneficence in preventing and managing hate speeches.

INDEX TERMS Semantic proximity, attitude polarization, echo chambers, polarization, hate speech, affective computing.

I. INTRODUCTION

Beneficence, from Latin *beneficentia* (*bene-facere*, i.e., acting favorably), means "kindness, generosity," and is commonly understood as doing something beneficial for others. Thus, beneficence is a pro-social act that may imply empathetic action. Beneficence actions include pro-social behaviors such as charity, volunteering (including virtual/digital/e-volunteering), donations, and taking care of others [13].

Beneficence has been extensively studied in the context of basic psychological needs. Recently, Martela and Ryan [13], [14] stressed the role of the subjective experience of beneficence. According to them, beneficence is "the sense of having a positive impact in the lives of other people" [13], and "a subjective feeling or evaluation about the [...] personal sense of having done good things to others" [14]. Beneficence is related to psychological well-being [14]. Some authors [13] include it among the basic psychological needs, next to Autonomy, Competence, and Relatedness [18], [21]. Simultaneously, it is related to positive emotions and attitudes, including compassion, satisfaction, kindness, and empathy. Beneficence can contribute to positive affect and vitality [14], life satisfaction [13], and meaningful work [12].

Despite its importance for wellness and well-being, beneficence has rarely been computationally addressed. To the best

The associate editor coordinating the review of this manuscript and approving it for publication was Hai Dong.

of our knowledge, before our preliminary work [11], of which this paper is an extension, there is no literature on non-verbal behaviors that might be specific to the sense (i.e., concept, meaning) of beneficence. Thus, proposing a computational approach to estimate beneficence is an open challenge that we addressed in this study.

Our approach offers a novel perspective for quantifying and analyzing beneficence, which can be applied to various domains beyond the specific context and topic of discussion. By exploring pro-social attitudes exhibited in social media interactions, our research contributes to a broader understanding of online discourse and its implications for social dynamics.

Specifically, in this work, we aim to study beneficence by analyzing social media users' comments on a specific topic that is significant to society but controversial. Focusing on social media activity seems a reasonable choice when considering Martela and Ryan definition of beneficence [13], [14], described as a "subjective feeling or evaluation." Online social networks are an excellent repository of personal opinions and beliefs in which individuals and groups (e.g., NGOs) can promote or perform voluntary initiatives and actions to influence or help others, including campaigns to promote a healthy lifestyle or online volunteering [9], [15]. Simultaneously, we can easily observe extremely polarized opinions on online social networks, where the opinion of a different party is considered not beneficial and not pro-social by the other party, resulting in hate speech [3].

In particular, in this respect, our goals are:

- (G1) to check whether social media users discussing controversial topics exhibit beneficence polarization (i.e., users have either very low or very high beneficence).
- (G2) to check whether social media users can be clustered in "echo chambers" for beneficence (see Section II: that is, they show beneficence similar to their neighborhood.

In this study, we used a dataset that included strongly polarized opinions on vaccination. The selection of this dataset is primarily motivated by the presence of distinct and polarized opinions rather than specifically focusing on vaccines as a specific topic, allowing for an examination of beneficence in online social media interactions across different ideological groups on topics related to ethics, which are more prone than others to show extreme affirmations (i.e., expressing strongly polarized opinions). Regarding the topic of vaccination, a prevalent argument shared by both proponents and opponents is the perceived impact on public health: advocates for vaccination believe that it enhances the health of both individuals and the community. Conversely, vaccine skeptics argue that vaccines do not benefit the community and may even harm those who are vaccinated; thus, vaccination will not enhance public health, and avoid herd immunity [16]. In this perspective, vaccination is a particularly relevant topic in the context of beneficence as a sense of having a positive impact on others' lives.

Consequently, we aimed to determine whether the polarization of social media users in terms of their opinions on vaccinations, can also be observed in their intent, that is, in the textual semantics of beneficence that we may detect in their posts and comments.

In terms of practical applications, the findings of this study could be useful to stakeholders involved in social dynamics, public health, policy-making, and community well-being.

By understanding the pro-social attitudes exhibited by users in social interactions through the semantics of beneficence, online platforms can develop strategies to promote positive interactions, mitigate toxic behavior, prevent hate speech, and foster a healthier online environment, which can help in designing targeted communication campaigns, addressing misinformation, and improving public health messaging strategies. The proposed methodology can serve as a foundation for further research in the fields of computational social science, sentiment analysis, and online discourse analysis to explore benefits in other domains, investigate the impact of online discussions on real-world outcomes, and contribute to a broader understanding of social dynamics in the digital age.

II. RELATED WORKS

Using web-based co-occurrence to identify semantic proximity is a well-known technique that is comparable in performance to semantics extracted from ontologies, with various advantages and limitations [5], [8], [10]. Semantic similarity can be used to evaluate closeness to the semantics of beneficence. A similar technique, with adaptations to vectors of words representing expressive terms in a sentence, has also been used in the context of affective computing and emotion recognition [7]; which assesses the emotional content of a sentence by combining information about each word in the sentence.

To identify controversial topics, "echo chambers," i.e., clusters of users having the same opinion about a given controversial topic, have been studied specifically in social media [6], rating and comparing users of different online social networks (e.g., Facebook, Reddit, Twitter) regarding their leaning towards controversial topics (e.g., gun control, Obamacare, abortion, vaccines). The results highlight that online users tend to form echo chambers, share information adhering to their worldviews, ignore dissenting information, and form polarized groups around shared narratives.

In addition, the correlation between personality traits and echo chambers has been studied [1]. The results show that personality traits are similarly distributed within polarized communities, except for the concept of "emotional stability," which is higher among users who support conspiracy narratives. Similar and significant correlations were found between personality traits within different echo chambers, indicating that the prevalent personality model was the same in both observed echo chambers (i.e., low extraversion, high emotional stability, low agreeableness, low conscientiousness, and high openness).

Regarding concepts that could be considered close to beneficence, a study [17] investigated the correlation between the inclination to be generous (i.e., donating to charity) and political orientation, finding a difference between users identifying political values in their polarized leaning to donate nationally or internationally.

In line with the studies outlined above, our research delves into the field of social media analysis by focusing on the content of users' comments. However, our approach differs in that we introduce a novel methodology using web-based semantic analysis of a user's social network [5], [8], [20], allowing us to assess the presence of particular concepts within textual data in an echo chamber, regardless of whether explicit terminology is used to denote them. Furthermore, to understand how benevolent behavior manifests in online interactions, our research uniquely focuses on examining beliefs in beneficence expressed in textual data.

III. DATA AND SOCIAL NETWORK

In this section, we describe the dataset and its network structure, which can be established from social interactions.

A. DATASET

The source dataset [19] was built using the Facebook Graph API. A targeted keyword search was conducted for the terms 'vaccine,' 'vaccines,' and 'vaccination' within the English language corpus, from January 1, 2010, to May 31, 2017. The inclusion criteria for Facebook pages required a minimum of ten English posts pertinent to the subject of vaccination.

The resulting dataset includes $\sim 300K$ posts by 243 pages, disseminated by $\sim$ $410K$ users. This dataset serves as a valuable resource for analyzing the dichotomy of perspectives on critical topics such as vaccination within social media networks. To facilitate a nuanced analysis, all Facebook pages were manually categorized into two predominant categories reflecting their stance on vaccination: pro-vax for those who were favorable to vaccines (145 pages) and anti-vax for those who were against them (98 pages). This classification is pivotal for understanding the narrative landscape of the vaccination discourse on various platforms.

Considering pages' narratives and posts' content, all Facebook pages were manually classified into two main groups: pro-vax (145) and anti-vax (98). A comprehensive breakdown of the dataset is presented in Table 1.

It is important to acknowledge the complexity of the vaccination debate as reflected in the dataset. The support or opposition to vaccines is not always absolute, and there can be varying degrees of agreement or disagreement within a group of users as highlighted by the 'echo chambers,' a phenomenon where users encounter information that reinforces their pre-existing views [19].

This dataset was chosen to provide a basis for exploring the nuances of opinion polarization within digital social ecosystems, rather than the specific topic of vaccination. The characteristics of this dataset provide a robust foundation for analyzing discourse dynamics, which is pertinent to our study's aim of measuring beneficence and its related pro-sociality in polarized communication.

The methods and insights derived from this analysis were designed to be topic-agnostic, ensuring broad applicability to various subjects that exhibit similar communication structures.

B. NETWORK GENERATION

From the source dataset $\mathcal{D}$, we start from the subsets $\mathcal{U} = \{u \in \mathbb{N} | u \text{ is a user in } \mathcal{D}\}$ and $\mathcal{P} = \{p \in \mathbb{N} | p \text{ is a post in } \mathcal{D}\}$, building a bipartite network $B = (\mathcal{U} \cup \mathcal{P}, \mathcal{L}) | l \in \mathcal{L}, l$ is a comment in $\mathcal{D}\}$, where the nodes are users and posts, and a link $l_k \in \mathcal{L}$ between user u_i and post p_j exists if user $u_i \in \mathcal{U}$ commented on post $p_j \in \mathcal{P}$. From the bipartite network B, by projecting onto $\mathcal{U}$ the pairs of links in B where two users commented on the same post, we build an undirected weighted co-occurrence network C, in which two users are connected if they both commented on the same post, as shown in the example in Figure 1. We will traverse the network C to identify polarization (i.e., leaning) toward the controversial topic, at a neighborhood level, and calculate the proximity to the concept of beneficence.

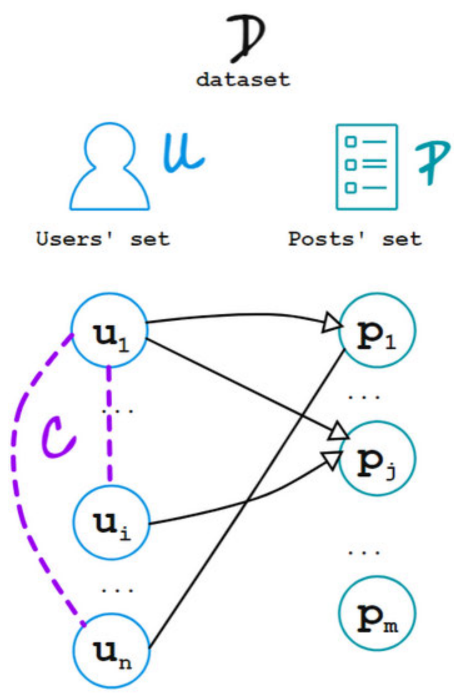

FIGURE 1. Network generation: a link between user u_i and post p_j exists if user u_i comments on p_j; projected users will be connected in the comment-weighted network C if they both commented on the same post.

C. USER LEANING

To measure the propensity of user i towards pro-vax or anti-vax content, we introduce a measure called 'user leaning' (i.e., polarization) x_i. Consider user i who expresses preferences by liking certain posts. Let $L_i = \{l_1, l_2, \ldots, l_{n_i}\}$ denote the set of likes generated by user i, where n_i denotes

TABLE 1. Data breakdown for the facebook dataset. The total vaccine-related pages were classified into pro-vax and anti-vax by [19].

Dataset	Time Span	Total Pages	Pro-vax Pages	Anti-vax Pages	Posts	Comments	Likes	Users
Facebook	1/1/2010 - 31/5/2017	243	145	98	153603	2095588	24155735	410062

the total number of likes. Each like l_j in this set is assigned a value from the set $\{-1, 1\}$, where -1 corresponds to a post from an anti-vax page and 1 corresponds to a post from a pro-vax page. This mapping is based on a source-based approach, with the assumption that the content of a page is a reliable indicator of its stance on vaccines.

A user's stance or leaning i, denoted by x_i, is quantified as the average value of likes, calculated as in Equation 1:

$$x_i = \frac{1}{n_i}\sum_{j=1}^{n_i} l_j \tag{1}$$

where x_i represents the average polarization of user i, with positive values indicating pro-vax leaning, negative values indicating anti-vax leaning, and zero indicating a neutral position or an equal number of likes for both leanings.

D. NEIGHBORHOOD LEANING

Similarly, we introduced the concept of 'neighborhood leaning' to quantify the collective tendency or orientation of a user's immediate social circle within a co-commenting network. Specifically, for a given user i, neighborhood leaning, denoted as x_i^N, is calculated as the mean value of the individual polarity of all users directly connected to user i. Equation 2 can be expressed as follows:

$$x_i^N = \frac{1}{k_i}\sum_{j=1}^{k_i} x_j \tag{2}$$

where k_i is the degree (i.e., number of connections) of user i within the unweighted co-commenting network C, and x_j corresponds to the leaning of the j^{th} neighbor connected to user i.

It is important to note that this calculation assumes an unweighted network structure, where all connections are considered equal in their contribution to the neighborhood leaning metric, as in the dataset.

This network captures the structure of social interactions and discussions, to study how users cluster around shared content of interest.

In particular, the dataset network shows that users leaning towards pro-vax or anti-vax content tend to have neighbors with the same leaning, forming echo chambers, as shown in Figure 2.

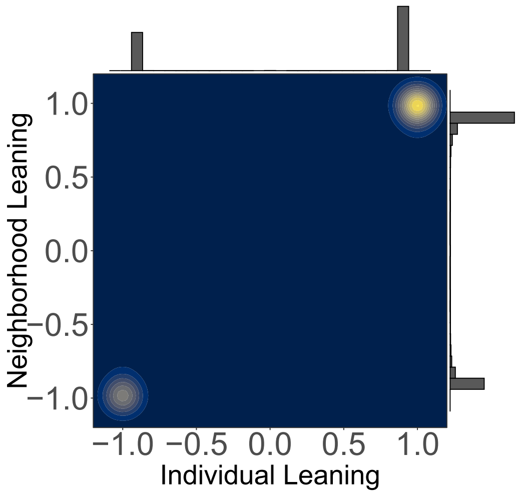

FIGURE 2. Joint distribution between the individual leaning and the neighborhood leaning towards anti-vax (represented by −1.0) or pro-vax (represented by 1.0). The light areas close to the extremes show the presence of distinct echo chambers.

IV. METHODOLOGY FOR BENEFICENCE EVALUATION

This section outlines our approach to quantifying the principle of beneficence in web content. To compute beneficence in online comments, we leverage well-known techniques of web-based similarity [5], [8], [20], extracting from a search engine the occurrence of words that users include in their comments, the co-occurrences of the term beneficence, and measuring their semantic proximity. The semantic proximity of posts to the concept of beneficence is computed by comparing three similarity measures: confidence [2], Normalized Google Distance (NGD) [5], and Pointwise Mutual Information (PMI) [4]. By aggregating proximity to beneficence of authors' comments, we obtained a proximity value for each user. We then compute the proximity distribution at the user's neighborhood level for each user. Finally, we look for local maxima of proximity distributions to check for polarization of beneficence in echo chambers in a dataset that already includes polarized echo chambers on other topics.

In the following subsections, we introduce specific semantic proximity measures to analyze the controversial dimensions of online user interactions. Details of specific settings and the use of these proximity measures for our specific case are provided in the experiments section.

A. WEB-BASED SEMANTICS

We collected data on terms' co-occurrence to calculate Web-based proximity measures, which have proven to be effective in the literature on any topic that is well represented on the Web in a given language [5], [8].

We used the Web as a knowledge base for semantic similarity, because of its comprehensive nature. The always-evolving collaborative objects of the Web also guarantee updated information [8]. This technique allows the incorporation of current and emerging concepts, trends, and terminologies, thereby enabling a more accurate semantic similarity assessment in real-time. However, the limit of web-based semantics is correlated with the time of the search, which shows the most recent state of knowledge, losing the history of semantics. Literature broadly covers these aspects, considering web-based semantics to be one of the most accurate because of the constant update of content [2], [5], [7], [8].

Furthermore, the Web covers many domains and disciplines, representing a vast amount of common knowledge. This broad coverage enables the extraction of semantic relationships spanning multiple domains, thereby capturing the subtleties and context-dependent meanings of terms. As a result, semantic similarity from the Web tends to be more representative of real-world knowledge than ontologies or dictionaries because it is not limited to a specific domain or predefined set of concepts. A large community of users dynamically updates the information on indexed terms almost in real-time as significant events occur.

Search engines can then be used as a practical method for extracting co-occurrences and calculating semantic proximity between pairs or groups of terms on the Web. The occurrence and co-occurrence data were collected by scraping web search pages, which provided more realistic data than when using the search engine's APIs [7]. To keep the results clean, the user profile data were deleted before each query, using an automated script to simulate the actions of a human user.

In the Facebook dataset, we considered users with 10 or more comments as active users who could express their personal opinions. For the selected users, we calculated the frequency of each word and retained the top ten words in the ranking. In this manner, a $\sim$ 20K word dictionary was built. For the users' comments, all the words that appeared in the dictionary were considered for semantic similarity and were analyzed for all the comments of active users.

B. SEMANTIC PROXIMITY MEASURES

Using the Web as a knowledge base for semantic proximity allows for a dynamic, up-to-date, and contextually rich analysis of the relationships between terms, providing an accurate and representative understanding of their semantic connections, in particular, to assess and quantify similarity or distance to a given concept [8], [10].

Formally, the approach uses a search engine, called S, as a black box for querying the Web and extracting statistics about the co-occurrences of terms.

Let $f(w) = S(w_1)$ and $f(w_1, w_2) = S(w_1, w_2)$ denote the number of results returned by S, with search terms w_1 and w_1, w_2 (in general, w is a *word*). N is defined as the total number of documents indexed by S. In our case, N was realistically approximated with a number higher than the maximum number of occurrences found in the data set.

The chosen proximity measures use the occurrences and co-occurrences of terms or a probability estimate that can be calculated directly from the frequency (i.e., from the occurrences), as $P(w) = \frac{f(w)}{N}$.

As we normalize all measures, both distance and similarity can be compared as complementary. The proximity measures examined and tested are as follows:

1) AVERAGE CONFIDENCE (AVGCONF)

Confidence (CF) [2] is an asymmetrical measure used in rule mining to measure the degree of confidence in rule $X \rightarrow Y$. Given the number of queries containing w_1, the confidence($X \rightarrow Y$) indicates the percentage of queries that also contain w_2. From a probabilistic perspective, confidence approximates the conditional probability:

$$\text{Conf}(X \rightarrow Y) = P(Y|X) = \frac{f(w_1, w_2)}{f(w_2)}. \tag{3}$$

where $\frac{f(w_1,w_2)}{f(w_1)}$, (i.e. $\frac{P(X \cap Y)}{P(X)}$) represents the a priori probability. The mutual (i.e., average) confidence [7], [8] can be defined as the average of the confidence (CF) between $Conf(w_1, w_2)$ and $Conf(w_2, w_1)$.

$$\text{AVGConf}(w_1, w_2) = \frac{\text{Conf}(w_1, w_2) + \text{Conf}(w_2, w_1)}{2}. \tag{4}$$

This equation provides a balanced and symmetrical measure.

2) NORMALIZED GOOGLE DISTANCE (NGD)

Normalized Google Distance (NGD) [5] is a measure of semantic proximity that relies on the idea that similar concepts occur together in a large number of documents on the Web. Therefore, the frequency of documents returned by S approximates the distance between related semantic concepts.

The NGD of two words w_1 and w_2, using their frequency $f(w_i)$ in the knowledge base, is formally defined as:

$$\text{NGD}(w_1, w_2) = \frac{\max(\log f(w_1), \log f(w_2)) - \log f(w_1, w_2)}{\log M - \min(\log f(w_1), \log f(w_2))} \tag{5}$$

3) POINTWISE MUTUAL INFORMATION (PMI)

PMI [4], [20] is a proximity measure used in statistics and information theory to quantify the association between two events. Our events are the co-occurrences of our two words w_1 and w_2:

$$\text{PMI}(w_1, w_2) = \log_2 \frac{f(w_1, w_2)}{f(w_1)f(w_2)}, \tag{6}$$

where $f(w_1, w_2)$ are the joint probabilities of w_1 and w_2 occurring together, and $f(w_1)$ and $f(w_2)$ are the marginal probabilities of w_1 and w_2 occurring independently (i.e. their frequency in the knowledge base).

Thus, PMI is a rough estimate of the amount of information that a word provides about another word in a pair. It measures the information supplied by the occurrence of event w_2 concerning the occurrence of event w_1 in particular. A high PMI rating suggests that the uncertainty has decreased. PMI has been effectively used to detect synonyms based only on word count.

However, it should be noted that PMI may not yield meaningful conclusions for low-frequency data. While PMI is an effective measure of independence, with values near zero suggesting that events occur independently, it may not be as useful in evaluating dependence, because the dependency score is connected to the frequency of specific words. In other words, PMI may be affected by the frequency of the terms being compared, thereby lowering the reliability level.

Nonetheless, PMI is beneficial in a variety of natural language processing and information retrieval applications [8]. It can assist in detecting significant correlations between words and show patterns in large corpora of text or web pages.

V. EXPERIMENTS: CALCULATING SEMANTIC PROXIMITY

This section details the specific application of the proximity measures presented in the Methodology section.

In our experiment, we aimed to calculate the semantic proximity of the significant terms in the user's comments (i.e., a vector of words) to the concept of beneficence. Therefore, in our measures of semantic proximity, w_1 is the term under consideration and w_2 is the term 'beneficence.' For *NGD*, *M* was fixed at 252,700,000,000 (i.e., 252.7 billion). Because *NGD* was the only measure of distance for our comparison, we used its inverse *(1-NGD)* to compare its results with the other measures of proximity.

The phases of the extraction include:

1) Preprocessing of comments (tokenization, stop words removal, filtering)
2) Web scraping through the Google search engine for word-level semantic proximity
3) Comment-level proximity modeling
4) User-level proximity modeling
5) Neighborhood-level proximity modeling

This process involves several steps for extracting relevant information from text data and estimating the semantic proximity between terms using web-based measures. The final step involves analyzing the echo chambers of the comment's proximity to beneficence.

A. PREPROCESSING OF COMMENTS

Before text mining the Facebook dataset, we performed preprocessing to refine the data and efficiently extract significant terms. Our focus was on identifying tokens (i.e., individual words) that have semantic meanings and are relevant to the concept of beneficence. To this end, we performed the following preprocessing tasks:

- Tokenization: We break the text down into its constituent tokens, that is, terms that may have meaning related to beneficence, isolating text units to be analyzed for semantic meaning.
- Stop word filtering: We remove punctuation and stop-words (e.g., non-semantic tokens that do not contribute to the understanding of beneficence), thereby optimizing the processing time. Stop words typically include articles and prepositions. In our case, we also included person names, symbols, and numbers.
- Emoji translation: We translated emojis into their respective short Unicode descriptions to ensure that the additional meaning conveyed by the emoji was preserved and could be part of the analysis.

Through these preprocessing steps, we ensured that the dataset was set up for an effective web search for terms that were significantly associated with the concept of beneficence while excluding extraneous data that could affect the mining process.

B. WORD-LEVEL SEMANTIC PROXIMITY

After preprocessing the comments, we represented each comment as a vector resulting from the preprocessing step, consisting of the remaining semantically relevant words. This vector was subsequently used in the vector space model.

We proceeded to the following steps to quantify the semantic relationship between the terms in our vector and the concept of beneficence.

- Automated Web Search: Each term from the pre-processed comments was entered into Google Search.[1] We recorded the frequency of appearance of each term in the search results as well as its co-occurrence with the word 'beneficence.'
- Data Scraping: We used the Selenium WebDriver to scrape the search results. Selenium simulates human-like interactions with web pages, which helps avoid potential bias from APIs, the results of which are limited and do not mirror the human experience [8].
- Search Query Management: To mimic natural browsing behavior and minimize the risk of being flagged by Google's anti-bot mechanisms, we introduced random delays between consecutive searches. Additionally, to prevent personalized search results, we set up Selenium to reset the search history and profile before any search, avoiding the previous results influencing the following searches.

Data on the occurrence and co-occurrence of each term with beneficence were stored in a structured dictionary for analysis. Through the implementation of these measures, we created a robust dataset that reflected the word-level semantic proximity between common terms used in Facebook comments and the concept of beneficence. This dataset will inform further analysis of semantic associations in social media debates.

At the end of this phase, three web-based similarity measures (see Subsection IV-B) were calculated for each pair comprising a term from the user comment and the term 'beneficence' to quantify the semantic closeness between the user-generated content and the concept of beneficence.

C. COMMENT-LEVEL PROXIMITY MODELING

In the subsequent analysis at both the comment and user levels (detailed in the following sections), we implemented a vector space model to represent the semantic proximity [7].

[1] The web scraping phase took place in mid-July 2022.

This model is characterized by a selection operator, denoted as $SEL \in \{MAX, AVG\}$, and a semantic proximity measure, denoted as P. In our experimental framework, P represents a set of proximity measures, specifically $proximity \in \{AVGConf, (1 - NGD), PMI\}$, where $(1 - NGD)$ represents the inverse of the Normalized Google Distance. The NGD is a measure of semantic distance, whereas the others are proximity measures. By transforming the normalized Google distance into an equivalent inverse proximity measure, we can compare the three.

This scheme, already used in image-based [8] and vector-based [7] emotion recognition to compute the similarity of terms to elements from an image or emotion model respectively, defines a class of semantic similarity functions where a different elementary distance and a different composition operator generate a single semantic similarity function within the class.

For a given task or dataset, this technique allows for flexibility in selecting the most appropriate measures. We can generate a unique semantic similarity function tailored to the specific requirements of the task or dataset by varying the elementary distance (proximity measure) and composition operator (selection method) for a total of six different proximity values that can be derived for each comment. The six values were encapsulated by the comment-level proximity metric CP, where C denotes a comment.

The proximity for a given selection method $sel1$ is computed as follows:

$$CP^{sel1} = SEL\{P(w_1, w_2), \forall w_1, w_2 \in C\} \tag{7}$$

where $sel1 \in \{MAX, AVG\}$. E.g., to refer to $sel1 = MAX$, we would write CP^{MAX}.

D. USER-LEVEL PROXIMITY MODELING

Building on modeling proximity at the comment level, we extended our analysis to the user level. To provide a comprehensive measure of a user's semantic proximity to beneficence, we aggregated the comment-level proximity values (CP^{sel1}) across all comments written by each user. This aggregation was performed using a second selection operator $sel2 = SEL$, resulting in 12 unique proximity values for each user referred to as user-level proximity (UP):

$$UP^{sel1,sel2} = SEL\{CP_i^{sel1}\} \tag{8}$$

where CP_i^{sel1} represents the collection of proximity values at the comment level for user i. These values were previously aggregated using the $sel1$ function. The operator $sel2$ was then applied to the set, resulting in a user-level proximity value. For example, if we have previously selected the maximum value of the proximity at the level of the comments ($sel1 = MAX$) and we now choose to average these maximum values at the level of the users ($sel2 = AVG$), we denote the user-proximity aggregation as $UP^{MAX,AVG}$.

Figure 3 shows the histograms for each proximity measure to visualize and analyze the distribution of the proximity values at the user level, providing insight into the distribution of users' semantic closeness to the concept of beneficence and allowing the observation of trends and comparison of the measures.

E. NEIGHBORHOOD-LEVEL PROXIMITY MODELING

According to our definition, given in section III-B (see also Figure 1), two users in the network $\mathcal{C}$ are connected if they comment on the same post. The term neighbor, therefore, refers to a user who has interacted with or commented on the same post as another user (e.g, if users u_1 and u_n both commented on the same post p_1, they are considered neighbors).

The concept of neighbors is used to analyze the patterns of interactions and examine the level of beneficence exhibited by users within their immediate social context. In other words, by studying the interactions between neighbors, we can gain insights into the dynamics of online discussions, the spread of opinions, and the similarities or differences in beneficence levels among users who engage with the same content.

At the neighborhood level, only AVG is used as an aggregation function of the user's user-level proximity (UP) for each neighbor j of user i:

$$NUP^{sel1,sel2} = AVG\{UP_{i,j}^{sel1,sel2}\} \tag{9}$$

where $UP_{i,j}^{sel1,sel2}$ are the sets of proximity values for the comments made by each neighbor j of user i in the data set. In this step, we obtained 12 values of the proximity of the user's neighbors. Finally, we computed the neighborhood-level proximity distribution of NUP. The resulting histograms, one for each proximity measure, are shown in Figure 4.

VI. RESULTS ANALYSIS AND DISCUSSION

We examined the plots of the proximity-measure distributions to determine the answer to our research questions for goals G1 and G2.

A. G1: CHECK WHETHER SOCIAL MEDIA USERS DISCUSSING CONTROVERSIAL TOPICS EXHIBIT POLARIZATION ON BENEFICENCE

According to the proposed methodology, polarization in beneficence, expressed by users and/or their neighbors, leads to at least two local maxima in the proximity measure distributions [6]. Instead, a lack of polarization corresponds to a distribution with a single local maximum. Our analysis shows a lack of polarization in the user and neighborhood-level proximity, as shown in Figures 3 and 4.

The results in Figure 3 show a lack of polarization in beneficence between the two groups (i.e., anti-vax vs. pro-vax), as all users tended to be concentrated, on average, in the same range of beneficence. Therefore, the answer to our first goal (G1 in Section I) is negative.

This plot pattern does not follow echo chambers regarding opinions on vaccinations, as presented in the source

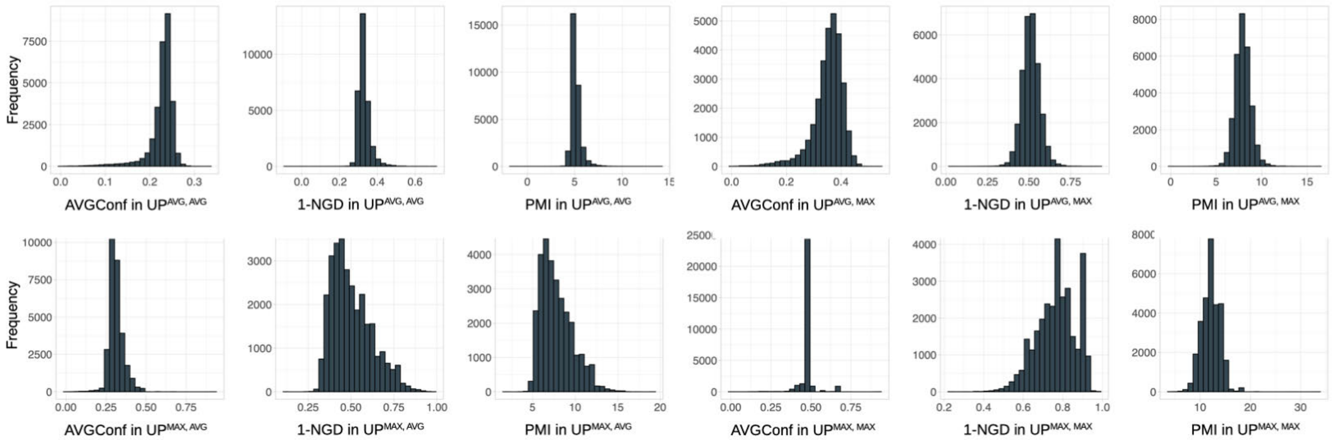

FIGURE 3. **User-level proximity distribution. In each plot, a value of proximity is reported on axis x. The naming convention is derived from the equations (see Equations 7, 8 and 9).**

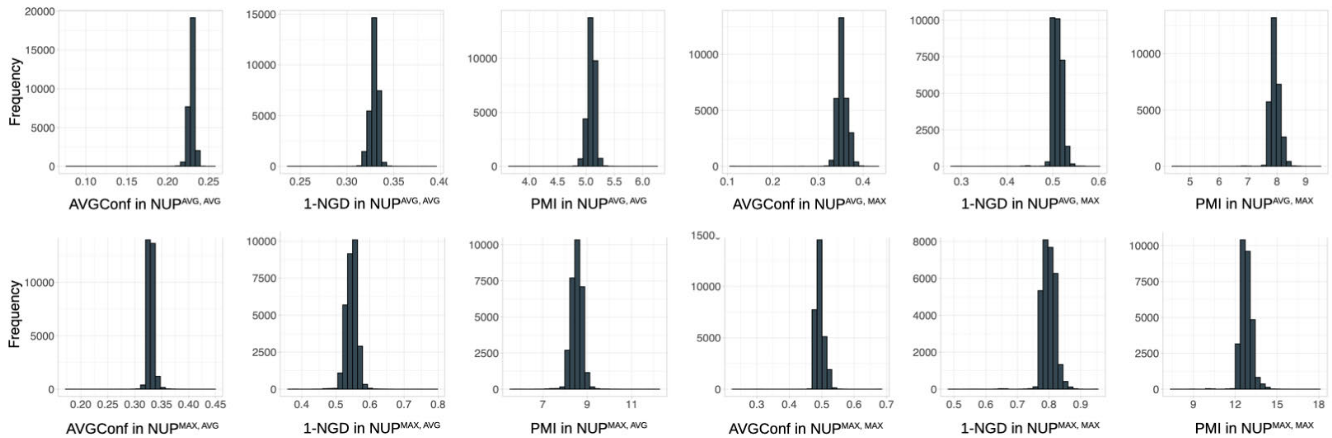

FIGURE 4. **Neighborhood-level proximity distribution. In each plot, a value of proximity is reported on axis x. The naming convention is derived from the equations (see Equations 7, 8 and 9).**

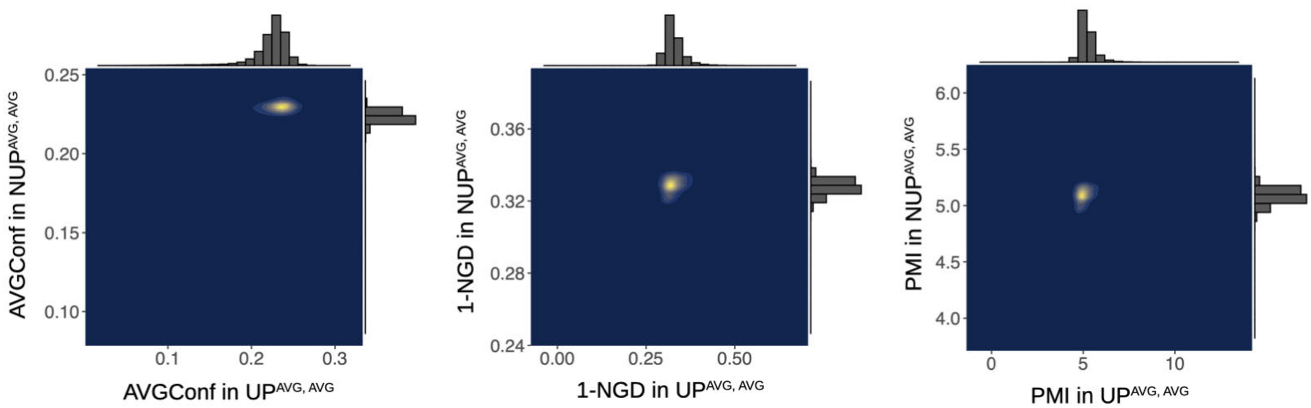

FIGURE 5. **Joint distribution between the average values of each user's *AVGConf*, $1 - NGD$, and *PMI* with the user's neighborhood average proximity measures [11]. The figure shows how the same beneficence measures characterize users and their neighborhoods with similar scores.**

dataset [19]. Thus, we need to note and consider that opinions and beneficence can emerge differently from the same subjective point of view, where the subjective polarized opinions on the controversial topic share a similar level of beneficence, that is, a pro-social attitude.

Based on the definition of beneficence [13], [14] is, thus, apparent that both polarized groups believe that their positions and actions can benefit the community. We can observe such a shared level of beneficence between the two groups in the dynamics of online discussion, where

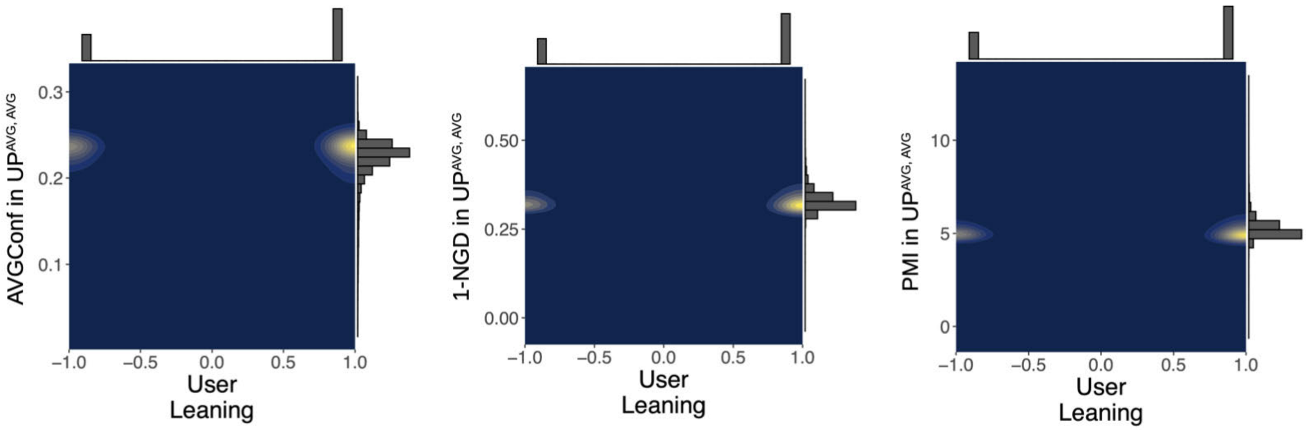

FIGURE 6. Joint distribution of the user's individual leaning in the anti-vax (-1) and pro-vax (+1) spectrum together with their *AVGConf*, $1 - NGD$, and *PMI* [11]. From a perspective of inclination, the distribution of the beneficence metrics unveils how users, despite opposite views on the topic of vaccination, are characterized by similar semantics in their comments. The similarity can be interpreted as a signal for their belief that their stance is beneficial for society.

each polarized group engages in the same content (i.e., comments of neighbors). Indeed, this inclination can be seen in many highly-valued comments (i.e., semantically close to the concept of beneficence) from both groups.

B. G2: CHECK WHETHER SOCIAL MEDIA USERS CAN BE CLUSTERED IN ECHO CHAMBERS FOR BENEFICENCE

In G2, an echo chamber relates to a situation in which users predominantly interact with and reinforce their own beliefs within a closed network, leading to limited exposure to diverse perspectives. Thus, we plotted the joint distribution between the proximity values of each user and those of the user's neighbors.

In Figure 5, we observe a high concentration of measures expressing beneficence in the comments of Facebook users. This concentration describes the tendency of users to display homogeneous beneficence, which is not reflected in their opinion polarization concerning the vaccination topic, as shown in Figure 2. This result confirms the previous analysis of G1. The answer to our second question G2 (see Section I) is thus negative: we cannot identify echo chambers in the topic of beneficence measured on users and their neighborhoods.

By plotting the joint distributions of the individual users' polarization (i.e., pro-vax and anti-vax) with the proximity values of each user (see Figure 6), we observe a prominent bimodality of users' opinions, according to which two main stances (pro-vax and anti-vax) emerge. Nonetheless, these two echo chambers about vaccines are still characterized by an akin propensity towards beneficence, testified by the average proximity to the beneficence of users' neighbors, which reach similar values, regardless of the individual leaning.

The analysis of terms frequently appearing in interactions between users (i.e., the similarity of users' comments to the concept of beneficence) showed a degree of agreement on the level of beneficence, even when the same users expressed extremely different opinions on the topic.

Finally, we observe from the plots that similarity with beneficence remains relatively low for each measure (i.e., within the first half of the range), as expected for a dataset on controversial topics, particularly in the context of public health debates.

In Table 2, we provide a brief snapshot of user comments from both pro- and anti-vaccination echo chambers that intuitively align with the observed patterns of beneficence derived from our similarity analyses, demonstrating variances in beneficence levels. This selection illustrates how expressions of support or empathy correlate with higher beneficence, whereas dismissive or exclusionary remarks align with lower beneficence. These examples serve as a microcosm of broader dataset dynamics, highlighting the prevalence of polarized yet consistently low-beneficence discourse within the studied controversial topic.

VII. STUDY LIMITATIONS AND FUTURE DIRECTIONS

Our study leverages web-based semantics to capture the knowledge present on the Web at a specific point in time. This snapshot approach offers up-to-date insights but naturally limits our ability to track dynamic changes in discourse over time. To address this limitation, future studies should explore methodologies for real-time data analysis.

Although our method provides deep analytical insights into social media discourse, it is computationally intensive and time-consuming, particularly for large-scale datasets. This aspect, intrinsic to our approach, suggests a tradeoff between the depth of analysis and processing time. However, the unique insights gained into social dynamics and their implications for policy-making highlight the value of this investment. Future implementations could focus on optimizing computational efficiency to balance these factors.

Our dataset predominantly featured comments with negative connotations and a generally low level of beneficence

TABLE 2. Sample comments for polarized pro/anti-vax echo chambers and higher/lower level of beneficence, illustrating how expressions of mutual support or empathy correlate with higher beneficence.

	Higher beneficence	Lower beneficence
Pro-vax	*Amazing! Well done Deborah. A fantastic result!*	*I would love it if those people all left the internet.*
Anti-vax	*So sorry for your loss. Prayers for you! Namaste!*	*Ugh... enough already! We get it! They want us dead!*

in posts, thus reflecting the controversial nature of the debate examined. In future works, it would be interesting to validate our approach on other datasets with different levels of polarization of opinion, when available.

The analysis of such limitations outlines a roadmap for enhancing research on web-based semantics and their applications in social media analysis.

VIII. CONCLUSION

The work provides a computational approach to measure beneficence, understood as "the concept of having a positive impact on others" in online personal opinions. To this end, we applied three measures of similarity (i.e., average Confidence, Normalized Google Distance, and Pointwise Mutual Information) to assess the proximity to the term beneficence of comments that users made on posts about the controversial topic of vaccinations and calculated the correlation with beneficence at user and neighborhood levels. To the best of our knowledge, this is the first attempt to estimate pro-social intentions and beliefs regarding beneficence.

Additionally, we applied these measures to study the polarization of social media users. We found that the polarization of media users regarding their opinions on vaccination was not confirmed in terms of beneficence as measured by web-based semantic proximity. Indeed, statements can be observed in the comments made by members of both groups showing pro-social attitudes, and we observe that the semantics of beneficence show comparable levels between the two groups.

Our methodology and results may contribute to a better understanding of the motivations and beliefs of various (polarized or non-polarized) groups in online communities. Therefore, vaccination was used as a use case in this study. The same methodology can be applied to study the benefits of different user groups and text sources (e.g., email and blogs), showing polarized opinions on general topics. The same approach can be used to study other pro-social attitudes.

We believe that our study can inspire other researchers to study psychological needs and pro-social attitudes and behaviors in online communities. It also contributes to a better understanding of well-being in digital social life.

ETHICAL IMPACT STATEMENT

Data are presented in an aggregated manner and their treatment complies with the terms, conditions, and privacy policies of the respective websites and EU General Data Protection Regulation.[2]

[2]https://gdpr-info.eu/

The source dataset from Facebook was part of a previously published article, and the computation of word occurrence was novel. Data in the source dataset were collected using publicly available Facebook Graph API.[3] The anonymized dataset of novel data about word occurrence and co-occurrence with *beneficence*, including the calculation of proximity measures, is publicly available in the Harvard Dataverse, named "Pro-Sociality: Beneficence in Social Network Posts": https://doi.org/10.7910/DVN/KAHXIS, and can be used citing this paper.

Figures 5 and 6 were also included in our preliminary work [11].

The work presented in this paper provides a novel computational approach for measuring beneficence: the authors do not intend to validate any particular opinion of users on the controversial topic of vaccination.

ACKNOWLEDGMENT AND AUTHORS' CONTRIBUTION

The authors would like to thank Walter Quattrociocchi and Paul Benjamin Ciurla from the Sapienza University of Rome for their contributions to the preliminary study [11]. DeepL Write AI[4] and Paperpal Preflight[5] have been used to cross-check the language and grammar in the manuscript.

Valentina Franzoni designed the study; Valentina Franzoni, Maurizio Mancini, and Radoslaw Niewiadomski developed the theoretical framework; Maurizio Mancini and Valentina Franzoni curated the experiments on semantic similarity; Matteo Cinelli and Gabriele Etta curated the experiments on echo chambers; all authors wrote the article; and Valentina Franzoni proofread and reviewed the article.

REFERENCES

[1] A. Bessi, "Personality traits and echo chambers on Facebook," *Comput. Hum. Behav.*, vol. 65, pp. 319–324, Dec. 2016.
[2] V. C. Cheng, C. H. C. Leung, J. Liu, and A. Milani, "Probabilistic aspect mining model for drug reviews," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 8, pp. 2002–2013, Aug. 2014.
[3] N. Chetty and S. Alathur, "Hate speech review in the context of online social networks," *Aggression Violent Behav.*, vol. 40, pp. 108–118, May 2018.
[4] K. W. Church and P. Hanks, "Word association norms, mutual information, and lexicography," *Comput. Linguistics*, vol. 16, no. 1, pp. 22–29, 1990.
[5] R. Cilibrasi and P. M. B. Vitányi, "The Google similarity distance," 2004, *arXiv:cs/0412098*.
[6] M. Cinelli, G. D. F. Morales, A. Galeazzi, W. Quattrociocchi, and M. Starnini, "The echo chamber effect on social media," *Proc. Nat. Acad. Sci. USA*, vol. 118, no. 9, 2021, Art. no.e2023301118.
[7] V. Franzoni, G. Biondi, and A. Milani, "A web-based system for emotion vector extraction," in *Computational Science and Its Applications—ICCSA*. Cham, Switzerland: Springer, 2017, pp. 653–668.

[3]https://developers.facebook.com/docs/graph-api
[4]https://www.deepl.com/write
[5]https://preflight.paperpal.com/

[8] V. Franzoni, Y. Li, and A. Milani, ''Set semantic similarity for image prosthetic knowledge exchange,'' in *Computational Science and Its Applications—ICCSA*. Cham, Switzerland: Springer, 2019, pp. 513–525.

[9] M.-A. Kaufhold and C. Reuter, ''The self-organization of digital volunteers across social media: The case of the 2013 European floods in Germany,'' *J. Homeland Secur. Emergency Manag.*, vol. 13, no. 1, pp. 137–166, Apr. 2016.

[10] C. H. C. Leung, A. W. S. Chan, A. Milani, J. Liu, and Y. Li, ''Intelligent social media indexing and sharing using an adaptive indexing search engine,'' *ACM Trans. Intell. Syst. Technol.*, vol. 3, no. 3, pp. 1–27, May 2012.

[11] M. Mancini, M. Cinelli, G. Etta, P. Ciurla, R. Niewiadomski, W. Quattrociocchi, and V. Franzoni, ''Computing beneficence: A study of pro-social attitudes in comments of online social media users,'' in *Proc. 11th Int. Conf. Affect. Comput. Intell. Interact. Workshops Demos (ACIIW)*, Sep. 2023, pp. 1–4.

[12] F. Martela and T. J. J. Riekki, ''Autonomy, competence, relatedness, and beneficence: A multicultural comparison of the four pathways to meaningful work,'' *Frontiers Psychol.*, vol. 9, Jul. 2018, Art. no. 327587.

[13] F. Martela and R. M. Ryan, ''Distinguishing between basic psychological needs and basic wellness enhancers: The case of beneficence as a candidate psychological need,'' *Motivat. Emotion*, vol. 44, no. 1, pp. 116–133, Feb. 2020.

[14] F. Martela and R. M. Ryan, ''The benefits of benevolence: Basic psychological needs, beneficence, and the enhancement of well-being,'' *J. Personality*, vol. 84, no. 6, pp. 750–764, Dec. 2016.

[15] B. McLennan, J. Whittaker, and J. Handmer, ''The changing landscape of disaster volunteering: Opportunities, responses and gaps in Australia,'' *Natural Hazards*, vol. 84, no. 3, pp. 2031–2048, Dec. 2016.

[16] R. Moore, R. S. Purvis, E. Hallgren, D. E. Willis, S. Hall, S. Reece, S. CarlLee, H. Judkins, and P. A. McElfish, ''Motivations to vaccinate among hesitant adopters of the COVID-19 vaccine,'' *J. Community Health*, vol. 47, no. 2, pp. 237–245, Apr. 2022.

[17] V. Pizziol, X. Demaj, R. Di Paolo, and V. Capraro, ''Political ideology and generosity around the globe,'' *Proc. Nat. Acad. Sci. USA*, vol. 120, no. 15, Apr. 2023, Art. no. e2219676120.

[18] R. M. Ryan and E. L. Deci, ''Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being,'' *Amer. Psychol.*, vol. 55, pp. 68–78, Jan. 2000.

[19] A. L. Schmidt, F. Zollo, A. Scala, C. Betsch, and W. Quattrociocchi, ''Polarization of the vaccination debate on Facebook,'' *Vaccine*, vol. 36, no. 25, pp. 3606–3612, Jun. 2018.

[20] P. D. Turney, ''Mining the web for synonyms: PMI-IR versus LSA on TOEFL,'' in *Proc. 12th Eur. Conf. Mach. Learn.*, Freiburg, Germany. Springer, Sep. 2001, pp. 491–502.

[21] N. Weinstein and R. Ryan, ''When helping helps: Autonomous motivation for prosocial behavior and its influence on well-being for the helper and recipient,'' *J. Personality Social Psychol.*, vol. 98, p. 222, Feb. 2010.

MATTEO CINELLI (Member, IEEE) is currently an Assistant Professor of computer science with the University of Rome ''La Sapienza,'' Rome, Italy. He is a member of the Center of Data Science and Complexity for Society (CDCS). His research interests include network and data science, mostly in the context of information diffusion, and social media.

GABRIELE ETTA (Member, IEEE) received the M.Sc. degree in data science from the University of Padua, Italy, in 2020. He is currently pursuing the Ph.D. degree in data driven modeling of social dynamics with the Sapienza University of Rome, Italy. His research interests include complex networks, information diffusion, and computational social science.

VALENTINA FRANZONI (Senior Member, IEEE) received the Ph.D. degree in engineering for computer science from the DIAG ''A. Ruberti'' Department of Computer, Control, and Management Engineering, University of Rome ''La Sapienza.'' She is currently a Tenure-Track Professor and the Coordinator of affective computing with the EmoRe Research Group, Department of Mathematics and Computer Science, University of Perugia, Italy, and a Lecturer with the Department of Computer Science, Hong Kong Baptist University. Her research interests include affective computing to ethics for artificial intelligence, complex networks, and SNA.

MAURIZIO MANCINI is currently an Associate Professor with the Department of Computer Science, Sapienza University of Rome. He is also the Co-Founder of the PsyComp Laboratory. He conducted research within several EU ICT Projects in FP5-7, H2020, and NextGenerationEU. His research interests include models and algorithms for automated expressive movement analysis and synthesis in the field of multimodal interfaces and HCI.

RADOSLAW NIEWIADOMSKI is currently an Assistant Professor with the University of Genoa and the Founder of the hAppI Laboratory. He has actively participated in several EU research projects, such as FP6 CALLAS, FP7 ILHAIRE, and H2020 DANCE. He has contributed to over 100 peer-reviewed conferences papers and journal articles. His research interests include emotion recognition, nonverbal behavior synthesis, and multimodal interaction.

Open Access funding provided by 'Università degli Studi di Perugia' within the CRUI CARE Agreement