



OPEN

Public concerns and attitudes towards autism on Chinese social media based on K-means algorithm

Qi Zhou, Yuling Lei, Hang Du & Yuexian Tao

To investigate the hot topics and attitudes of autism in the larger community. In this study, we analyzed and summarized experimental texts from the social media platform Zhihu using the TF-IDF algorithm and K-means clustering approach. Based on the analysis of the 1,740,826-word experimental text, we found that the popularity of autism has steadily risen over recent years. Sufferers and their parents primarily discuss autism. The K-means clustering algorithm revealed that the most popular topics are divided into four categories: self-experience of individuals with autism, external views of individuals with autism, caring and stressful behaviors of caregivers, and information about autism. This study concluded that people with autism face more incredible negative emotions, external cognitive evaluations of the autistic group reflect stereotypes, the caregiver's family suffers high financial and psychological stress, and disorders caused by disease in autistic individuals.

Autism Spectrum Disease (ASD) is a severe neurodevelopmental disorder with an unclear cause and a poor prognosis, marked by difficulties in verbal communication, abnormalities in social interaction, and stereotyped behavior¹⁻³. According to the World Health Organization, the prevalence of autism has risen at an alarming rate over the two decades, reaching 6.5 per 1000 persons⁴. Autism prevalence among children in the American Health Interview Survey increased from 1.68% in 2014 to 2.48% in 2017, meaning 1 in 59 persons^{5,6}. The Chinese Autism Education and Rehabilitation Industry Development Report estimates that there are over 10 million autistic persons in China, with over 2 million children⁷. Moreover, some researchers find that China may have a higher prevalence of autism due to a shortage of medical workers, public misinformation, and parental concealment, resulting in fewer diagnoses for people with autism in remote areas⁸. Wu's research also states that the Chinese ASD community is in a neglected position and urgently deserves the attention of the public⁹.

However, most autism-related research has focused on high-income countries, and there is a lack of research on the low- and middle-income countries where most people with autism live¹⁰⁻¹². Studies have found that public attitudes towards ASD are very different in the US and China, with 86–91% of US citizens showing adequate awareness of autism, compared to 57–65% in China¹³. In addition, researchers have debated the involvement of people with autism in public engagement, with some believing that the illness makes it difficult for the autistic community to engage in public engagement and others believing that people with autism are capable of actively participating in public engagement¹⁴. However, the different perspectives of the public and professionals lead to different views, so it is essential to investigate the public's attitudes toward people with autism¹⁵.

Due to the inclusiveness, social media are now the channel where people express their perspectives on ASD allowing professionals to share their experiences and enabling the public to comprehend crucial information¹⁶. Researchers demonstrate that television and social media platforms such as Twitter^{17,18}, Google¹⁹, YouTube^{20,21}, and Facebook²² are now equally the primary sources of information for individuals with autism and their stakeholders. As a result, information disseminated on social media will influence public perceptions of ASD, no matter facts or suggestions. Negative portrayals of ASD in the media may reinforce public misconceptions^{23,24}; however, the media may also improve public attitudes^{13,25,26}. Some researchers have begun to focus on the analysis of autism data on social media platforms such as Twitter and YouTube in an attempt to raise public awareness of autism^{16,18,27}. However, these studies are limited to English-language media, lacking of those on Chinese media.

Researchers have recently applied text-mining algorithms to analyze potential information in the medical field²⁸⁻³⁰. Density-based clustering methods exhibit good robustness against noise and outliers²⁹, whereas hierarchical clustering has the inherent ability to determine the optimal number of clusters, thereby ensuring a higher level of stability³⁰. However, K-means is more efficient than hierarchical and density-based clustering methods due to its lower computational complexity, especially when dealing with massive text data. Furthermore, the

Hangzhou Normal University, 2318th Yuhang Tang Avenue, Hangzhou 311121, Zhejiang Province, China. email: 20021065@hznu.edu.cn

cluster results represented by their centroids provide a more intuitive depiction, facilitating a clearer understanding of the clustering characteristics and underlying topics³¹. Given the vast scale of internet comment texts, this study uses the K-means algorithm for topic clustering. However, the K-means clustering algorithm has the disadvantage of requiring the number of clusters to be defined in advance; therefore, the elbow calculation method is used to obtain the K values in this study. Moreover, by combining PCA dimensionality reduction and TF-IDF techniques, this research effectively captures the distinctive characteristics of text data while simultaneously reducing the dimensionality of the feature space, thereby enhancing the efficacy of clustering.

In summary, this study aims to fill the gap in autism research in Chinese social media, investigate potential autism misconceptions, and raise the public's awareness and tolerance of autism. The rest of the sections are as follows: the second part is about the methods, which include PCA dimensionality reduction, TF-IDF algorithm, and K-Means algorithm; the third part is about results; the fourth part is discussion; and the fifth part is the conclusion.

Method

Study subjects. Considering the need to design the text materials to align with the context of Chinese social media³², we collected the replies to questions related to “autism” from Zhihu. These replies were obtained as experimental text. The data collection was completed by February 20, 2022.

Data collection. In this study, we use Python 3.7 to gather the dataset of autism-related Q&A content, editing time, self-presentation, number of views, approvals, and secondary ratings. To ensure that the dataset data fits the algorithm requirements, we merge the dataset's data and remove duplicate text and blank fields.

Data processing. *Text cleaning.* In this study, the experimental text was first processed with noise reduction to remove irrelevant content, such as advertisement and web links, to improve the word separation effect and accuracy. Second, we construct a specialized deactivation word list for the experimental text by combining the deactivation word lists of HIT and Baidu with the initial word classification findings. After loading the user dictionary, we use Python's Jieba database to separate words while using the deactivated word list to delete the inactive terms.

TF-IDF algorithm and PCA decomposition. 1. Calculate Term Frequency (TF), which measures the frequency of keywords appearing in the text, divided by the total number of words in the document to prevent bias towards longer documents. 2. Calculate Inverse Document Frequency (IDF), which is obtained by dividing the total number of documents by the number of documents containing the specific keyword, and then taking the logarithm of the quotient to indicate the discriminative ability of the keyword across different document categories. 3. Multiply TF and IDF to generate the high-weighted TF-IDF, which represents the importance of a term based on its frequency in the document and across the corpus.

Principal Component Analysis (PCA) is a dimensionality reduction approach frequently employed to transform high-dimensional data into a lower-dimensional representation. PCA dimensionality reduction can effectively reduce the dimensions of the TF-IDF matrix while preserving the essential data information.

This research implements the function with Sklearn. The procedure is as follows: Initially, transform the words in the text into a term frequency matrix using CountVectorizer. Second, calculate the TF-IDF weight of each word using TfidfTransformer. TfidfVectorizer combines CountVectorizer and TfidfTransformer, enabling direct generation of TF-IDF values. Finally, transform the high-dimensional TF-IDF feature vector space into a lower-dimensional representation using StandardScaler and PCA decomposition.

K-means algorithm analysis. The K-means clustering algorithm is a common and well-researched exploratory data analysis technique applied and validated in Song and Wang's research for Chinese text mining^{33,34}. The K-means algorithm enables researchers to understand text data more deeply and uncover counterintuitive insights and patterns. It is particularly beneficial for extracting usable information and knowledge from text data, classifying text data into meaningful groups, and simplifying text data via dimensionality reduction and denoising. The K-means algorithm steps are as follows:

First, Initialize cluster centers: In the KMeansClusterer of the NLTK library, the number of cluster centers can be set using the num_means parameter. This study uses the elbow algorithm to calculate inertia values for different K values to find the perfect cluster number. The inertia value decreases rapidly for small values of K, but the rate of decrease slows down as K increases. We select the elbow curve as the optimal value of K. Secondly, Iterative optimization: This method clusters the data based on a predefined data set and some clustering centers and iteratively optimizes the clustering centers' locations. Thirdly, Termination requirements: NLTK's K-means clustering is mainly controlled by the REPEATS parameter. This parameter describes the number of times the algorithm should be executed, each with a distinct initial clustering centre²². Based on the results obtained through the elbow method and iterative optimization, we conclude that K = 4 better meets the requirements of this research, as shown in Fig. 1.

K-Means algorithm result description. We use TF-IDF for feature extraction, which can provide acceptable input data for better K-means clustering results. However, their objectives are distinct. The K-means algorithm itself requires that the TF-IDF weights be normalized. Moreover, the clustering effect of K-means relies on the high-frequency words of individual features and the combined expression of the entire feature vector. Source words ranked low in the TF-IDF high-frequency words may still have a significant impact³⁵.

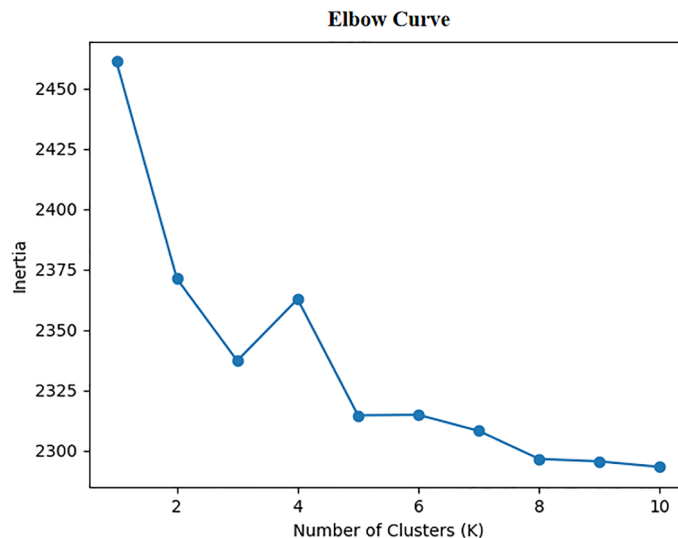


Figure 1. Elbow curve.

The K-means algorithm yields a source vocabulary and a clustering vocabulary as output. The clustering source vocabulary is the initial vocabulary set to initialize cluster centers. The source vocabulary for clustering is typically a representative set of words whose results can be used as hot topics, whereas the individual clustering vocabularies are sets of words within each clustering cluster that may be more diverse and precise. The vocabulary in the clusters that result from clustering will differ from the clustering source vocabulary, but there is a connection between them.

Results

Description of textual material. This study comprised 191 questions concerning autism from the Zhihu Q&A community as of February 20, 2022, and we analyzed the data to create a final experimental text of 4338 comments totaling 1,740,826 words.

Research trend analysis

Responses posted on the Zhihu platform involve both initial and numerous revisions. This study showed that 75.2% of initial posts and 24.8% of subsequent modifications were updated. Respondents' increased incentive to add questions is associated with their identification and the duration of autism. This phenomenon also indicates that autism groups require long-term social attention. The increase in the number of topics about autism around April each year may be related to the fact that April 2 is World Autism Day.

Basic characteristics of Zhihu users. Homemakers, individuals with autism, psychologists, rehabilitation therapists, teachers from traditional or special schools, and social media practitioners are the primary respondents to the autism topic. Homemakers and people with autism are the main posters and editors, eager to share their progress, feelings, interventions, and encouraging comments. Psychologists, rehabilitation therapists, teachers, and social media practitioners use their professional and educational experiences as the basis for information output, actively participate in community discussions and guide people with autism and their families through scientific rehabilitation interventions.

Basic characteristics of autism topic keywords. After obtaining the results of Jieba word division, the TF-IDF algorithm was used to perform word frequency statistics and mine the keywords in the text to calculate their weights, as shown in Table 1. The results of intelligent topic segmentation based on K-Means clustering analysis are shown in Table 2.

Discussion

This study used cluster analysis to identify four popular topics for user discussion: self-experience of individuals with autism, external views of individuals with autism, caring and stressful behaviors of caregivers, and information about autism.

Self-experience of people with autism. *People with autism are more susceptible to negative emotions and are unable to express them.* Anxiety, depression, humiliation, sadness, and fatigue indicate the unfavorable orientation of the autistic group's emotional life and emotional responses. While words like mask, cover-up, away, control, compelled, and powerless indicate that the autistic group is more prone to adopt negative attitudes to cope or not cope with the catharsis of their emotions in response to negative emotional influences. This study

Ranking	Keywords	Weighting
1	Autism	0.03862
2	Kids	0.03575
3	Children	0.01452
4	Intervention	0.01293
5	Disorder	0.01084
6	Language	0.01076
7	Ability	0.00962
8	Parenting	0.00903
9	Socialization	0.00737
10	Performance	0.00733

Table 1. Top 10 keywords and weights.

Source terms	Frequency statistics	Clustered vocabulary		
		Noun	Verb	Adjective
Autism	26	Emotional life, emotional reactions, reticence, communication, dissociation, glossiness	Cover up, conceal, keep away, control, forced, get better, social interaction, care, criticize	Powerless, stiff, anxious, ashamed, difficult, sad, cruel, sad, tired, depressed, cautious, indifferent, introverted, unhelpful, bewildered, helpless, reticent
Other people	53	Alien, corner, teacher, friend, campus college, fellow classmate, school, poster, TV series, environment, occasion, TV, computer, public, human dignity, society, tone, idiot, monster, world, willingness	Bully, corporal abuse, walk away, withdrawal, submit, punish, reinstate, dump, doubt, influence, taunt, distort, insult, perceive, ingratiate	Aggravated, bad, indifferent, hard, cranky, sensitive, choked, disgusted, strange, twisted, chilled, angry, naive, ridiculous
Child	35	Primary school, children, whole class, regular people, children, childhood, classroom, control, imagination, egoism, older brother and sister	Cry and fuss, run around, treat, integrate, impose chaperone, spend money, trick or learn, weak	Polite, painful, anxious, sensitive, tortured, depressed, familiar, cruel, weak, confused, embarrassed, indifferent, silent
Disorders	8	Barriers, Society, Identification, Intelligence, Worldliness	Train, diagnose, imitate, socialize, express, speak, communicate, lose control	Broken, anxious, struggling, uncomfortable, in pain

Table 2. Cluster analysis results.

indicated that only individuals with mild autism or those who had recovered successfully from scientific interventions actively shared their experiences on the Zhihu platform. During childhood, this group had an active social mindset and did not feel different from other children their age. However, due to unwarranted verbal and behavioral bullying, different worldviews, and ignorance by parents or teachers, autistic individuals may blame the illness for their situation and then magnify the negative emotions caused by the illness, causing them to turn closed off. Terrible childhood experiences increase the probability that autistic individuals will spend their adult lives alone, less socially engaged, and in a world of their creation. To address this problem, we consider that professionals can build and validate autism-related assessment observation scales to recognize emotional changes in autistic individuals, focus on changes in autistic individuals' inner worlds, and support their emotional needs.

Individuals with autism continue to express optimism in facing life's challenges. Words like reticent, apathetic, and introverted reveal the autistic group's opinion of their situation. Favorable terms such as improvement, social engagement, and communication suggest that the autistic population has a constructive outlook. A small percentage of the autistic population knows they may be weaker than average in learning ability, communication ability, emotional control, psychological tolerance, and even self-care. However, they have a nonchalant attitude and are proactive in accepting rehabilitation interventions to care for their family and friends uniquely. This phenomenon is similar to Uddin and Heseltin's findings in that the autistic group develops resilience when overcoming adversity^{36,37}.

External cognitive evaluation of the autistic group contains stereotypes. Schools, the media, and the general public are the primary sources of external stereotypes concerning autistic persons.

The school condition of the autistic people. Words like alien, bullying, aggravation, dreadful, bullying, corporal punishment, walking away, dropping out, in the corner, discipline, school, and reinstatement indicate that school life for the autistic community is not as pleasant. Due to their disease, people with mild autism struggle academically, while those with severe autism are even discouraged from attending school. In addition to the disease's susceptibility to peer indifference and bullying, some teachers, classmates, and parents of classmates disapprove of autistic children attending traditional schools, fearing they may disturb classroom discipline and the quality

of education³⁸. The study by Yang Guangxue also shows that teachers cannot give particular care to the autistic group and struggle to deal with the stress and destructive emotions caused by parents³⁹. All of the circumstances above compel us to consider how to guarantee the right of autistic children to primary education in traditional schools and how to enhance the special education system for the autistic group. Therefore, we believe that autistic groups in China need a combination of primary and supplemental education for schooling to be effective. Creating traditional and special education campuses must prioritize school environment construction, teacher capacity development, and general parental guidance⁴⁰. We may use the Australian experience in the classroom to focus on creating a barrier-free environment⁴¹. Teachers actively pursue tenure in autism special education, creating rich autistic curricular resources and systematic special teaching methods based on their general school teaching qualifications⁴⁰. Schools can invite students' parents to autism education workshops to alter their stereotypical view of disability.

Mass media in which autism is frequently discussed. Words such as television series, posters, lectures, environments, occasions, and computers indicate that autism is receiving more attention in the mass media. Nevertheless, we observe that the media now portrays autism information negatively. Exaggerating the dangers of autism, stigmatizing autism, and misinforming about autism are instances of biased news that not only make people with autism feel humiliated but also aggravate the public stereotype of the autistic community, generating ostracism and disgust. Many scholars also know that negative prejudice and stigma in media coverage of autism may exacerbate social isolation among people with autism^{42,43}. In response to this phenomenon, we believe the media should capitalize on its widespread transmission to play a constructive role. As the primary means by which the general public learns about the autistic community, the mass media should be responsible for scientifically popularizing professional autism information and paying attention to the quality and quantity of their coverage⁴⁴. Professionals in the media should create an excellent online environment for persons with autism, minimize stigmatizing expressions about autism, and protect autistic people's privacy^{13,26}.

Public stereotypes of autistic groups in society. Words such as mockery, malice, fool, monster, dump, and distortion reflect the public's explicit and implicit prejudice and rejection of the autistic group. We find that the general public has a superficial awareness of autism, and their ignorance and preconceptions make it easy for them to feel scared when interacting with autistic groups. The cause of this phenomenon might be misinformation from the media or because autism has recently attracted attention in our country⁴⁵. Therefore the correct perceptions and impressions have not yet formed. In addition, the stereotyped influence of the social public on people with autism is reflected in a vast social gap, which restricts several rights of the autistic group, including work, education, and medical care. One researcher surveyed the public's perceptions of the autistic community. Sixty-five percent of those polled believe society is unwilling to accept the autistic group and are aware of job discrimination against this group⁴⁶. Therefore, we should respond to the United Nations' initiative to educate people from all social classes about autism, eliminate prejudice and stereotypes against autistic people, and provide autistic individuals with equal access to employment, education, and medical care.

Financial and mental strain, self-stigma, and external stigma exist for caregiver families. *Families of caregivers experience both financial and psychological strain.* Words such as reality, self-care, accompanying, tormenting, directing, intervention, treatment, problem-solving, diagnosis, and evaluation indicate the overwhelming economic and psychological pressure on parents and family members to care for the autistic group. Parents must accompany their children to expert rehabilitation interventions and pay for expensive, long-term rehabilitation instruction. The situation is consistent with Cristiane et al.'s research, which found that parents in the typical autistic family struggle to balance job and family responsibilities and that autistic families in underprivileged areas may even stop getting treatment due to the cost^{46–48}. The essence of the status quo is that most autistic families in China are currently in the self-help stage, and only a small number of families are eligible for outpatient reimbursement of 70 percent^{49,50}. As a result, we must improve the social support system and develop more socialized and private rehabilitation institutes to serve a larger Population. On a psychological level, the lack of knowledge about the daily life care and professional rehabilitation of the autistic group puts a higher burden on parents. As parents age, they are anxious about spreading the stress of daily care to other family members. In addition, parents are concerned that their lack of experience and access to the most recent information about the illness might impede their children's rehabilitation⁵¹. The situation agrees with the findings of Wang et al. We claim that through the engagement of relevant national policies, we must develop a service system for the caregiving obligations of the autistic group and provide caregivers with psychological support in the form of social solidarity⁵². Let the state shoulder the rehabilitation training system for autistic groups, and alleviate a portion of the load of autistic caretakers⁵³.

Self-stigma and collateral stigma among caregivers. Words such as genetic, neurological, root cause, and avoidance are linked with autism to reflect the self-stigma of parents who blame themselves for their children's suffering and assume they caused the illness. We find this phenomenon results from unknown etiology and causes. Without a scientific explanation, parents are reluctant to admit their child has a disability and blame themselves. The concept of karma also contributes to parental guilt about their children's misfortunes⁵⁰. Words linked with autism, such as sociability, embarrassment, commoner, low self-esteem, and inclusion, reflect social discrimination and caregiver stigma may induce collateral stigma. This phenomenon necessitates researchers to consider the interaction between individuals and societal groups. Caregivers of individuals with autism may experience social exclusion and isolation, resulting in significant psychological pressure and their voluntary or subtle marginalization by society⁵⁴. Kevin et al. found that self-stigma and linked stigma were positively associated with

parental symptoms of depression and anxiety, which led to social pressure and isolation^{55,56}. Corcoran et al. even concluded that parents of autistic children are more likely to suffer from emotional disorders⁵⁷. We observe that professionals and the public must avoid stigmatizing caregivers. Professionals must be able to rapidly identify parental confidence deficits, emotions of despair, and social isolation and provide practical skills such as self-compassion and positive thinking approaches by constructing long-lasting support systems^{58–60}. The public should also reflect critically on the correct attitudes, perceptions, and behaviors necessary to eliminate discriminatory behavior⁶¹.

Barriers encountered by individuals with autism. People with autism have severe social and language impairments, and according to cluster analysis, the words connected with autism are impairment, social, training, socialization, expression, speech, imitation, and communication. Xiao Fufang et al. concluded that the autistic group exhibited adverse reactions to social impairment, language impairment, and stereotypical behavior⁶². However, the results of this study showed that the autistic group did not actively mention the presence of stereotyped behavior. We suggest that the phenomenon's cause may be related to the ability to actively share their experiences on Internet platforms, which is mild autism. It is also likely that they need to gain knowledge of people with autism and their parents to accurately identify stereotyped behavior as a dull habit. Some autistic people are conscious of their stereotypes and are compelled by societal pressure to imitate others intentionally⁶⁰. We believe that in the early diagnosis stage, focus on screening children with family history and pay attention to premature attention disengagement, repetitive behaviors, and other risk signals in children^{63,64}. In the later rehabilitation period, the government can rely on the community and schools, families, and institutions to establish an autistic support system⁶⁵.

Conclusion

On the Chinese Internet, acceptance and awareness of the autism community are gradually growing. We must address the appropriate medical and employment demands of autistic individuals, pay attention to poor autistic families, and establish a nationwide autism support system. Professionals and the media should contribute to the distribution of accurate autism information.

Data availability

The data that support the findings of this study are available from The Zhihu platform but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of The Zhihu platform. Please contact the corresponding author Tao if you want to request the data from this study.

Received: 25 April 2023; Accepted: 9 September 2023

Published online: 13 September 2023

References

- Manning-Courtney, P. et al. Autism spectrum disorders. *Autism Spectr. Disord.* **43**, 2–11 (2013).
- Rinaldi, A. Piecing together a different picture: A host of new studies on autism have begun decoding the longstanding puzzle of its causes. *EMBO Rep.* **17**, 1690–1695 (2016).
- Feng, L. et al. Autism spectrum disorder in Chinese populations: A brief review. *Asia Pac. Psychiatry* **5**, 54–60 (2013).
- World Health Organization. *Autism spectrum disorders 2019*. <https://www.who.int/news-room/fact-sheets/detail/autismspectrum-disorders> (2020).
- Xu, G., Strathearn, L., Liu, B. & Bao, W. Corrected prevalence of autism spectrum disorder among US children and adolescents. *JAMA* **319**, 505 (2018).
- Xu, G., Strathearn, L., Liu, B. & Bao, W. Prevalence of autism spectrum disorder among US children and adolescents, 2014–2016. *J. Nurs. Educ.* **319**, 81 (2006).
- People's Political Consultative Conference. Report on the Development of the Autism Education and Rehabilitation Industry in China. <http://www.rmzxb.com.cn/c/2022-04-02/3086586.shtml> (2022).
- Luo, Z. et al. Current status of and challenges posed by autism spectrum disorders in China: Prevalence, legal issues, and public awareness. *Biosci. Trends* **16**, 167–171 (2022).
- Wu, J. Y. Autism, an area that needs public attention and investment. *Sci. China Life Sci.* **58**, 931–932 (2015).
- Han, E. et al. The changing epidemiology of autism spectrum disorders. *Thrita J. Med. Sci.* **38**, 136236132211435 (2017).
- Rice, C. E. & Lee, L.-C. Expanding the global reach of research in autism. *Nurs. Ethics* **21**, 515–517 (2017).
- Happé, F. & Frith, U. Annual research review: Looking back to look forward—changes in the concept of autism and implications for future research. *J. Child Psychol. Psychiatry* **61**, 218–232 (2020).
- Yu, L., Stronach, S. & Harrison, A. J. Public knowledge and stigma of autism spectrum disorder: Comparing China with the United States. *Autism Int. J. Res. Pract.* **24**, 1531–1545 (2020).
- Hollin, G. & Pearce, W. Autism scientists' reflections on the opportunities and challenges of public engagement: A qualitative analysis. *J. Autism Dev. Disord.* **49**, 809–818 (2019).
- Kenny, L. et al. Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism* **20**, 442–462 (2016).
- Bakombo, S., Ewalefo, P. & Konkle, A. T. M. The influence of social media on the perception of autism spectrum disorders: Content analysis of public discourse on YouTube videos. *Int. J. Environ. Res. Public Health* **20**, 3246 (2023).
- Bellon-Harn, M. L., Ni, J. & Manchaiah, V. Twitter usage about autism spectrum disorder. *Autism Int. J. Res. Pract.* **24**, 1805–1816 (2020).
- Corti, L., Zanetti, M., Tricella, G. & Bonati, M. Social media analysis of Twitter tweets related to ASD in 2019–2020, with particular attention to COVID-19: Topic modelling and sentiment analysis. *Res. Nurs. Health* **25**, 176–188 (2002).
- Tárraga-Mínguez, R., Gómez-Marí, I. & Sanz-Cervera, P. What motivates internet users to search for asperger syndrome and autism on Google?. *Int. J. Environ. Res. Public Health* **17**, E9386 (2020).
- Lacruz-Pérez, I., Sanz-Cervera, P., Pastor-Cerezuela, G., Gómez-Marí, I. & Tárraga-Mínguez, R. Is it possible to educate, intervene or 'cure' autism spectrum disorder? A content analysis of YouTube videos. *Int. J. Environ. Res. Public Health* **18**, 2350 (2021).

21. Bellon-Harn, M. L., Manchaiah, V. & Morris, L. R. A cross-sectional descriptive analysis of portrayal of autism spectrum disorders in YouTube videos: A short report. *Autism* **24**, 263–268 (2020).
22. Zhao, Y., Zhang, J. & Wu, M. Finding users' voice on social media: An investigation of online support groups for autism-affected users on Facebook. *Int. J. Environ. Res. Public Health* **16**, E4804 (2019).
23. Luzinski, C. The web and public confidence in MMR vaccination in Italy. *JONA J. Nurs. Adm.* **35**, 4494–4498 (2017).
24. Gottlieb, L. N., Gottlieb, B., Bitzas, V. & Tam-Seto, L. Communicating science to the public: MMR vaccine and autism. *Vaccine* **22**, 1–6 (2003).
25. Namuli, J. D., Nakimuli-Mpungu, E., Mwesiga, E. K. & Joyce, N. S. Knowledge gaps about autism spectrum disorders and its clinical management among child and adolescent health care workers in Uganda: A cross-sectional study. *EC Psychol. Psychiatry* **9**, 112–121 (2020).
26. Hartwell, M. *et al.* Brief report: Public awareness of asperger syndrome following Greta Thunberg appearances. *J. Autism Dev. Disord.* **51**, 2104–2108 (2021).
27. Sonbhadra, S. K. *et al.* Online attitudes and information-seeking behavior on autism, asperger syndrome, and Greta Thunberg. *Int. J. Environ. Res. Public Health* **18**, 4981 (2020).
28. Nadif, M. & Role, F. Unsupervised and self-supervised deep learning approaches for biomedical text mining. *Brief. Bioinform.* **22**, 1592–1603 (2021).
29. Gao, X., Ding, X., Han, T. & Kang, Y. Analysis of influencing factors on excellent teachers' professional growth based on DB-Kmeans method. *EURASIP J. Adv. Signal Process.* **2022**, 1–11 (2022).
30. Nichols, L. *et al.* In simulated data and health records, latent class analysis was the optimum multimorbidity clustering algorithm. *J. Clin. Epidemiol.* **152**, 164–175 (2022).
31. Zhang, H.-B. User reviews clustering method based on topic analysis. *Comput. Sci.* **46**, 50–55 (2019).
32. Savargiv, M. & Bastanfard, A. Text material design for fuzzy emotional speech corpus based on Persian semantic and structure. In *2013 International Conference on Fuzzy Theory and Its Applications (iFUZZY)* 380–384 (2013). <https://doi.org/10.1109/iFUZZY.2013.6825469>
33. Arnold, S. R. *et al.* Quantitative analysis of food safety policy—Based on text mining methods. *Foods* **11**, 3421 (2022).
34. Poorchangizi, B. *et al.* The public information needs of COVID-19 vaccine: A study based on online Q&A communities and portals in China. *Front. Psychol.* **13**, 961181 (2017).
35. Chang, I.-C., Yu, T.-K., Chang, Y.-J. & Yu, T.-Y. Applying text mining, clustering analysis, and latent Dirichlet allocation techniques for topic classification of environmental education journals. *Sustainability* **13**, 10856 (2021).
36. Uddin, L. Q. Exceptional abilities in autism: Theories and open questions. *Curr. Dir. Psychol. Sci.* **12**, 509–517 (2022).
37. Heselton, G. A., Rempel, G. R. & Nicholas, D. B. 'Realizing the problem wasn't necessarily me': the meaning of childhood adversity and resilience in the lives of autistic adults. *Int. J. Qual. Stud. Health Well-Being* **17**, 2051237 (2022).
38. Ochi, M. *et al.* School refusal and bullying in children with autism spectrum disorder. *Child Adolesc. Psychiatry Ment. Health* **14**, 17 (2020).
39. Yang, G., Guo, E. & Qian, X. Survey and analysis of the professional status of teachers in autism rehabilitation institutions. *Spec. Educ. China* **11**, 66–71 (2011).
40. Cao, S. *et al.* Key competencies of students with autism spectrum disorders: Perspectives of Chinese teachers and parents. *Front. Psychol.* **13**, 1054249 (2022).
41. Karola, D., Julie-Ann, J. & Lyn, M. School's out forever: Postsecondary educational trajectories of students with autism. *Aust. Psychol.* **51**, 304–315 (2016).
42. Lord, C., Elsabbagh, M., Baird, G. & Veenstra-Vanderweele, J. Autism spectrum disorder. *Lancet* **392**, 508–520 (2018).
43. Rafiei, M., Nakhostin-Ansari, A., Meshkat, S., Khosravi, A. & Memari, A. H. Public awareness and stigma of autism spectrum disorder in Iran; An online survey. *Res. Nurs. Health* **134**, 104441 (2023).
44. Westphal, A. Public perception, autism, and the importance of violence subtypes. *J. Am. Acad. Child Adolesc. Psychiatry* **56**, 462–463 (2017).
45. Huang, Y. Research on communication and empowerment in the perspective of new media technology. *J. Hubei Univ.* <https://doi.org/10.13793/j.cnki.42-1020/c.2016.06.022> (2016).
46. Yıldız, M., Demir, Y., Kırcah, A. & İncedere, A. Caregiver burden in schizophrenia and autism spectrum disorders: A comparative study. *Psychiatry Investig.* **18**, 1180–1187 (2021).
47. Paula, C. S. *et al.* Challenges, priorities, barriers to care, and stigma in families of people with autism: Similarities and differences among six Latin American countries. *Autism* **24**, 2228–2242 (2020).
48. Cao, M. *et al.* Socioeconomic factors and autism among 16- to 30-month-old children: Evidence from a national survey of China. *Autism* <https://doi.org/10.1177/13623613221132743> (2022).
49. Yi, K. China's public libraries conduct research on services for autistic groups. *Lior:Forum* **37**, 109–115 (2017).
50. Liu, Y. *et al.* Struggle for recognition, rights, and redistribution: Understanding the identity of parents of children with autism spectrum disorder in China. *Front. Psychol.* **13**, 981986 (2022).
51. Alibekova, R. *et al.* Stress, anxiety and depression in parents of children with autism spectrum disorders in Kazakhstan: Prevalence and associated factors. *Glob. Ment. Health* **9**, 472–482 (2022).
52. Poorchangizi, B. *et al.* Parents' knowledge and attitude and behavior toward autism: a survey of Chinese families having children with autism spectrum disorder. *Transl. Pediatr.* **11**, 1445–1457 (2022).
53. Wang, H., Liu, S., Xu, J., Hu, X. & Han, Z. R. Daily experiences and well-being of Chinese parents of children with autism. *Thrita J. Med. Sci.* <https://doi.org/10.1177/13623613221144191> (2015).
54. Hajarian, M., Bastanfard, A., Mohammadzadeh, J. & Khalilian, M. SNEFL: Social network explicit fuzzy like dataset and its application for Incel detection. *Multimed. Tools Appl.* **78**, 33457–33486 (2019).
55. Oti-Boadi, M., Dankyi, E. & Kwakye-Nuako, C. O. Stigma and forgiveness in Ghanaian mothers of children with autism spectrum disorders (ASD). *J. Autism Dev. Disord.* **50**, 1391–1400 (2020).
56. Chan, K. K. S. & Leung, D. C. K. Linking child autism to parental depression and anxiety: The mediating roles of enacted and felt stigma. *J. Autism Dev. Disord.* **51**, 527–537 (2021).
57. Corcoran, J., Berry, A. & Hill, S. The lived experience of US parents of children with autism spectrum disorders: A systematic review and meta-synthesis. *J. Intellect. Disabil. JOID* **19**, 356–366 (2015).
58. Liao, X., Lei, X. & Li, Y. Stigma among parents of children with autism: A literature review. *Asian J. Psychiatry* **45**, 88–94 (2019).
59. Papadopoulos, C., Lodder, A., Constantinou, G. & Randhawa, G. Systematic review of the relationship between autism stigma and informal caregiver mental health. *J. Autism Dev. Disord.* **49**, 1665–1685 (2019).
60. Ng, C. S. M. & Ng, S. S. L. A qualitative study on the experience of stigma for Chinese parents of children with autism spectrum disorder. *Sci. Rep.* **12**, 19550 (2022).
61. Han, E. *et al.* Development of stigma-related support for autistic adults: Insights from the autism community. *Thrita J. Med. Sci.* <https://doi.org/10.1177/13623613221143590> (2023).
62. Xiao, F.-F. & Shin, H.-Y. A review of the current prevalence and research on autism spectrum disorders. *Educ. Guide* <https://doi.org/10.16215/j.cnki.cn44-1371/g4.2010.07.002> (2010).
63. Aloisio, D. & Huron, R. F. Autism as representative of disability. *Pediatr. Clin. N. Am.* **67**, 341–355 (2020).

64. Canu, D. *et al.* Early non-social behavioural indicators of autism spectrum disorder (ASD) in siblings at elevated likelihood for ASD: A systematic review. *Eur. Child Adolesc. Psychiatry* **30**, 497–538 (2021).
65. Zhou, B. *et al.* Motor impairments in Chinese toddlers with autism spectrum disorder and its relationship with social communicative skills. *Front. Psychiatry* **13**, 938047 (2022).

Author contributions

Q.Z.: Conceptualization, Methodology, Software, Investigation, Writing—Original Draft; Y.L.: Methodology, Software, Writing—Original Draft; H.D.: Methodology, Software, Writing—Original Draft; Supervision, Writing—Review & Editing. Y.T.: Conceptualization, Resources, Supervision, Writing—Review & Editing.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Y.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023