2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)

# A Tale of the Social-side of ASD

Shameem Ahmed*
*Computer Science Department*
*Western Washington University*
Bellingham, USA
shameem.ahmed@wwu.edu

Md. Forhad Hossain*
*Computer Science Department*
*Missouri State University*
Springfield, USA
forhad.hossain.cse@gmail.com

Kurt Price
*Computer Science Department*
*Western Washington University*
Bellingham, USA
pricek21@wwu.edu

Cody Pranger
*Computer Science Department*
*Western Washington University*
Bellingham, USA
prangec3@wwu.edu

Md. Monsur Hossain
*We Work*
Seattle, USA
monsur.hossain@wework.com

Moushumi Sharmin
*Computer Science Department*
*Western Washington University*
Bellingham, USA
moushumi.sharmin@wwu.edu

*Abstract*— **Individuals with Autism Spectrum Disorder (ASD) experience difficulty in social interaction resulting in a lack of close social connections. Online social networking sites show promise in enabling individuals with ASD to make connections as the media of communication is often void of complexities present in real-life situations (e.g., real-time interpretation of complex social and emotional cues). We conducted a systematic content analysis of the Autism and Asperger's subreddits to understand user types, interaction patterns and trends, and user's experience, expectations, and unmet needs pertaining to these online communities. Our analysis revealed several interesting findings including the presence of 15 different topic areas, the tone and preference of the community, the difference in communication style between individuals with ASD and other related members (e.g., parents, therapists). Our analysis signified a need for an accessible and inclusive social networking site design that will cater to the specific needs of this community.**

**Keywords— ASD, Autism, Asperger's, Online Social Network, Reddit, Content Analysis, Visual Analysis.**

## I. Introduction

Autism Spectrum Disorder (ASD) affects one's ability to interact with others in social situations and affects the way people with ASD interact and communicate with others. Individuals with ASD often experience difficulty with understanding non-verbal cues in conversation in social environments, resulting in a lack of close social connections (friends). Online social networking sites (e.g., Facebook, Twitter, and Reddit) show promise in enabling individuals with ASD to make connections with others as the media of communication is void of complexities present in real-life situations. For example, online social media frees individuals with ASD to recognize emotion and social cues and come up with an appropriate response in real-time. However, social networking sites are built around with an expectation that people will follow the traditionally accepted norms of communication. The difference in communication style and capabilities in neuro-typical individuals and individuals with ASD indicates that careful design is needed to make social networking sites inclusive and accessible to everyone including individuals with ASD.

The overarching goal of our research is to develop a set of actionable guidelines for designing inclusive social network sites, especially focusing on individuals with ASD. As the first step in this process, we conducted a systematic content analysis of the Autism and Asperger's subreddits of Reddit, a popular social networking site. This paper reports the findings of our systematic content analysis. More specifically, in this phase of research, our goal is to understand who are the typical users involved in these communities, in which ways these users are utilizing social networking sites, what motivates them to engage in these communities, and their experience, expectations, and unmet needs pertaining to these communities.

Research on health-related information seeking and sharing is gaining popularity as many users are utilizing various online communities for their healthcare needs. Burton et al. reported that autism is one of the top health-related information that people seek and/or discuss in social media [1]. At the same time, Ammari et al. reported that although parents of children with special needs rely on social media (e.g., Facebook and Yahoo! groups) for accessing information and getting social support [2, 3], social media, in general, failed to connect these families. Research also suggests that individuals with ASD and/or families with children with ASD still face social stigma and may not feel comfortable to disclose their/their children's condition to others in real-life or online communities openly [29].

To understand how individuals with ASD or people related to them (e.g., parents, siblings, friends, therapists) are utilizing online social communities, we analyzed two autism-related subreddits in Reddit where users interact anonymously and share personal information without revealing any details about themselves. By requiring minimal information to create an account and participate in the community, Reddit ensures users' privacy and protects their identities. We conducted content and visual analytics to gain deeper insight about the users, usage pattern, and the role of the platform in supporting user interaction. To the best of our knowledge, our paper offers the first such systematic content analysis of ASD-related interaction data available on two subreddits (Autism and Asperger's).

Our analysis revealed several interesting findings including how people self-identify and disclose some personal information using badges (e.g., parent of an autistic child, autistic adult), the different types of posts that get attention from the community (celebration of milestones), tone and language used (e.g., people-centric vs. disease-centric, first-person vs. third-person experience sharing), and communication focused on a limited set of topics (e.g., success stories and diagnosis information seeking). Visual analytics revealed interesting patterns and trends, which indicated that people with similar badges interact more with each other and there is a clear preference for celebratory posts among users in the Autism community but not in the Asperger's community. Our analysis signifies a need for better-designed online communities that will cater to the specific needs and expectations of individuals with ASD and individuals related to them.

*Authors contributed equally to this work

978-1-7281-2607-4/19/$31.00 ©2019 IEEE
DOI 10.1109/COMPSAC.2019.00098

## II. RELATED WORK

The increased availability of information on the Internet is raising public-awareness related to ASD and is providing knowledge and support for concerned parents and individuals related to them [12]. Online discussion forums work as supportive communities, helping to learn about strategies, diagnostic opportunities, and resources pertaining to ASD and reducing the isolation experienced by parents and individuals with ASD [12].

### *A. ASD and Social Media*

Recently social media became popular among researchers to discover new insights about ASD. For instance, Beykikhoshk et al. showed that Twitter could be used as a data-mining source to learn about ASD community – their behavior, concerns, and needs [11, 14, 20]. Saha et al. systematically analyzed the activity of autism blogger communities on Twitter and found that these communities provide extensive social support to its community members [4]. Ben-Sasson and Yom-Tov focused on building machine-learning models to classify a child's risk of ASD based on parent's online narrative about their children [19].

Hong reported that over-reliance on a small set of people, typically a primary caregiver, is not only a barrier to the independence of the young adults with ASD but also a burden on the caregivers [8]. Hong focused on social networking services that help young adults to collect information or to get advice regarding complete and subtle real-world situations an adolescent or adult with ASD may encounter [8]. In addition, Hong et al. found that technology-supported communication particularly strengthened the relationship between individuals with autism and extended network members, mitigating concerns about over-reliance on primary caregivers [9].

Moorhead et al. discovered twelve limitations of social media for health communication [27]. One of them is a lack of reliability, confidentiality, and privacy [27]. Their findings were based on a systematic literature review they conducted on 98 original studies to understand how users (e.g., the public, patients, and healthcare professional) use the social media for health communication [27]. Beykikhoshk et al. exploited large-scale ASD-related data collected from Twitter [20], however, did not consider the relationship of a poster with autism. Moreover, the limitation of tweet-size (maximum of 140 characters) hindered the content specific data mining and analysis. In addition, several of the prior studies related to ASD and social media utilized only a limited set of self-reported ASD-related data for analysis.

In this paper, we systematically analyzed a large-scale ASD-related data available in a social networking site. To collect this data, we chose Reddit, more specifically two subreddits, for the following reasons: (a) Reddit is very popular social media (330 million users worldwide[10]); (b) Reddit has been extensively used for ASD-related discussion (53k+ users in Autism and Asperger's subreddits); (c) Unlike Twitter, there is no restriction on post size; (d) Unlike other popular social media, Reddit allows its user to share information (e.g., health-related private information) keeping their identity secret; and (e) Reddit enables researchers to determine user's relationship with autism (e.g. individuals with ASD, parents of a child with ASD) using it's badge system.

### *B. ASD and Visualization*

Visual information has emerged as a critical form of support for children with ASD due, in part, to documented strengths of this form for information processing [15]. Further, the use of technology has the potential to alleviate some apprehension experienced by many children with ASD when interacting with people [16]. According to Grandin [17], there are three types of autistic/asperger's cognitive types: Photo-realistic visual thinkers, Pattern thinkers, and Word-fact thinkers. In a different thread of research, Hailpern et al. developed a real-time voice visualization system (Vocsyl) to help children with ASD and speech delays to learn multisyllabic word production [18, 30]. However, there is limited research on how to utilize visual analytics to understand the users of online ASD communities, their experiences, needs, and challenges.

Prior research pertaining to ASD and visualization is based on the physical interaction of an individual with ASD with the real world (e.g., how they communicate with others by using their hands, eye movements, and multisyllabic speech). On the contrary, in this paper, we present visualizations of social media interactions of individuals with ASD (or someone related to that person) with others about their experiences, concerns, and thoughts on ASD. These visualizations provide insight pertaining to ongoing interaction in these communities and why and how individuals with ASD and people related to them are utilizing these communities.

## III. RESEARCH METHODOLOGY

For our research, we focused on Reddit, a social media platform where subscribed users can submit contents. These contents are organized into sub-communities called "subreddits." Posts are given scores, called "karma", by the community through "upvotes" and "downvotes", with upvotes adding a point to the post's score and vice versa. In this research, we mainly focused on two subreddits titled "/r/autism" and "/r/aspergers." Users on these subreddits share their stories and experiences as well as provide advice regarding different aspects of Autism and Asperger's. During the data collection phase, the Autism subreddit had 21,725 subscribed users and the Asperger's subreddit had 32,224 subscribed users.

### *A. Data Collection*

We used PRAW (The Python Reddit API Wrapper) [21] for collecting data from the two subreddits mentioned above. We used a MySQL database to store, access, and analyze the data. The dataset includes but is not limited to user's posts, comments, upvotes, downvotes, post dates, comment dates, and user information (limited). The Reddit API limits crawling historical posts on any subreddit to most recent 1,000 posts. We utilized different types of query to collect a robust dataset. For example, we collected most recent, hot, new, rising, controversial, and top posts from the previous week. Our analysis revealed that, on average, the total post frequency per

week on both subreddits are less than one thousand. To overcome Reddit API's constraint pertaining to the collection of historical data, we collected data for several weeks using the PRAW API. We started collecting data on November 2017 and continued data collection for the next eight weeks. We also collected all comments associated with those posts. We extracted all users/username from all posts and comments. These lists of users/username enabled us to collect all posts submitted by these users on the "r/autism" and "r/aspergers" subreddits (using each user post query of the PRAW API). Our final corpus contained a total of 20,217 posts (6,398 from "r/autism" and 13,819 from "r/aspergers") and 135,704 comments (38,309 from "r/autism" and 97,395 from "r/aspergers").

### *B. Data Preprocessing*

Before training our model, we transformed the raw data collected from two subreddits into our desired format. After converting all words to lowercase and then tokenizing those words, we removed stop words and irrelevant words (e.g., HTTP, WWW). To remove stop words, we used python stop-words library for the English language [22]. Then we removed all single-character words and numbers. However, we kept words that contain numbers. Next, we Lemmatized all the words in documents by using WordNetLemmatizer library [24]. To make the topics more meaningful, we added bigrams and trigrams to the documents (only ones that appear 15 times or more). Then we removed all extreme and rare words by filtering out the words that appeared less than 20 times in the documents or appears in more than 45% of the documents. Our final corpus included 12,605 unique tokens in 18,185 posts.

## IV. Results

In this section, we discuss findings pertaining to user distribution, their interaction with the community, post types, patterns and the tone and nature of these communities.

### *A. Distribution of Users*

The Autism subreddit had 21,725 subscribed users and several badges (provided by the subreddit) that these users could select to represent themselves. These badges appeared next to a user's name on all of their posts. These badges include Parent of Autistic Child (PAC), Autistic Adult (AA), Parent of Autistic Children (PACH), High Functioning Autism (HFA), Autistic Parent of an Autistic Child (APAC), Aspergers (AS), PDD-NOS/Aspergers (PAS), Moderator (MO), Sibling of an Autistic Individual (SA), Autism (AM), Seeking Diagnosis (SD), Educator (ER), Friend of Person with Autism (FPA), Parent of Child with PDD-NOS (PCP), Carer of a Child with Autism (CCA), NT, PDD-NOS (PN), Autistic Parent of Autistic Children (APAC), and Sibling of someone with PDD/NOS (SSP).

Fig. 1 shows the distribution of users who choose to identify themselves with a badge. One interesting observation is that 29% of the users of these subreddits do not identify as individuals with autism. Rather, they are related to or close to someone in the spectrum.

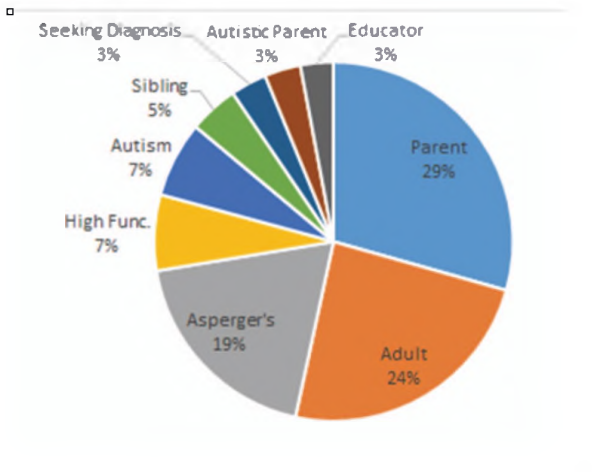


Fig 1. Distribution of self-identified users (based-on badges)

However, the majority (67%) of the users who self-identify with a badge identify as being on the spectrum, ranging from High Functioning Autism to Autistic Parents of Children with Autism. In contrast, the Asperger's subreddit, which had 32,224 users, does not provide this capability of self-identification, limiting our ability to classify its users into specific categories.

### *B. Categorization of Posts*

Analyzing all contents created by the communities may reveal interesting information about users' experiences in the communities and nature of the community. While useful, such analysis is often challenging due to the design of the community. Good organization and categorization of posts and comments (topic modeling) may alleviate some of these challenges. For example, one can randomly pick N number of posts from each topic and perform a qualitative study on that. This is more feasible than performing a qualitative study on an entire unorganized dataset. To identify what types of topic members discussed in these communities, we utilized Latent Dirichlet Allocation (LDA) [26] for categorizing the post content, which resulted in 15 different topic categories. We used Gensim and Scikit-learn Python Library for training our model [23]. In the following subsections, we provide details about our approach of findings related to post content and post categorization.

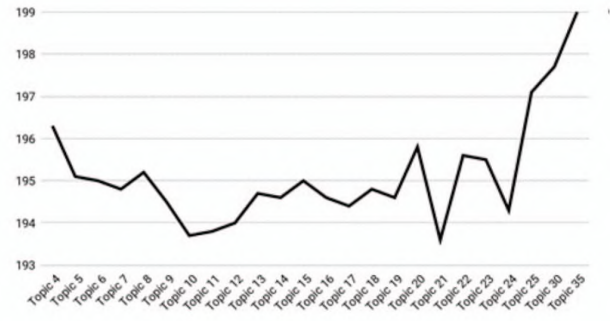


Fig 2: Perplexity for different category amount, K. Perplexity values are plotted in Y-axis while amount of category is plotted in Y-axis.

TABLE I: CATEGORIZATION OF POSTS PER TOPIC

| | Main Topics Discussed | Number of posts | Example Posts with Title (posts are included verbatim) |
|---|---|---|---|
| Topic 1 | Expressing feelings (e.g., depression, anger, nonplussed, self- harm) | 125 | A bunch of my friends have a habit of pointing out my social errors... even if they were harmless. ……..<br>I know I'm weird, I know I'm awkward. Please stop pointing it out. :( |
| Topic 2 | Shared links to different blogs | 82 | http://latrobepsy.az1.qualtrics.com/SE/?SID=SV_3mkhnrHbg2LXz9j Thanks, the more studies the better. |
| Topic 3 | Intellectual ability of ASD | 32 | Who here knows their IQs? Is it above average? I know that higher IQ is a common result of AS. Is there anyone here with AS and only average or below average IQ? My IQ is around 140. |
| Topic 4 | Movies, TV shows, ASD related documentaries, music, YouTube videos. | 50 | I just found a short movies about a robot's struggle to make friends. https://www.youtube.com/watch?v=zeU68vY_vDE Any others similar like this? |
| Topic 5 | Public school, kindergarten related experiences, problems, suggestion | 50 | Hello. I have a 4 year old son with autism. He's currently attending preschool at the local elementary, but my wife and I have been discussing whether to home school in the future. |
| Topic 6 | Different problems (e.g., meltdown, stress, eating disorder) | 161 | Along with my Asperger's, I also have selective eating disorder and can't eat fruits or vegetables without throwing up. |
| Topic 7 | Romantic partner, attraction to opposite sex, sexual relationships | 87 | Does anyone feel extremely sad when they look back at not being able to have the normal romantic relationships that everyone else got to have from 12-18? |
| Topic 8 | Religion, religious beliefs, Jesus, and gender difference (Male, female, Transgender). | 12 | I go to Christian school, have since 7th grade. Early on, I was a strong believer and as time went on I fell away and felt like religion was being shoved down my throat on a constant basis. This was further exacerbated due to the fact that im the only goth and metalhead in the entire school and i'm on the spectrum too. ... |
| Topic 9 | Job post, finding a job, asking help for job selection, job-related problem, tuition | 66 | I cannot find a job that suits me. I am currently a temp somewhere waiting to be hired, but was just informed that they probably won't be hiring the position I'm in now for a while (at least another 6 months) before they figure more things out. |
| Topic 10 | Aggressive behavior towards autistic people, being bullied, the negative attitude. | 18 | "Flashbacks of being bullied as a kid (and throughout school) are rushing back to my mind... http://nypost.com/2016/10/06/special-needs-boy-set-on-fire-by-bullies-he-thought-were-friends/" |
| Topic 11 | Fabric/clothing, wardrobe suggestion, skin touch sensitivity | 42 | This is why I buy my clothes at the Goodwill. |
| Topic 12 | ASD diagnosis, syndrome, affordable diagnosis, importance of diagnosis, and diagnosis result (Biased/ right/ wrong) | 91 | "I see here a lot of confusion about Asperger's syndrome, even though here, more than anywhere else, comes handy the ancient Greek aphorism ""know thyself""... |
| Topic 13 | Different newspapers, social media or blogs links about Autism or Asperger's | 68 | Michael Burry, M.D. (neurology) was one of the few people in the world who foresaw the sub-prime mortgage crisis. Michael Burry also learned that he had Asperger's. http://en.m.wikipedia.org/wiki/Michael_Burry |
| Topic 14 | Research links about Autism or relationship with Autism. | 66 | Why has it taken so long for Autism Speaks to announce vaccines do not cause autism? |
| Topic 15 | Problem with communication, their habits, speaking properly, problem related to speaking | 51 | Either I have no clue what someone is saying to me, or I know all too well what they are saying. I can't articulate myself very well so I don't know if I'm making sense here or not. |

*1) Identifying optimal number of content categories*

To identify the optimal number of topics, we used perplexity as a metric. Latent Dirichlet allocation (LDA) approach is computationally expensive. Therefore, instead of using the entire corpus, we randomly selected 50% of the documents as our data source. We split the data source into a training set containing 80% of the documents and a testing set containing the remaining 20%. We trained our model with a different category from k = 4 to k = 35 and compared the resulting performance. We observed that two values of k (10 and 21) gave us the lowest perplexity value. Moreover, our testing suggested that 10 to 21 topic categories might be optimal for our modeling (Fig. 2). Although perplexity values indicated an initial range of topics, a human evaluation was necessary for validating the accuracy of this categorization. We used different k values between 10 to 20 and observed the outcome using LDAVis library [25]. We started with k=10 and found that some topics can be further divided. We found that k = 15 provides the optimal outcome.

*2) Topic Distribution*

We trained our model for k = 15 topics. Our training involved 200 passes and 2,500 iterations to attain sufficient convergence. After training, we analyzed the category distribution and co-relation between each document in each category (see Fig. 3). LDA assumes that each document is a mixture of topics and distributes probability values over topics/categories for each document. We used the highest probability value over topics for each post and assigned them to that topic. This does not necessarily mean that a particular post only belongs to that particular topic. Fig. 3 shows further details which indicates that 2,615 posts were classified as topic 1 and 470 posts were classified as topic 8.

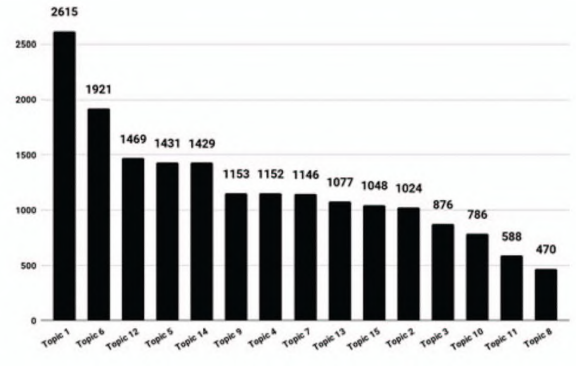


Fig 3: Distribution of post on different topics. X-axis shows topic categories and Y-axis shows number of topics associated with each category.

To further analyze each topic and how relevant posts are on each topic category, we examined each post that has been classified with more than 90% confidence level to a particular topic (see Table 1 for details about posts in each topic with example).

*3) Topic analysis by actual post*

To evaluate the performance of our model, we randomly selected 15 posts from each topic category that has been classified to that topic with a 90% confidence level. This gave us a corpus of 225 posts that the research team members manually analyzed. This additional human analysis helped us to validate the accuracy of automated categorization. Our human analysis confirmed that posts classified within the same category are very similar to each other. For instance, the two following posts were identified as topic 1.

*"Alone, stupid can't find meaning in my life"*

*"Why do they have to point out when I do something awkward?"*


Fig 4. Post contribution by individual users. The X-axis represents the number of posts a user made and the Y-axis represents how many users contributed to that number of posts.

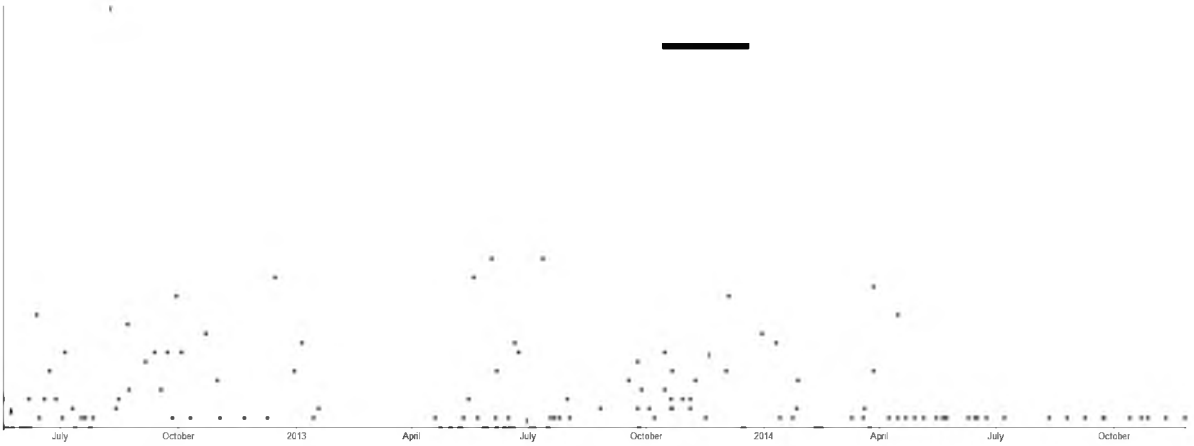

Fig 5. Timeline graph of the top poster's interaction in the community. This top poster submitted more than 140 posts in a two year period. X-axis represents post time and Y-axis represents interaction (upvotes and comments) associated with these posts. Posts with higher interaction are placed higher than posts with lower interaction.

In these two posts, users focused on expressing their feelings, sorrows, and frustration. Similarly, the following two posts were identified as topic 11, which are focused on clothing and fashions.

*"This is why I buy my clothes at the Goodwill."*
*"What kinda shoes do you wear?"*

### C. *Identification of Key Players*

In terms of posts, the Autism subreddit closely follow the 1% rule of participation in social media [13], which states that only a few users contribute the majority of the content. Fig. 4 reveals that there is a concentration of 30 users who contributed 14.73% of the content in the group. Interestingly, users who contributed the greatest number of posts are not necessarily the "key players." For instance, we found two users who contributed a significant number of posts (the rightmost on the X-axis of Fig. 4). Despite posting many contents, post of these two users rarely generated significant traction in terms of karma and comments. To understand the interaction between frequent/top posters and the community, we visualized all posts by the topmost poster (who created 140 posts) and the upvotes and comments received from the community (see Fig. 5).

Our visualization indicates that this top poster saw a decrease in the karma score of his/her posts over time. All but one of his/her posts failed to generate a lot of traction from the community and lasted for less than three days. Upon close inspection of the only outlier (a post that received much attention), we discovered that this is a post about a missing child with ASD, which made the thread extremely important to the community. This analysis highlights that the number of posts is a poor indicator for identifying key.

### D. *Identification of Top Posts*

Easy identification of top posts would enable new members of this group to learn about the culture, content, and tone of the community. Creating a temporal representation of all posts along with their upvotes and interaction (comments) may help to identify posts that were picked up (liked) by the community. This will also enable identifying patterns of posts and community sentiment.

As the first step in this direction, we visualized all posts of the group from the beginning of the group's history, ranking them based on their karma scores (see Fig. 6). Their shapes (square or circle) denote the time the posts were first made and the end of the post's last comment (lifespan). This allowed us to build an intuitive model of the interactions that are typically seen in this community and the response that they receive.

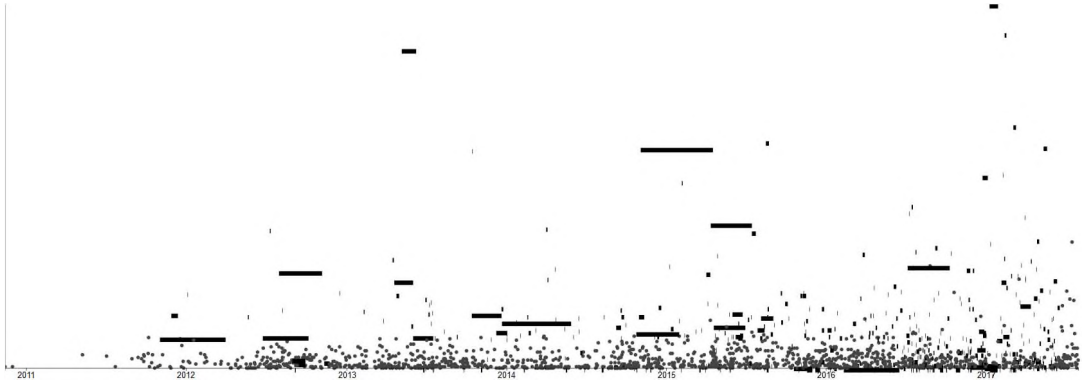

Fig 6. Interaction history of all posts in the autism subreddit between 2011 and 2017. The X-axis represents creation and interaction time of individual posts and the Y-axis represents the number of upvotes received by a post. Length of posts indicates the lifespan of a post (interaction with the community).

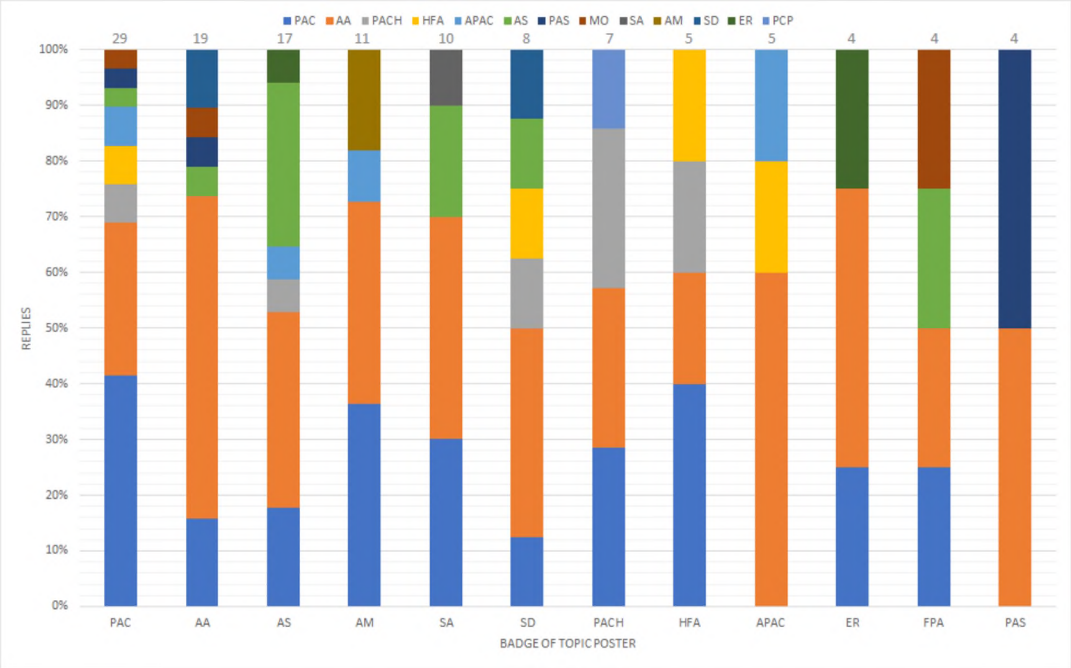


Fig 7. Relationship between posters and commenters using badge information in the Autism SubReddit. These normalized stacked bars highlight interaction among user-groups (based on badge). Each stacked bar in X-axis represents number of replies received from different user-groups. Y-axis represents relative contribution of different groups in reaction to posts submitted by posters belonging specific groups.

We identified the top three posts from Fig. 6 and closely analyzed their content and the comments. We learned that each of these top posts focused on celebrating some form of milestones, change in life, or other achievements. Fig. 6 also highlights that as the community grew in size, these outliers gained large karma scores that pushed them much further over the average posts. However, the growing size of the community did not affect the lifespan of average posts or even many of the more popular posts.

In contrast to the celebratory posts on the Autism subreddit, the top posts (in terms of upvotes) in the Asperger's subreddit tend to be rather fond of lamentation. Within the top 20 posts in this community, multiple threads can be seen highlighting many sources of stress and dissatisfaction within the Asperger's community, such as the use of "autistic" as a slur, to someone stressing about a recent diagnosis of Asperger's, etc. Of note, the sixth-most upvoted thread of the Asperger's subreddit denounces the celebratory threads as problematic and unhelpful to the Autistic community for distorting the public perspective of those in the spectrum. In contrast, such celebratory posts are widespread and well-liked in the Autism subreddit.

### *E. Identifying "Birds-of-a-feather"*

Our analysis reveals that users who use "same/similar" badge interact more frequently (comment on each other's post) than other users who use dissimilar badges. Fig. 7 indicates that users belonging to the same group (same/similar badge) interact more compared to users with different badges.

By analyzing the amount of interaction (comment volume) of different groups of users on topics posted by users with specific badges we learned that this affinity is noticeable in almost all groups of users. For example, the first bar represents the comment volume of different groups of users on topics posted by Parent of Autistic Child (PAC) and the distribution of comments (represented using colors) show that a significant amount of comments was posted by users with same badge (PAC). We observed a similar phenomenon in the second bar where threads initiated by Autistic Adults (AA) badge received the highest comments from the same group of users. This pattern is visible for almost all these bars, which indicates that users may benefit from novel organizations of posted contents which may facilitate accessing and interacting with posts created by specific groups of users.

## V. Discussion

### *A. Identification of Key Players is Complex*

In any online community, access to key players may be beneficial as they can help posts to get noticed and make the community aware of important information. In communities where a particular user's real-life identity is associated with his/her online account (e.g., Facebook and Twitter accounts of celebrities), it is easier to reach a large group of users by posts

promoted by these well-known key players. In contrast, Reddit does not allow users to associate their real-life identities with their online accounts. Our findings suggest that identifying key players within these subreddit communities is complex and may involve looking beyond the number of posts they make and consider a metric consisting of karma, number of replies, number of unique users commenting on their posts, post types, and post lifespan. Further research is needed to facilitate key player identification without changing the participation norms of these subreddits.

### B. *User Types and Their Needs Shape the Community*

Our analysis reveals that the Autism subreddit community preferred celebratory posts while the Asperger's community focused more on problems associated with everyday experience of individuals with autism. It is important to note that a large portion of the users of the Autism subreddit is parents, family members, and friends of individuals with ASD. Research suggests that parents or caregivers of autistic children experience unique stressors and that support groups for parents or caregivers may improve the quality of their lives [7, 8]. We can speculate that in their stressful lives, celebratory posts may offer a sense of hope, which may have contributed to the preference for celebratory posts. Asperger's subreddit, on the other hand, lacks any badge information. However, analysis of posts revealed that many members of this community are individuals with Autism (Asperger's) and top posts in this community focuses on challenges and experiences of everyday lives and how these can be improved.

### C. *Effective Identification of User Groups and Their Contents may Facilitate Interaction*

At present, there is no easy way to access contents posted by specific categories of users on Autism subreddit. Our analysis reveals that people who identify as a specific group have similar interests and/or common concerns and as such may benefit by easy access to contents created by "similar people." Self-Determination theory [28] suggests that people feel more empowered and motivated when they find other "similar" people experiencing similar situations. We speculate enabling users to connect to other "similar users" may provide a better experience in these online communities. In addition, providing easy access to contents posted by similar users may reduce search time for specific posts or topics.

## VI. Conclusion

Autism is a life-long journey and it not only affects the lives of individuals with Autism but also the lives of parents, family members, and caregivers [7, 8]. Online communities such as Autism and Asperger's subreddits can offer a platform where individuals with ASD and their friends, families, and caregivers can interact, share experience, and learn from each other, effectively creating a support network. In these communities, posts that highlight successes or milestones can show possibilities and strategies that worked to achieve a goal or address a challenge. In addition, such communities can work as a community knowledge-base where people can share information related to diagnosis, research, and available resources. Our research indicates that these communities interact with and react to posts differently and it is difficult to utilize known models of key players' identification (e.g., well-known celebrities in Facebook) or trending topics identification in such groups. Further research is needed to identify how to best design such online networking sites to provide a better experience to the users of these communities.

## References


[1] S. H. Burton, C. V Tew, S. S. Cueva, C. G. Giraud-Carrier, and R. Thackeray, "Social Moms and Health: A Multi-platform Analysis of Mommy Communities," *ASONAM*, 2013.

[2] T. Ammari, M. R. Morris, and S. Y. Schoenebeck, "Accessing Social Support and Overcoming Judgment on Social Media among Parents of Children with Special Needs," *Int. AAAI Conf. Weblogs Soc. Media*, 2014, pp. 22–31.

[3] T. Ammari and S. Schoenebeck, "Networked Empowerment on Facebook among Parents of Children with Special Needs," *CHI*, 2015.

[4] A. Saha and N. Agarwal, "Demonstrating Social Support from Autism Bloggers Community on Twitter," *ASONAM*, 2015, pp. 1053–1056.

[5] M. O. Mazurek, "Social media use among adults with autism spectrum disorders," *Comput. Human Behav.*, vol. 29, no. 4, pp. 1709–1714, Jul. 2013.

[6] A. Estes, J. Munson, G. Dawson, E. Koehler, X.-H. Zhou, and R. Abbott, "Parenting stress and psychological functioning among mothers of preschool children with autism and developmental delay," *Autism*, vol. 13, no. 4, pp. 375–387, June. 2009.

[7] L. A. Schieve, S. J. Blumberg, C. Rice, S. N. Visser, and C. Boyle, "The Relationship Between Autism and Parenting Stress," *Pediatrics*, vol. 119, no. Supplement 1, p. S114 LP-S121, Feb. 2007.

[8] H. Hong, "Specializing social networking services for young adults with autism," *CSCW Companion*, 2014, pp. 65–68.

[9] H. Hong, S. Yarosh, J. G. Kim, G. D. Abowd, and R. I. Arriaga, "Investigating the use of circles in social networks to support independence of individuals with autism," *CHI*, 2013.

[10] 71 Amazing Reddit Statistics and Facts (March 2018), https://expandedramblings.com/index.php/reddit-stats/

[11] A. Beykikhoshk and T. Caelli. "Data-Mining Twitter and the Autism Spectrum Disorder: A Pilot Study." *ASONAM, 2014*, pp. 349–356.

[12] C. J. Jordan, "Evolution of autism support and understanding via the World Wide Web," *Intellect. Dev. Disabil.*, vol. 48, no. 3, pp. 220–227, 2010.

[13] T. van Mierlo, "The 1% Rule in Four Digital Health Social Networks: An Observational Study," *J. Med. Internet Res.*, vol. 16, no. 2, p. e33, Feb. 2014.

[14] A. Beykikhoshk, O. Arandjelović, D. Phung and S. Venkatesh, "Overcoming data scarcity of Twitter: Using tweets as bootstrap with application to autism-related topic content analysis," *ASONAM, 2015*, pp. 1354-1361.

[15] N. J. Minshew, G. Goldstein, and D. J. Siegel, "Neuropsychologic functioning in autism: Profile of a complex information processing disorder," Journal of the International Neuropsychological Society 3, 1997.

[16] C. B. Baskett, "The effect of live interactive video on the communicative behavior in children with autism", Ph.D. thesis, UNC-Chapel Hill, 1996.

[17] T. Grandin, "How does visual thinking work in the mind of a person with autism? A personal account," *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 364, no. 1522, pp. 1437–1442, 2009.

[18] J. Hailpern, A. Harris, R. La Botz, B. Birman, and K. Karahalios, "Designing visualizations to facilitate multisyllabic speech with children with autism and speech delays," *DIS 2012*.

[19] A. Ben-Sasson and E. Yom-Tov, "Online concerns of parents suspecting autism spectrum disorder in their child: Content analysis of signs and automated prediction of risk," *J. Med. Internet Res.*, vol. 18, no. 11, 2016.

[20] A. Beykikhoshk, O. Arandjelovi, D. Phung, S. Venkatesh, and T. Caelli, "Using Twitter to Learn about the Autism Community," Social Network Analysis and Mining, 5(1), 22, 2015.

[21] Praw python Reddit API wrapper. https://github.com/praw-dev/praw

[22] Python Stop-words. https://pypi.python.org/pypi/stop-words

[23] Gensim - Topic modeling python library. https://radimrehurek.com/gensim/

[24] Wordnet - WordNet Lemmatizer. http://www.nltk.org/_modules/nltk/stem/wordnet.html

[25] pyLDAvis - Python library for interactive topic model visualization. https://github.com/bmabey/pyLDAvis

[26] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. Latent Dirichlet allocation. J. Mach. Learn. Res. 3 (March 2003), 993-1022.

[27] S. A. Moorhead, D. E. Hazlett, L. Harrison, J. K. Carroll, A. Irwin, and C. Hoving, "A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication.," *J. Med. Internet Res.*, vol. 15, no. 4, p. e85, Apr. 2013.

[28] Deci EL, Ryan RM. Intrinsic motivation and self-determination in human behavior. New York: Plenum Publishing Co.; 1985.

[29] Redacted for double-blind review.

[30] J. Hailpern, K. Karahalios, L. DeThorne, and J. Halle, "Vocsyl: Visualizing syllable production for children with ASD and speech delays." ASSETS, 2010. pp. 297-298. .