What Snippets Feel: Depression, Search, and Snippets

Ashlee Milton
ashleemilton@u.boisestate.edu
PIReT – People and Information Research Team
Boise State University, Boise, Idaho, USA

Maria Soledad Pera
solepera@boisestate.edu
PIReT – People and Information Research Team
Boise State University, Boise, Idaho, USA

ABSTRACT

Mental health disorders (MHD) is a rising, yet stigmatized, topic in the United States. Individuals suffering from MHD are slowing starting to overcome this stigma by discussing how technology affects them. Researchers have explored behavioral nuances that emerge from interactions of individuals affected by MHD with persuasive technologies, mainly social media. Yet, there is a gap in the analysis pertaining to search engines, another persuasive technology, which is part of their everyday lives. In this paper, we report the results of an initial exploratory analysis conducted to understand the sentiment/emotion profiles of search engines handling the information needs of searchers with MHD.

CCS CONCEPTS

• **Information systems** → *Sentiment analysis.*

KEYWORDS

Mental Health, Search Engines, Snippets, Emotion

1 INTRODUCTION

Persuasive technologies can change the behaviors or attitudes of individuals [9], but not all people are affected in the same manner. Consider people suffering from mental health disorders (**MHD**). They tend to be more sensitive or easily influenced, making it natural to think they would interact with and be affected by persuasive technologies in different manner than average individuals. Mental illness is a rising issue in modern society. Over the last couple of years people have started to more openly discuss mental illness and how it effects their lives [12]. As this discussion continues, it leads to the question: are there consequences that can unknowingly occur as a result of individuals suffering from MHD engaging with persuasive technologies?

Search engines (**SE**) are an example of an ubiquitous persuasive technology. Yet, there is very little information regarding how users with MHD interact or are effected by resources accessed via SE. Consider Figure 1, which captures a snapshot of snippets generated by Google for the query "waste of space", a phrase a person with depression may say. Among the resources retrieved we see mostly dictionaries, which on the surface seem pretty benign. However when examining the snippets of the resources we see phrases like, "worthless person", "He's a complete wast of space", "fat bastard", "goddam wast of space", and "I'm just a waste of space". Now imagine being someone battling a MHD like depression, how do you think being exposed to these phrases would effect you? As reported by the National Institute of Mental Health, 1 in 5 adults in the united states suffer from a MHD [13]. With millions of individuals turning to SE regularly [1, 5], SE users affected by MHD become a large population, which makes it imperative to explore how this diverse group of users interact with and are effected by SE, so we may better support their online quests for information.

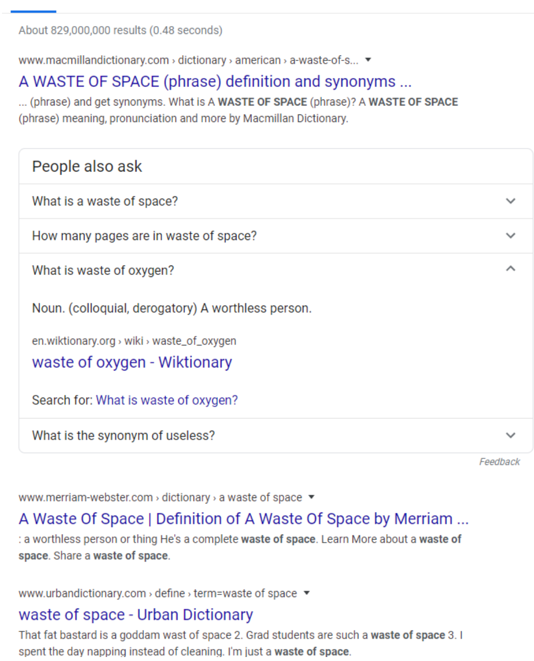

Figure 1: Google's SERP for the query "waste of space"

To set the ground work necessary to address this important need, we focus our initial exploration on search engine results pages (**SERP**) and feelings. Specifically, we examine the feelings that emerge from SERP, as they present a first impression of the type of resources users will be exposed to, in response to their queries. To do so, we follow the framework set forth by Kazai et al. [11] for exploring the sentiment/emotion profiles of SE responding to queries crafted by traditional searchers. In our case, we compare and contrast profiles generated by examining snippets of resources retrieved in response to inquires associated to individuals with MHD. Given that the interactions of people with MHD and SE are not available for analysis, we built a synthetic query log that mimics these interactions. For this purpose, we use posts from Reddit. Given that we take advantage of this plarform to create our synthetic queries and that the age range for 90% of Reddit users is 18-49, the center of attention for our analysis are adults. Further, we posit that the degree to which sentiments/emotions effect users with MHD depends on the kind of MHD they have. Thus, to narrow scope and not overgeneralize, we only explore depression and anxiety. In fact, from here on, when we use MHD we only refer to depression and anxiety.

We describe below the empirical analysis we conducted to understand the intensity in which feelings are present in snippets responding non-traditional users. Outcomes from this analysis will act as a springboard for future work in SE design to support individuals affected by diverse MHD across ages.

2 RELATED WORK

From a persuasive technology perspective, MHD literature is focused on the social media domain. Mainly, the ability to identify users with depression from social media posts [6, 7, 14, 15, 18]. Depressed users have not been the only ones considered, as researchers have also studied linguistics of social media posts by users with other MHD, such as schizophrenia [3]. Some of the most common attributes examined in studying social media in regards to users with MHD are the vocabulary and syntax of posts, as well as interactions with the platforms themselves (e.g., number of posts and retweets). Findings from the research conducted thus far not only highlight the vocabulary and syntax used by MHD individuals online but also the trends in interactions with other people on social media platforms and with the platforms themselves [19].

Research exploring the relationship between MHD and SE is in its infancy. From a user perspective, Campbell et al. [4] discuss help seeking behavior (i.e., looking for resources to understand and help with MHD) of users with MHD. Zhu et al. [24], on the other hand, use query logs from a university web server to predict users suffering from depression. Similarly, Zaman et al. [23] identify users with self-esteem issues from user-provided Google search histories. Xu et al. [20] instead evaluate the degree to which mood influences users' interactions with SE. From a resource perspective, some researchers have looked at the emotion, sentiment, and opinion emerging from Web resources [8, 11], but only those retrieved for traditional users. Unfortunately, none of the aforementioned works shine a light on the potential that SE' responses (i.e., resources retrieved) have to alter the mental well-being of users with MHD.

In our work, we take a first step towards addressing the gap we see in the literature by exploring what snippets feel when reacting to users with MHD, as to determine what feelings are being push onto this population that is already struggling with their own emotions.

3 DATA AND METHOD

In our exploration, we use the data and method below.

Data. Query logs from mainstream SE are seldom available for research and, to the best of our knowledge, non-existent from searchers with MHD. Consequently, to enable sentiment/emotion exploration we created two synthetic query logs: **TQL** and **MQL**, which emulate <query, SERP> pairs generated by traditional searchers and individuals affected by MHD, respectively.

For **TQL**, we use Yahoo Webscope's search query tiny sample [21], which includes 4,458 queries (1,211 unigrams; the remaining n-grams). Each query is associated with the corresponding SERP (top-10 resources, as users do not often go past the first page when looking at resources [16]), retrieved using Google API. For **MQL**, we use 4,418 synthetic queries (1,200 unigrams, the remaining n-grams to follow the query distribution of **TQL**) which we generate from Reddit posts. For each query, we also include the corresponding SERP generated using Google API. (Note that for **MQL** and **TQL** we keep only the resources that lead to snippets in English.)

In creating synthetic queries from searchers with MHD, we turn to Reddit, since its posts capture the language and topics used by our target population in an online forum environment, and it has been used in other MHD studies, but in the social media domain [7]. Specifically, we use 2,600 posts extracted from 10 subreddits, including *r/getting_over_it*, *r/depression*, and *r/suicidewatch*. We selected these subreddits as they have users that self-identify as having MHD. From the aforementioned posts we extract the 4,418 most frequent n-grams ($1 \leq n \leq 4$), using NLTK noun phrase chunking, which we use as synthetic queries for **MQL**.Topics and language across subreddits related to MHD greatly vary. Thus, we grouped <query, SERP> pairs in **MQL** into 3 categories based on the levels of severity of MHD [13, 25]: (1) **MQL-M** refers to queries from subreddits that have *mentions* of MHD but MHD is not the only focus, (2) **MQL-E** identifies queries from subreddits that are *explicitly* for users and topics of MHD, and (3) **MQL-S** captures queries from subreddits focused on topics of *self-harm and suicide*.

Method. To build sentiment/emotion profiles we follow the framework in [11]; using **MQL** and **TQL** in lieu of SE query logs. We create a vector capturing the sentiment/emotion of each snippet in **MQL** and **TQL** by averaging the sentiment/emotion vectors for each word in the snippet. For sentiment, we use SentiWordNet [2], which represents words as vectors of *positive*, *negative*, and *objective*. For emotion, we use EmoLexData [17], which represents words as vectors of *afraid*, *amused*, *angry*, *annoyed*, *dont_care*, *happy*, *inspired*, and *sad*. Words that do not appear in the lexicons are set to 0 on all vector elements, except objective and dont_care, which are set to 1.

4 ANALYSIS

We explore **TQL** and **MQL** profiles from different perspectives.

By Resources. We first average the sentiment/emotion vectors representations of all resources in SERP from **MQL** and **TQL** (rows 1 -5 in Table 1). Objective is the predominant sentiment, whereas the emotions with the highest scores are dont_care and happy. We notice some significant differences between the profiles generated for **TQL** and **MQL**. The sentiment vector of **MQL** is less objective than its counterpart from **TQL**, whereas **MQL**'s emotion vector includes higher scores for angry, annoyed, inspired, and sad. The emotion vector of **MQL** has lower scores for amused and dont_care than **TQL**'s. When exploring the profiles of **MQL**'s categories, we observe that the emotion profile of **MQL-M** yields scores that are lower for afraid, slightly higher for happy, and remain the same for sad when compared to the emotion profile of **TQL**.

By Queries. We also average sentiment/emotion profiles on a per query basis (rows 6-10 in Table 1). We observe a similar pattern as the one discussed for sentiment/emotion profiles generated at result level. When digging into categories within **MQL** a few differences do emerge. The emotion profile of **MQL-M** has a lower value for afraid than **TQL**, but now happy is aligned with the profile of **TQL**. The emotion profile of **MQL-E** has also changed, with amused and sad scores now being attune with **TQL**'s.

By Top-Ranked Result. As top-ranked resources in response to queries are the first users encounter on a SERP, we are interested in the sentiment/emotion that emerges from them. From profiles reported in rows 11 - 15 of Table 1, we observe that the sentiment of top-ranked resources remains consistent with both previous analysis. However, the emotion profile of **MQL** has changed, as it now

Averaged by	Row	Source	Sentiment			Emotion							
			Pos	Neg	Obj	Afraid	Amused	Angry	Annoyed	Dont_care	Happy	Inspired	Sad
Resources	1	**TQL**	5.547	3.330	91.123	3.560	6.561	4.913	4.584	43.853	25.360	6.555	4.614
	2	**MQL**	6.032β	4.170β	89.798β	3.511	6.317β	5.294β	4.889β	40.572β	25.478	9.113β	4.828β
	3	**MQL-M**	6.131β	4.030β	89.839β	3.307β	6.275β	5.358β	4.918β	40.713β	25.609α	9.155β	4.666
	4	**MQL-E**	6.007$\beta\gamma$	4.279$\beta\delta$	89.714β	3.634δ	6.404α	5.239β	4.852β	40.411β	25.391	9.233β	4.835$\beta\gamma$
	5	**MQL-S**	5.962$\beta\delta$	4.200$\beta\delta$	89.839β	3.589δ	6.273β	5.284β	4.895β	40.589β	25.437	8.958$\beta\eta$	4.974$\beta\delta$
Queries	6	**TQL**	5.543	3.330	91.127	3.561	6.552	4.897	4.577	43.921	25.325	6.557	4.609
	7	**MQL**	6.057β	4.227β	89.716β	3.535	6.319β	5.345β	4.898β	40.604β	25.367	9.069β	4.863β
	8	**MQL-M**	6.130β	4.028β	89.842β	3.307β	6.275α	5.360β	4.918β	40.708β	25.612	9.153β	4.667
	9	**MQL-E**	6.002β	4.277$\beta\delta$	89.721β	3.634δ	6.400	5.247β	4.852β	40.403β	25.402	9.231β	4.831
	10	**MQL-S**	5.961$\beta\gamma$	4.198$\beta\gamma$	89.840β	3.588δ	6.273β	5.286β	4.896β	40.589β	25.439	8.957β	4.972$\beta\gamma$
Top-Ranked	11	**TQL**	5.679	3.251	91.070	3.463	6.382	4.908	4.524	43.867	25.387	6.557	4.911
	12	**MQL**	5.993β	4.254β	89.753β	3.285	6.342	5.545β	5.108β	40.081β	26.033α	8.640β	4.966β
	13	**MQL-M**	6.237β	4.035β	89.727β	3.074α	6.466	5.588β	5.143β	40.116β	26.178α	8.551β	4.884
	14	**MQL-E**	5.845γ	4.413$\beta\delta$	89.742β	3.243	6.218	5.507β	5.039β	40.346β	25.922	8.822β	4.903
	15	**MQL-S**	5.901γ	4.311$\beta\gamma$	89.788β	3.527γ	6.343	5.540β	5.140β	39.795β	26.000	8.552β	5.104

Table 1: Sentiment/emotion profiles inferred from TQL and MQL (SERP generated using Google), where vector sum to 100. For statistical significance (two tail t test), α p < 0.05 and β p < 0.01 with respect to scores generated from TQL for the respective average type. Further, γ p < 0.05 and δ p < 0.01 indicate significance with respect to scores generated from MQL-M for the respective average type. Lastly, η p < 0.05 and p < 0.01 for to scores generated from MQL-M and MQL-E, respectively.

Averaged by	Unigram		N-gram	
	TQL	**MQL**	**TQL**	**MQL**
Resources	6.087	6.100	6.713	6.398β
Queries	6.123	6.141	6.708	6.397β

Table 2: Amused scores for unigrams vs. n-grams. For statistical significance (two tail t test), β p < 0.01 w.r.t. TQL for the respective average type and query length.

Query Type	Source	Positive	Angry	Happy
Unigram	**TQL**	5.416	4.609	23.467
	MQL	6.403β	7.234β	24.677α
N-gram	**TQL**	5.780	5.016	26.088
	MQL	5.840	4.915	26.539

Table 3: Positive, angry, and happy scores for top-ranked results. Statistical significance (two tail t test), β p < 0.01 w.r.t. TQL for the analogous query length.

has higher scores for angry, annoyed, happy, and inspired, as well as a lower score in dont_care when compared to **TQL**'s emotion profile. **MQL**'s emotion vector at the top-ranked result has amused and sad scores inline with **TQL**'s, yet the happy scores increased. When examining **MQL-M**, **MQL-E**, and **MQL-S**, we discovered a change from their sentiment profiles. The sentiment profiles of **MQL-E** and **MQL-S** no longer have significant difference for positive sentiment over **TQL**'s profile. The emotion profile of **MQL-M** has a lower score for afraid and a higher score for happy than **TQL**, but the difference for amused across the two profiles is no longer significant. Similarly, the profiles of **MQL-E** and **MQL-S** have amused and sad consistent with **TQL**.

Unigram vs N-grams. We explored the variations, if any, exhibited in profiles emerging from unigram and n-gram queries. Among the most interesting results, captured in Table 2, we see that the emotion vector of **MQL** has a lower amused score then **TQL** when looking at n-grams, while with unigrams, **MQL** profile is aligned with **TQL**'s. The most notable differences between unigrams and n-grams with regard to top-ranked resources are summarized in Table 3. **MQL** n-grams are alike in positive, angry, and happy with **TQL** n-grams, whereas **MQL** unigrams result in significantly higher values in positive, angry, and happy when compared to **TQL** unigrams.

MHD Categories. As MHD have varying levels of severity, we look into the profiles emerging from **MQL-M**, **MQL-E**, and **MQL-S**, reported in rows 3-5, 8-10, and 13-15 of Table 1, to determine if SE respond abnormally to these categories. Most significant differences across categories occur when comparing **MQL-M** with respect to **MQL-E** and **MQL-S**. The sentiment profiles generated from **MQL-E** and **MQL-S** are less positive and more negative than **MQL-M**'s, except when averaging by query, where only **MQL-S** is less positive, i.e., all categories are less positive overall. In terms of emotion profiles, the only significant differences across the 3 categories are for afraid, inspired, and sad. **MQL-E** and **MQL-S** emotion vectors have higher scores in afraid than **MQL-M**'s when averaging all result and by query, but only **MQL-S**'s has higher scores for afraid when averaging by top-ranked resources. However, only averaging all result shows both **MQL-E** and **MQL-S**'s emotion profiles having larger scores for sad than that of **MQL-M**. When averaging by query, the emotion vector of **MQL-S** has a higher score for sad than **MQL-M**'s. **MQL-S**'s inspired score is lower than both **MQL-M**'s and **MQL-E**'s when averaging all resources.

Safe Search. It is worth mentioning that we also examined sentiment/emotion profiles emerging when using Google SafeSearch to generate the corresponding SERP for each query in **TQL** and **MQL**. We do so, based on our interest in examining how safe search functionality handles inquiries related to MHD. Upon initial exploration of sentiment/emotion profiles we did not observe statistically significant changes with respect to the profiles surfacing without the use of safe search. For this reason, and due to space constrains, we exclude detailed findings from this analysis.

5 DISCUSSION

Results from our analysis reveal significant changed between the sentiment/emotion profiles of SE when responding to queries from users suffering from MHD, when compared to traditional searchers. Most notably, the sentiment/emotion profiles originating from **MQL** express more sentiment and emotion overall. Our findings concur with the results reported by Kazai et al. [11] in regards to web resources having a high proportion of objective sentiment and dont-care, happy, and inspired emotions, yet, we witness more changes in positive and negative sentiment, as well as angry and annoyed emotions. The increase towards more polar sentiments and negatively charged emotions is worrisome. It shows that users with MHD are encountering resources that have the potential to negatively effect their mental health and may not provide them with objective information to make decisions.

The evident changes in responses to users with MHD in regards to anger and annoyance are concerning, especially for top-ranked resources. SE resources conveying such emotions to users with MHD on a daily basis could have a negative effect on their mental health [10]. The presence of anger and annoyance is also problematic when we look at the emotion profile of **MQL-S**, where we also notice a higher score in sadness based on resources retrieved for queries from traditional users and users with less severe MHD. Recall that synthetic queries in **MQL-S** contain language from posts related to suicide and self-harm. While resources being higher in scores for angry, annoyed, and sad when responding to users with severe MHD compared to traditional users is anticipated, being exposed to these emotions may have a negative affect on users mental health. There is also a surprising result when investigating **MQL-E**, as when this category is averaged by query the value for amusement increases and the one for sadness decreases. This finding prompted us to look into the queries that caused such changes. We found that queries like "miserable", "disorder", "no fun", "new doctor", and "bed", lead to profiles with high amusement and low sadness scores. Out of context, some of these queries could be perceived as jesting *you're "no fun"*, but in the context of MHD they take on different connotations: *having "no fun"* or *not being able to get out of "bed"* are realities of users of suffering from MHD and are generally not looked at in an amused light.

When exploring the sentiment/emotion profiles of **MQL** per category, we discovered that **MQL-E**'s and **MQL-S**'s profiles are less positive and more negative, afraid, and sad than the profile of **MQL-M**, showcasing that different levels of severity of MHD changes the sentiment/emotion profiles of content retrieved by SE. While this pattern remains true for top-ranked resources on the SERP, only **MQL-S**'s profiles show more of the afraid emotion. This finding is interesting, given that we identify less significant changes overall when looking at just the top-ranked result in response to all queries. The change in profile for top-ranked resources could be due to Google pushing help hot-lines to the top-ranked resources when dealing with queries explicitly related to suicide and self harm, however, more investigation is needed to confirm this theory.

Findings emerging from our exploration reveals there is much work to be done for SE to accommodate users with MHD. To start, SE would require the ability to recognize this compromised population. Currently, SE like Google and Bing acknowledge these users by providing the number to the suicide lifeline in response to queries include words directly related to committing suicide. While a step in the right direction, this only accounts for a small percentage of SE interactions initiated by users with MHD. Further, we showed that the resources provided to these users in response to their inquiries are more emotionally charged. Hence, there is a demand to adjust SE retrieval and ranking algorithms to dampen the emotional weight MHD searchers are exposed to. Although beyond the scope of this work, given the correlation between our finding and those found in the social media domain, investigating how users with MHD interact with SE interfaces and how the interfaces must adapt to enable support is crucial.

6 LIMITATIONS AND FUTURE WORK

We discuss limitations and future work emerging from our work.

SE. In this preliminary analysis, we examine resources retrieved using Google (and its safe search functionality). Exploring a single SE enabled us to set foundation in this area. Due to the varied retrieval and ranking strategies adopted by popular SE, we believe it is necessary to extend analysis to alternative SE.

Language. Currently, we only consider resources written in English. Given the world-wide adoption of SE, extending our exploration to other languages is a must.

Analysis. We are aware that machine learning strategies are available for sentiment/emotion analysis. As our exploration follows the framework presented by Kazai et al. [11], we use the same lexicon strategy they adopt. We plan to explore machine learning techniques, in addition to available lexicons explicitly for depression (which have been built from social media resources [22]), in future iterations of our work.

Data. Query logs are hard to obtain, especially for non-traditional users, e.g., searchers affected by MHD. The lack of access to this resource is what prompted us to create synthetic queries that would enable preliminary inspections. To do so, we assumed that users of subreddits we target are individuals who suffer from MHD, an assumption we cannot confirm.

7 CONCLUSION

We have presented the discoveries that have arisen from exploring the sentiment/emotion profiles of SE (Google in our case) for inquiries common among users with MHD. Preliminary findings reveal significant differences across the sentiment/emotion profiles of SERP created by SE for searchers suffering from MHD vs. traditional SE users. While these results are not surprising given the context of our work, there was no prior documentation highlighting the differences in emotional/sentiment profiles of SE results when responding to users with MHD. With SE being a persuasive technology, presenting resources that evoke emotions in individuals experiencing MHD as opposed to remaining objectivity, could have an effect on the mental health and decision making abilities of users with MHD. Outcomes from this work reveal new research questions to be addressed by the information retrieval and human-computer interaction communities, including how query suggestions, ranking, and interface design can influence users with MHD. Further, how existing SE should be adapted to recognize and better serve this user group.

REFERENCES

[1] 99firms.com. 2019. Search Engine Statistics. Available at: https://99firms.com/blog/search-engine-statistics/. (accessed March 20, 2020).

[2] Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. 2010. Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining.. In *Lrec*, Vol. 10. 2200–2204.

[3] Michael L Birnbaum, Sindhu Kiranmai Ernala, Asra F Rizvi, Munmun De Choudhury, and John M Kane. 2017. A collaborative approach to identifying social media markers of schizophrenia by employing machine learning and clinical appraisals. *Journal of medical Internet research* 19, 8 (2017), e289.

[4] Andrew Campbell, Brad Ridout, Melina Linden, Brian Collyer, and John Dalgleish. 2018. A preliminary understanding of search words used by children, teenagers and young adults in seeking information about depression and anxiety online. *Journal of Technology in Human Services* 36, 4 (2018), 208–221.

[5] J. Clement. 2019. U.S. internet usage penetration 2019, by age group. Available at: https://www.statista.com/statistics/266587/percentage-of-internet-users-by-age-groups-in-the-us/. (accessed March 20, 2020).

[6] Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. 2013. Predicting depression via social media. In *Seventh international AAAI conference on weblogs and social media*.

[7] Munmun De Choudhury, Emre Kiciman, Mark Dredze, Glen Coppersmith, and Mrinal Kumar. 2016. Discovering Shifts to Suicidal Ideation from Mental Health Content in Social Media. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 2098–2110. https://doi.org/10.1145/2858036.2858207

[8] Gianluca Demartini and Stefan Siersdorfer. 2010. Dear Search Engine: What's Your Opinion about...? Sentiment Analysis for Semantic Enrichment of Web Search Results. In *Proceedings of the 3rd International Semantic Search Workshop* (Raleigh, North Carolina, USA) *(SEMSEARCH '10)*. Association for Computing Machinery, New York, NY, USA, Article 4, 7 pages. https://doi.org/10.1145/1863879.1863883

[9] BJ Fogg. 2003. *Persuasive technology: using computers to change what we think and do.* Morgan Kaufmann.

[10] Markham Heid. 2018. You Asked: Is It Bad for You to Read the News Constantly? Available at: https://time.com/5125894/is-reading-news-bad-for-you/. (accessed March 18, 2020).

[11] Gabriella Kazai, Paul Thomas, and Nick Craswell. 2019. The Emotion Profile of Web Search. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1097–1100.

[12] Zina Moukheiber. 2019. Mental Health Awareness Is On The Rise, But Access To Professionals Remains Dismal. Available at: forbes.com/sites/zinamoukheiber/2019/05/22/mental-health-awareness-is-on-the-rise-but-access-to-professionals-remains-dismal/#300040c37ba5. (accessed March 19, 2020).

[13] National Institute of Mental Health. 2019. Mental Illness. Available at: https://www.nimh.nih.gov/health/statistics/mental-illness.shtml. (accessed March 19, 2020).

[14] Esteban A Ríssola, David E Losada, and Fabio Crestani. 2019. Discovering Latent Depression Patterns in Online Social Media. In *Proceedings of the 10th Italian Information Retrieval Workshop* (Padova, Italy) *(IIR '19)*. 13–16. http://ceur-ws.org/Vol-2441/paper10.pdf

[15] H Andrew Schwartz, Johannes Eichstaedt, Margaret Kern, Gregory Park, Maarten Sap, David Stillwell, Michal Kosinski, and Lyle Ungar. 2014. Towards assessing changes in degree of depression through facebook. In *Proceedings of the workshop on computational linguistics and clinical psychology: from linguistic signal to clinical reality*. 118–125.

[16] Dushyant Sharma, Rishabh Shukla, Anil Kumar Giri, and Sumit Kumar. 2019. A Brief Review on Search Engine Optimization. In *2019 9th International Conference on Cloud Computing, Data Science and Engineering (Confluence)*. IEEE, 687–692.

[17] Kaisong Song, Wei Gao, Ling Chen, Shi Feng, Daling Wang, and Chengqi Zhang. 2016. Build emotion lexicon from the mood of crowd via topic-assisted joint non-negative matrix factorization. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*. 773–776.

[18] Sho Tsugawa, Yusuke Kikuchi, Fumio Kishino, Kosuke Nakajima, Yuichi Itoh, and Hiroyuki Ohsaki. 2015. Recognizing Depression from Twitter Activity. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 3187–3196. https://doi.org/10.1145/2702123.2702280

[19] Nikhita Vedula and Srinivasan Parthasarathy. 2017. Emotional and Linguistic Cues of Depression from Social Media. In *Proceedings of the 2017 International Conference on Digital Health* (London, United Kingdom) *(DH '17)*. Association for Computing Machinery, New York, NY, USA, 127–136. https://doi.org/10.1145/3079452.3079465

[20] Luyan Xu, Xuan Zhou, and Ujwal Gadiraju. 2019. Revealing the Role of User Moods in Struggling Search Tasks. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval* (Paris, France) *(SIGIR'19)*. Association for Computing Machinery, New York, NY, USA, 1249–1252. https://doi.org/10.1145/3331184.3331353

[21] Yahoo! 2019. Yahoo! Datasets. Available at: https://webscope.sandbox.yahoo.com/catalog.php?datatype=l. (accessed November 6, 2019).

[22] Amir Hossein Yazdavar, Hussein S. Al-Olimat, Monireh Ebrahimi, Goonmeet Bajaj, Tanvi Banerjee, Krishnaprasad Thirunarayan, Jyotishman Pathak, and Amit Sheth. 2017. Semi-Supervised Approach to Monitoring Clinical Depressive Symptoms in Social Media. In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017* (Sydney, Australia) *(ASONAM '17)*. Association for Computing Machinery, New York, NY, USA, 1191–1198. https://doi.org/10.1145/3110025.3123028

[23] Anis Zaman, Rupam Acharyya, Henry Kautz, and Vincent Silenzio. 2019. Detecting Low Self-Esteem in Youths from Web Search Data. In *The World Wide Web Conference* (San Francisco, CA, USA) *(WWW '19)*. Association for Computing Machinery, New York, NY, USA, 2270–2280. https://doi.org/10.1145/3308558.3313557

[24] Changye Zhu, Baobin Li, Ang Li, and Tingshao Zhu. 2016. Predicting depression from internet behaviors by time-frequency features. In *2016 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*. IEEE, 383–390.

[25] Mark Zimmerman, Caroline Balling, Iwona Chelminski, and Kristy Dalrymple. 2019. Symptom presence versus symptom intensity in understanding the severity of depression: Implications for documentation in electronic medical records. *Journal of affective disorders* 256 (2019), 344–347.