

# Why Do They Leave: Modeling Participation in Online Depression Forums

**Farig Sadeque**

School of Information  
University of Arizona  
Tucson, AZ 85721  
farig@email.arizona.edu

**Ted Pedersen**

Dept. of Computer Science  
University of Minnesota, Duluth  
Duluth, MN 55812-3036  
tpederse@d.umn.edu

**Thamar Solorio**

Dept. of Computer Science  
University of Houston  
Houston, TX 77204-3010  
solorio@cs.uh.edu

**Prasha Shrestha**

Dept. of Computer Science  
University of Houston  
Houston, TX 77204-3010  
pshrestha3@uh.edu

**Nicolas Rey-Villamizar**

Dept. of Computer Science  
University of Houston  
Houston, TX 77204-3010  
nrey@uh.edu

**Steven Bethard**

School of Information  
University of Arizona  
Tucson, AZ 85721  
bethard@email.arizona.edu

## Abstract

Depression is a major threat to public health, accounting for almost 12% of all disabilities and claiming the life of 1 out of 5 patients suffering from it. Since depression is often signaled by decreasing social interaction, we explored how analysis of online health forums may help identify such episodes. We collected posts and replies from users of several forums on *healthboards.com* and analyzed changes in their use of language and activity levels over time. We found that users in the Depression forum use fewer social words, and have some revealing phrases associated with their last posts (e.g., *cut myself*). Our models based on these findings achieved 94  $F_1$  for detecting users who will withdraw from a Depression forum by the end of a 1-year observation period.

## 1 Introduction

According to the World Health Organization, 30.8% of all years lived with disability (YLDs) are due to mental and neurological conditions (WHO, 2001). Among these conditions, depression alone accounts for a staggering 11.9% of all the disability. The Global Burden of Diseases, Injuries, and Risk Factors Study estimated that depression is responsible for 4.4% of the Disability-Adjusted Life Years (DALYs) lost, and if the demographic and epidemiological transition trends continue, by the year 2020 depression will be the second leading cause of DALYs lost, behind only ischaemic heart disease (WHO, 2003).

Although depression carries a significant amount of the total burden of all the diseases, this is not its most tragic outcome. Depression claims the lives of

15-20% of all its patients through suicide (Goodwin and Jamison, 1990), one of the most common yet avoidable outcomes of this disorder. Early detection of depression has been a topic of interest among researchers for some time now (Halfin, 2007), but the cost of detection or diagnosis is extremely high, as 30% of world governments who provide primary health care services do not have this type of program (Detels, 2009), and these diagnoses are done based on patients' self-reported experiences and surveys.

Social media offers an additional avenue to search for solutions to this problem. Posts, comments, or replies on different social media sites, e.g., Facebook or Twitter, in conjunction with natural language processing techniques, can capture behavioral attributes that assist in detecting depression among the users (De Choudhury et al., 2013). One property of depression that has not been well explored in social media is its temporal aspect: it may be episodic, recurrent, or chronic, with a recurrence rate of 35% within 2 years. Thus it is critical, when using social media to look at depression, to study how behavioral patterns change over a detailed timeline of the user. For example, decreased social interaction, increased negativity, and decreased energy may all be signals of depression.

In this work, we make the following contributions:

1. We collect a large dataset of user interactions over time from online health forums about depression and other related conditions.
2. We identify phrases (e.g. *cut myself*, *depression medication*) that are highly associated with the last post or reply of a user in a depression forum.
3. We show that users in depression forums have a substantially lower use of social words than

users of related forums.

4. We show that user demographics, activity levels, and timeline information can accurately predict which users will withdraw from a forum.

While these contributions obviously do not represent a solution to depression, we believe they form a significant first step towards understanding how the study of social media timelines can contribute.

There are several other works that have analyzed participation continuation problems in different online social paradigms using different approaches, i.e. friendship relationship among users (Ngonmang et al., 2012), psycholinguistic word usage (Mahmud et al., 2014), linguistic change (Danescu-Niculescu-Mizil et al., 2013), activity timelines (Sinha et al., 2014), and combinations of the above (Sadeque et al., 2015). Also there are numerous works that contribute to the mental health research (De Choudhury et al., 2016; De Choudhury, 2015; Gkotsis et al., 2016; Colombo et al., 2016; Desmet and Hoste, 2013) We believe ours is the first work to integrate language and timeline analysis for studying decreasing social interaction in depression forums.

## 2 Data

Our data is collected from HealthBoards<sup>1</sup>, one of the oldest and largest support group based online social networks with hundreds of support groups dedicated to people suffering from physical or mental ailments. Users in these forums can either initiate a thread, or reply to a thread initiated by others.

We focused on the forums Depression, Relationship Health, and Brain/Nervous System Disorders. While depression remains our main focus, the other two forums represent related conditions and serve as control groups to which we can compare the Depression forum. Relationship Health includes social factors that interact heavily with mental health. Brain/Nervous System Disorder considers a more physical perspective, including the neuropsychiatric disorders Arachnoiditis, Alzheimer’s Disease and Dementia, Amyotrophic Lateral Sclerosis (ALS), Aneurysm, Bell’s Palsy, Brain and Head Injury, Brain and Nervous System Disorders, Brain Tumors, Cerebral Palsy and Dizziness/Vertigo.

<sup>1</sup><http://www.healthboards.com/>

	Depression	Relationship	Brain/Nervous
Posts	19535	17810	13244
Replies	105427	199430	74974
Users	15340	12352	14072
Reply/Post	5.4 (0.1)	11.2 (0.1)	5.6 (0.1)
Post/User	1.3 (0.03)	1.4 (0.03)	0.9 (0.03)
Reply/User	6.9 (0.4)	16.1 (1.3)	5.3 (0.7)
Gender: male	20.77%	22.15%	22.99%
Gender: female	54.07%	57.52%	59.16%
Gender: unspecified	25.16%	20.33%	17.85%

Table 1: Summary of the data collected from HealthBoards. Numbers in parentheses are standard errors.

We crawled all of the posts (thread initiations) and replies to existing threads for these support groups from the earliest available post until the end of April 2016. The posts and replies were downloaded as HTML files, one per thread, where each thread contains an initial post and zero or more replies. The HTML files were parsed and filtered for scripts and navigation elements to collect the actors, contents and general information about the thread. We stored this collected information using the JSON-based Activity Stream 2.0 specification from the World Wide Web Consortium (W3C, 2015). All collected contents were part-of-speech tagged using the Stanford part-of-speech tagger (Manning et al., 2014) and all words were tagged with their respective psycholinguistic categories by matching them against the Linguistic Inquiry and Word Count (LIWC)<sup>2</sup> lexicon.

Table 1 gives descriptive statistics of the dataset. All three forums are roughly similar in number of users. However, users in the Depression forum are less engaged than users in Relationship Health, having a lower average number of replies per post and a lower average number of replies per user. While the Depression forum is similar to the Brain/Nervous System Disorder forum in terms of posts and replies per user, there are more users in the Depression forums that choose not to specify their gender.

### 2.1 Language Analysis

We hypothesized that the final post of a user might include linguistic cues of their decreasing social interaction. For this experiment, we considered users who were inactive for at least the one year preceding the day of data collection. We gathered the contents of their posts, and used pointwise mutual information (PMI) between unigrams (and bigrams) and last posts

<sup>2</sup><http://www.liwc.net/>

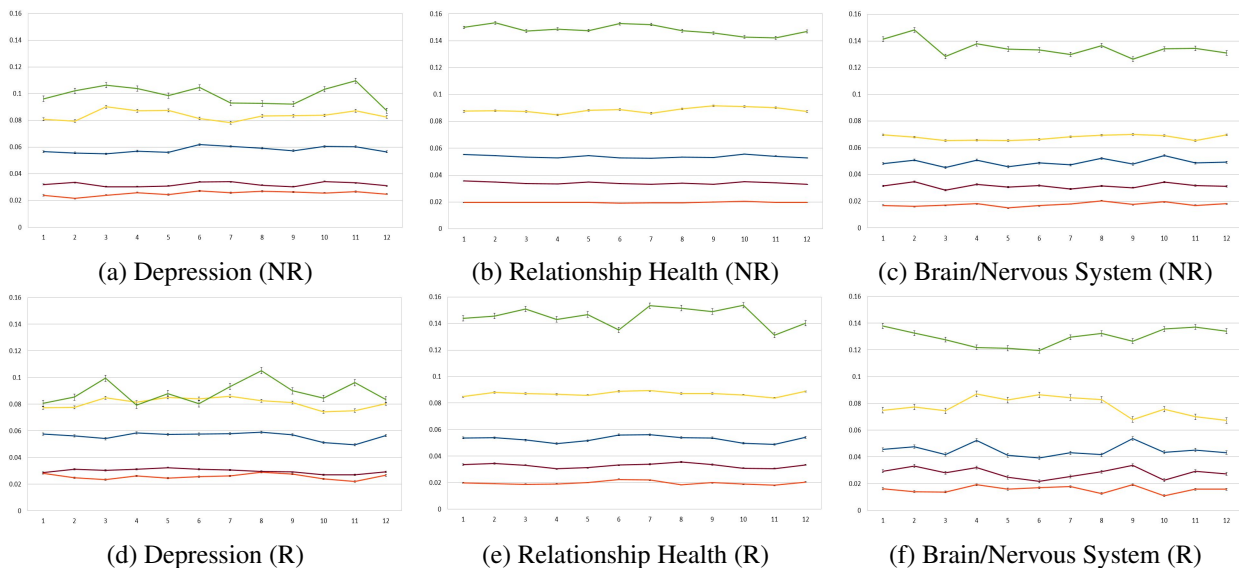


Figure 1: The final 12 months of word use by category: Social (top; green), Cognition (2nd from top; yellow), Affect (middle; blue), Positive Emotion (crimson; 2nd from bottom), and Negative Emotion (orange; bottom).

Depression		Relationship		Brain/Nervous	
Bigrams	PMI	Bigrams	PMI	Bigrams	PMI
I+Feel	0.54	i+no	0.77	got+Bells	0.76
of+Pristiq	0.53	this+disorder	0.74	centre+of	0.73
My+fiance	0.52	a+narcissist	0.70	neural+canal	0.72
My+partner	0.50	wife+said	0.67	prominence+of	0.72
depression+med.	0.48	He+constantly	0.66	ears+from	0.68
in+middle	0.47	dad+does	0.65	bulge+with	0.68
'm+suffering	0.47	confessed+that	0.65	mild+posterior	0.67
slept+with	0.46	Just+recently	0.64	small+intestine	0.67
cut+myself	0.46	my+fiance	0.63	your+biggest	0.67
Any+help	0.46	she+continued	0.63	are+increasing	0.67

Table 2: Top 10 bigrams from each forum based on their PMI with last posts. For space, we abbreviated medication as *med.*

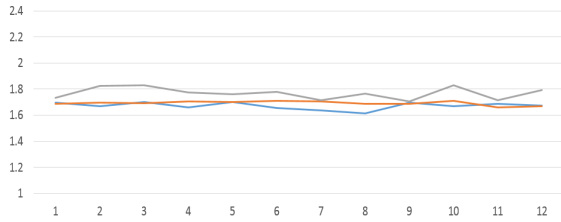
to identify words (and bigrams) that occurred in last posts more often than would be expected by chance. We excluded unigrams occurring less than 50 times and bigrams occurring less than 10 times.

Table 2 lists the top 10 bigrams most associated with last posts in each forum. These n-grams suggest differences in reasons for leaving different types of forums. Depression has some especially revealing phrases: people appear to withdraw from the forum after starting treatment (*of Pristiq, depression medication*), but also after apparent calls for help (*'m suffering, cut myself, Any help*).

We next hypothesized that there may be observable changes over time in the language of users who are

disengaging from the community. We identified the five LIWC psycholinguistic classes that were most associated with last posts across the forums (using PMI as above): Social, Cognition, Affect, Positive Emotion, and Negative Emotion. We selected the most active users that (1) posted in at least two different years, and (2) made at least 100 posts or replies. We then divided these users into two cohorts. The first cohort included the top 100 users who were inactive for at least one year preceding the day of data collection, which we call the non-returning (NR) cohort. The second cohort included the top 100 users with high activity but not marked as inactive yet, which we call the returning (R) cohort.

We then considered the 12 months ending at the user's last post or reply, and graphed the frequency that words from the five psycholinguistic classes were used. Figure 1 shows the use of words from different psycholinguistic classes over time. For most word classes, usage is fairly constant over time and similar across the forums. However, use of social words in the Depression forum is about 40% lower than in Relationship Health or Brain/Nervous System Disorder. This reduced use of social words may indicate less social interaction and less energy, consistent with signs of recurring depressive episodes. Interestingly, both the returning (R) cohort and the non-returning (NR) cohort exhibit this behavior.



(a) Sentiment score (NR)

Figure 2: Sentiment score of activities over the final 12 months for three forums for Depression forums (blue), Relationship Health forums (orange), and Brain/Nervous System Disorder forums (grey).

## 2.2 Sentiment Analysis

We hypothesized that sentiment scores of users’ later activities might provide some insight into their decreasing social interaction. We encountered many posts with negative sentiment after which the user stopped participating in the forum, for example:

... I was really frightened of what was happening to me, my Mum took me straight back to the doctors, to a different one, they were useless, they put me straight on zoloft, I took the zoloft for about 3 days when everything got worse, I couldn’t eat, I kept throwing up, I was having constant panic attacks I just wanted to sleep but lived in fear when I was alone...<sup>3</sup>

To investigate, we took the same users from the language analysis and calculated sentiment for all of their posts and replies using the Stanford CoreNLP (Manning et al., 2014) sentiment analyzer (Socher et al., 2013). The analyzer scores each sentence from 0 to 4, with 0 being extremely negative and 4 being extremely positive. We take as the score for an activity (post or reply) the average of the sentence-level scores for all sentences in that activity.

We graphed the average activity sentiment scores for each forum by averaging the scores over the last 12 months of each user. Figure 2 shows that there was little change in sentiment scores over time, and the lines closely follow the average score for respective forums (Depression: 1.68, Relationship Health: 1.70, Brain/Nervous System Disorder: 1.78). (The figure

<sup>3</sup><http://www.healthboards.com/boards/2346283-post1.html>

	Depression		Relationship		Brain/Nervous	
	Return	Non-return	Return	Non-return	Return	Non-return
AvgInit	114.3	192.5	139.3	420.1	236.3	332.8
AvgMax	218.7	453.3	215.8	492.8	225.2	445.0
AvgMed	1.5	8.8	1.1	15.9	2.7	3.0

Table 3: Average initial (AvgInit), maximum (AvgMax), and median (AvgMed) idle time (in days) for users in the forums.

shows non-returning users; returning users were similar.) This finding refutes our hypothesis regarding sentiment scores of later activities being useful for predicting continued participation.

## 2.3 Idle Time Analysis

We hypothesized that a user’s idle time predict whether they remain socially engaged in the forum. We define idle time as the time between two sequential activities (posting or replying). For each forum, we identified all users who posted in at least two different years, and selected 50 random users who were active within the one year preceding the day of data collection, and 50 random users who were not. We then calculated the initial idle time (from account creation to first activity), maximum idle time, and median idle time. Table 3 shows average initial, maximum, and median idle times across the forums. In general, non-returning users wait longer before their first activity, and have larger maximum and median idle times. Depression forum users have smaller initial idle times than Relationship Health or Brain/Nervous System Disorder users, both for returning and non-returning users.

## 3 Prediction Task

Having observed linguistic and timeline features that suggest when a user is withdrawing from a Depression forum, we began to construct a predictive model for identifying users that are entering such episodes. This task is similar to the *continued participation* prediction task introduced by Sadeque et al. (2015). Formally, we consider the model

$$m_{\Delta t}(u) = \begin{cases} 1 & \text{if } \exists a \in \text{activities}(u) : \\ & \text{start}(u) + \Delta t < \text{time}(a) < \\ & \text{start}(u) + \Delta t + \text{maxtime}(u) \\ 0 & \text{otherwise} \end{cases}$$

Obs	BL	D	A	T	DAT	DATP	DATU	DATB	DATG	DATS
1	78.6	78.7	66.8	81.9	83.3	82.9	78.4	81.2	79.3	83.4
6	87.3	82.9	75.5	90.2	91.1	90.7	90.9	91.1	91.1	91.2
12	92.5	84.1	78.8	93.7	94.0	94.0	94.1	94.0	94.0	94.1

Table 4: F-1 scores predicting which users will stop participating in the Depression forum, for different observation periods (Obs; in months) and different feature sets. BL: Classifier that predicts all inactive.

where  $\Delta t$  is the observation period,  $u$  is a user,  $start(u)$  is the time at which the user  $u$  created an account,  $activities(u)$  is the set of all activities of user  $u$ ,  $time(a)$  is the time of the activity  $a$ , and

$$maxime(u) = \max_{a \in activities(u)} time(a)$$

Intuitively,  $m$  should predict 0 iff  $\Delta t$  time has elapsed since the user created their account and the user will be inactive in the forum for longer than ever before.

We considered the following classes of features:

- D** User profile demographics: gender and whether a location and/or an avatar image was provided.
- A** Activity information: number of thread initiations, number of replies posted, number of replies received from others, number of self-replies.
- T** Timeline information: initial, final, maximum and median idle times.
- U/B/G** Bag of unigrams/bigrams/1-skip-2-grams from the last post of the observation period.
- P** Counts of words for each LIWC psycholinguistic class in the last post of the observation period.
- S** Sentiment score of the last post of the observation period

We trained an L2 regularized logistic regression from LibLinear (Fan et al., 2008) using the data collected from the Depression forum. Throwaway accounts (Leavitt, 2015), defined as accounts with activity levels below the median (2 posts or replies), were excluded from training and testing, though their replies to other users were included for feature extraction. After removing such accounts, 8398 user accounts remained, of which we used 6000 for training our model, and 2398 for testing.

Table 4 shows the performance of this model on different observation periods (1 month, 6 months, 12 months) and different combinations of the feature classes. It also shows the performance of a baseline model (BL) that predicts that all users will be inac-

tive, the most common classification. We measure performance in terms of  $F_1$  (the harmonic mean of precision and recall) on identifying users who withdraw from the forum by the end of the observation period. The most predictive features are the timeline (T) features, resulting in  $F_1$  of 93.7 for a 12 month observation period. Though demographic (D) and activity (A) features underperform the baseline alone, adding them to the timeline features (DAT column) yields a 6% error reduction: 94.0  $F_1$ . The improvement is larger for 1 and 6 month observation periods: 8% and 10% error reductions, respectively.

Adding the language-based features (the DATP, DATU, DATB, DATG, DATS columns) does not increase performance. This is despite our findings in section 2.1 that some phrases were associated with final posts in the forum, but consistent with our findings in Section 2.2 that sentiment analysis was not a strong predictor. This failure of linguistic features may be due to the relatively modest associations; for example, *cut myself* had a PMI of 0.46, and is thus only 38% more likely to show up in a last post than expected by chance. It may also be due to the simplicity of our linguistic features. Consider *Im getting to that rock bottom phase again and im scared*. By PMI, *rock bottom* is not highly associated with last posts, since people often talk about recovering from *rock bottom*. Only present tense *rock bottom* is concerning, but none of our features capture this kind of temporal phenomenon.

## 4 Conclusion

Our analysis of user language and activities in depression-oriented health forums showed that certain phrases and a decline in the use of social words are associated with decreased social interaction in these forums. Our predictive models, based on this analysis, accurately identify users who are withdrawing from the forum, and we found that while demographic, activity, and timeline features were predictive, simple linguistic features did not provide additional benefits. We believe that better understanding of the attributes that contribute to the lack of social engagement in online social media can provide valuable insights for predicting medical issues like depressive episodes, and we hope that our current work helps to form a foundation for such future research.

## References

- Gualtiero B. Colombo, Pete Burnap, Andrei Hodorog, and Jonathan Scourfield. 2016. Analysing the connectivity and communication of suicidal users on twitter. *Computer Communications*, 73, Part B:291 – 300. Online Social Networks.
- Cristian Danescu-Niculescu-Mizil, Robert West, Dan Jurafsky, Jure Leskovec, and Christopher Potts. 2013. No country for old members: User lifecycle and linguistic change in online communities. In *Proceedings of the 22Nd International Conference on World Wide Web, WWW '13*, pages 307–318, Republic and Canton of Geneva, Switzerland. International World Wide Web Conferences Steering Committee.
- Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. 2013. Predicting depression via social media. In *ICWSM*, page 2.
- 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*, pages 2098–2110. ACM.
- Munmun De Choudhury. 2015. Social media for mental illness risk assessment, prevention and support. In *Proceedings of the 1st ACM Workshop on Social Media World Sensors*, pages 1–1. ACM.
- Bart Desmet and Véronique Hoste. 2013. Emotion detection in suicide notes. *Expert Syst. Appl.*, 40(16):6351–6358, November.
- Roger Detels. 2009. *The Scope and Concerns of Public Health*. Oxford University Press Inc., New York.
- Rong-En Fan, Kai-Wei Chang, Cho-Jui Hsieh, Xiang-Rui Wang, and Chih-Jen Lin. 2008. Liblinear: A library for large linear classification. *The Journal of Machine Learning Research*, 9:1871–1874.
- George Gkotsis, Anika Oellrich, Tim JP Hubbard, Richard JB Dobson, Maria Liakata, Sumithra Velupillai, and Rina Dutta. 2016. The language of mental health problems in social media. In *Third Computational Linguistics and Clinical Psychology Workshop (NAACL)*, pages 63–73. Association for Computational Linguistics.
- Frederick K. Goodwin and Kay Redfield Jamison. 1990. *Manic-Depressive Illness: Bipolar Disorder and Recurring Depression*. Oxford University Press Inc., New York.
- Aron Halfin. 2007. Depression: the benefits of early and appropriate treatment. *The American journal of managed care*, 13(4 Suppl):S927, November.
- Alex Leavitt. 2015. "this is a throwaway account": Temporary technical identities and perceptions of anonymity in a massive online community. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, CSCW '15*, pages 317–327, New York, NY, USA. ACM.
- Jalal Mahmud, Jilin Chen, and Jeffrey Nichols. 2014. Why are you more engaged? predicting social engagement from word use. *CoRR*, abs/1402.6690.
- Christopher Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven Bethard, and David McClosky. 2014. The stanford corenlp natural language processing toolkit. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 55–60, Baltimore, Maryland, June. Association for Computational Linguistics.
- Blaise Ngonmang, Emmanuel Viennet, and Maurice Tchunte. 2012. Churn prediction in a real online social network using local community analysis. In *Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2012)*, ASONAM '12, pages 282–288, Washington, DC, USA. IEEE Computer Society.
- Farig Sadeque, Thamar Solorio, Ted Pedersen, Prasha Shrestha, and Steven Bethard. 2015. Predicting continued participation in online health forums. In *SIXTH INTERNATIONAL WORKSHOP ON HEALTH TEXT MINING AND INFORMATION ANALYSIS (LOUHI)*, page 12.
- Tanmay Sinha, Nan Li, Patrick Jermann, and Pierre Dillenbourg. 2014. Capturing attrition intensifying structural traits from didactic interaction sequences of mooc learners. In *Proceedings of the 2014 Empirical Methods in Natural Language Processing Workshop on Modeling Large Scale Social Interaction in Massively Open Online Courses*.
- Richard Socher, Alex Perelygin, Jean Y Wu, Jason Chuang, Christopher D Manning, Andrew Y Ng, and Christopher Potts Potts. 2013. Recursive deep models for semantic compositionality over a sentiment treebank. In *EMNLP*.
- W3C. 2015. Activity streams 2.0: Working Draft 15 December 2015. <http://www.w3.org/TR/2015/WD-activitystreams-core-20151215/>.
- World Health Organization WHO. 2001. The world health report 2001 mental health: New understanding, new hope. [http://www.who.int/whr/2001/en/whr01\\_en.pdf?ua=1](http://www.who.int/whr/2001/en/whr01_en.pdf?ua=1). Last Accessed: 2016-05-29.
- World Health Organization WHO. 2003. Global burden of disease (GBD) 2000: version 3 estimates. <http://www.who.int/entity/healthinfo/gbdwhoregionyld2000v3.xls?ua=1>. Last Accessed: 2016-05-28.