

# Crowdsourcing-based Annotation of Emotions in Filipino and English Tweets

Fermin Roberto G. Lapitan<sup>1</sup> Riza Batista-Navarro<sup>1,2</sup> Eliezer A. Albacea<sup>1</sup>

<sup>1</sup>Institute of Computer Science, University of the Philippines Los Baños, Philippines

<sup>2</sup>School of Computer Science, University of Manchester, United Kingdom

{fglapitan, eaalbacea}@up.edu.ph, riza.batista@manchester.ac.uk

## Abstract

The automatic analysis of emotions conveyed in social media content, e.g., tweets, has many beneficial applications. In the Philippines, one of the most disaster-prone countries in the world, such methods could potentially enable first responders to make timely decisions despite the risk of data deluge. However, recognising emotions expressed in Philippine-generated tweets, which are mostly written in Filipino, English or a mix of both, is a non-trivial task. In order to facilitate the development of natural language processing (NLP) methods that will automate such type of analysis, we have built a corpus of tweets whose predominant emotions have been manually annotated by means of crowdsourcing. Defining measures ensuring that only high-quality annotations were retained, we have produced a gold standard corpus of 1,146 emotion-labelled Filipino and English tweets. We validate the value of this manually produced resource by demonstrating that an automatic emotion-prediction method based on the use of a publicly available word-emotion association lexicon was unable to reproduce the labels assigned via crowdsourcing. While we are planning to make a few extensions to the corpus in the near future, its current version has been made publicly available in order to foster the development of emotion analysis methods based on advanced Filipino and English NLP.

## 1 Introduction

Social media platforms are integral to the lives of Filipinos. In terms of time spent on using social media, Filipinos currently rank first, with an average of 3.7 hours of usage per day (Kemp, 2016). Social media penetration is at 47% of the population which means that almost half of 102 million Filipinos have social media access. Among the most commonly used social media platforms, Twitter ranks sixth with 16% of Filipinos on social media using it. As of May 2016, there are 7.56 million active Twitter users in the Philippines, making it the world's tenth country with the most number of Twitter users.

The Philippines is known not only for being the social media capital (Cameron, 2016), but also for being one of the world's five most natural disaster-prone countries (Esplanada, 2015). Each year, around twenty typhoons enter the Philippine Area of Responsibility (PAR), of which eight to nine make landfall. Aside from typhoons, earthquakes and volcanic eruptions also occur frequently as the country is located within the Pacific Ring of Fire. The local and national government have utilised social media as a means for communicating with citizens during times of disaster. For example, Project NOAH (Lagmay, 2012) of the Philippine Atmospheric Geophysical and Astronomical Services Administration (PAGASA) has created a dedicated Twitter account for announcing weather updates via tweets. Some heads of municipalities and provinces post announcements, e.g., those pertaining to suspension of classes or work, on Facebook and Twitter. Meanwhile, ordinary citizens tweet about traffic situations, current conditions in their local area, as well as share how they feel as these events unfold. Tweets circulating during disasters can thus aid responders obtain meaningful feedback on the current situation in particular areas as well as assess the emotional states of those affected.

---

This work is licensed under a Creative Commons Attribution 4.0 International Licence. Licence details: <http://creativecommons.org/licenses/by/4.0/>

Emotions conveyed in tweets could inform decisions pertinent to disaster risk reduction and management (DRRM). However, such decisions often need to be made urgently. This poses a challenge considering the large volume of tweets that Filipinos generate especially in the event of natural disasters. Automating the identification of emotions in tweets is therefore beneficial, potentially leading to more efficient and timely decision-making. Nevertheless this is considered a difficult natural language processing (NLP) task primarily due to the noisy textual content of tweets. With a 140-character limit per tweet, Twitter users often compact their messages with the use of ungrammatical sentence fragments, intentionally misspelled words and abbreviations. Furthermore, often very little contextual information is expressed in tweets, with each one typically containing only a few words. In the Philippines, another complication arises from the fact that tweets are expressed in either of the country’s two official languages: Filipino (the official name for Tagalog) or English, or even in a mix of both (i.e., “Taglish”). As Filipino is a low-resourced language, not many dictionaries and corpora that could potentially support Filipino NLP are available.

In order to support the development of advanced automatic methods for recognising emotions in tweets generated in the Philippines, we constructed an emotion-annotated corpus of 1,146 disaster-relevant tweets from the country. It consists of Filipino and English tweets which were annotated according to the eight primary emotions identified by Plutchik: anger, anticipation, joy, sadness, trust, surprise, disgust and fear (Plutchik, 2001). In this work, we demonstrate how crowdsourcing facilitated the efficient collection of human-supplied annotations, and describe our measures for ensuring that data quality and reliability were not compromised. A discussion of our results is then presented followed by an analysis that emphasises the value of our newly developed corpus in the context of supporting the development of Filipino and English NLP methods for emotion identification.

## 2 Related work

Sentiment analysis, the automatic classification of pieces of text according to positive, negative or neutral sentiment, has been an active area of NLP research (Pang and Lee, 2008). Some efforts have however further addressed finer-grained classification, in which the specific emotion conveyed by a piece of text is identified. Strapparava and Mihalcea (2007) built a corpus of news titles (i.e., headlines) extracted from news web sites and classified them according to six predefined emotions (Anger, Disgust, Fear, Joy, Sadness, and Surprise) and valence (Positive or Negative). A web interface was developed, allowing annotators to use slider widgets in assigning values between 0 and 100, to indicate how much any of the six emotions of interest is conveyed in each of 1,250 headlines. Six annotators carried out the task, guided by sample annotated headlines including ones expressing multiple emotions. The resulting corpus, split into development and test sets (containing 250 and 1,000 headlines respectively), was employed as gold standard data in the Affective Text shared task of the SemEval 2007 Workshop.

Microblogs generated by social media have also attracted active research on sentiment and emotion analysis. Wen and Wan (2014) sought to classify Chinese microblog texts into one of eight emotion categories (i.e., Anger, Disgust, Fear, Happiness, Like, Sadness, Surprise and None). To support the development of their methods, they constructed a data set consisting of 13,252 sentences from 4,000 microblog texts sourced from Sina Weibo, a popular Chinese microblogging site. Similarly, De Leon and Estuar (2013) aimed to automatically analyse emotions in social media posts, specifically in tweets generated in the Philippines which are mostly written in Filipino or English. To this end, they gathered hundreds of thousands of tweets in both languages, during some of the country’s most prominent disasters. While the resulting data set is undoubtedly a valuable resource, it does not contain any manually produced annotations and thus cannot serve as a gold standard for the development or evaluation of NLP methods.

Indeed, manually labelling emotions in a huge number of tweets is a daunting effort. If done in the traditional manner, i.e., by a small team of human annotators, the task can quickly turn into a burden, potentially leading to the generation of inconsistent annotations. Crowdsourcing, the process of soliciting judgements from contributors (crowds) over the internet, lends itself well to the task of analysing emotions expressed in text. Mohammad and Turney (2013) used Amazon’s crowdsourcing platform,

Mechanical Turk<sup>1</sup>, to build EmoLex, a lexical resource capturing associations between words and any of Plutchik’s eight primary emotions. However, given the risk of attracting underperforming annotators, a few issues with quality control arose, which the proponents attempted to address by keeping annotation instructions simple and easy to understand.

In this study, we cast the analysis of emotions in social media content as a crowdsourcing-based task. We employed the CrowdFlower platform<sup>2</sup>, allowing us to define measures for ensuring that high-quality annotations on tweets are produced. As a result, we have constructed the first gold standard emotion-annotated corpus of Filipino and English tweets, which can facilitate the development of advanced NLP methods for emotion analysis.

### 3 Methods

In this section, we present details on how the annotation of emotions in Filipino and English tweets was carried out. We first describe the data preparation methods employed and then proceed to a discussion of our annotation schema. Finally, we focus on the design and configuration of the task in our chosen crowdsourcing platform.

#### 3.1 Data preparation

Upon request, we obtained a corpus of 660,000 tweets from the Ateneo de Manila University’s Social Computing Science Laboratory who provided us with the data set in compliance with Twitter’s terms and conditions<sup>3</sup>. That is, the original data set was exported to a spreadsheet format which was split into smaller spreadsheets with 50,000 tweets each, provided to us on a one-spreadsheet-per-day basis over a total of 14 days. These tweets were gathered from the 7th to the 9th of August 2012 during which the Philippines’ largest island, Luzon, was hit by heavy southwest monsoon rain (locally known as “habagat”). We first randomly selected 2300 tweets from the whole set. Two automatic pre-processing steps were then carried out on these tweets, namely, duplicate removal and language detection (using Google Spreadsheets’ `detectlanguage` `detectlanguage` function). As our interest is in obtaining annotations on Filipino and English tweets, with the intention to acquire more for the former—given that it is lower-resourced, we finally included 778 Filipino and 570 English tweets in our selection, for a total of 1,348 tweets.

#### 3.2 Definition of emotion classification schema and guidelines

In defining our schema for classifying tweets according to emotion, we adopted the eight primary types identified by Robert Plutchik: Anger, Anticipation, Joy, Sadness, Trust, Surprise, Disgust and Fear (Plutchik, 2001). His wheel of emotions, shown in Figure 1, illustrates how other emotions are just varying intensities of the eight primary ones, or derived through combinations. For example, Ecstasy is a more intense feeling of Joy while Serenity is its less intense variant. Love, on the other hand, is Joy and Trust combined. Apart from the eight emotion types, an additional category Other was included in the classification scheme to account for tweets which are judged as not expressing any emotion.

In order to elucidate the specifications of our task, we formulated a few guidelines. Firstly, only one of the nine categories mentioned above can be assigned to any given tweet; in cases where multiple emotions are conveyed, annotators were asked to select the emotion that is most strongly expressed. Where the identified emotion (e.g., Contempt) falls between two primary emotions (e.g., Anger and Disgust), the annotator should use his/her best judgement to select the emotion which is more strongly conveyed. Emoticons contained in tweets can be considered as valid indicators of predominant emotions. Finally, we define Other as a catch-all category; when a tweet does not express any emotion or if it was written in an unfamiliar language, this category should be selected. Sufficient examples were provided to illustrate each of these guidelines.

---

<sup>1</sup><https://www.mturk.com/mturk/welcome>

<sup>2</sup><https://www.crowdflower.com>

<sup>3</sup><https://dev.twitter.com/overview/terms/agreement-and-policy>

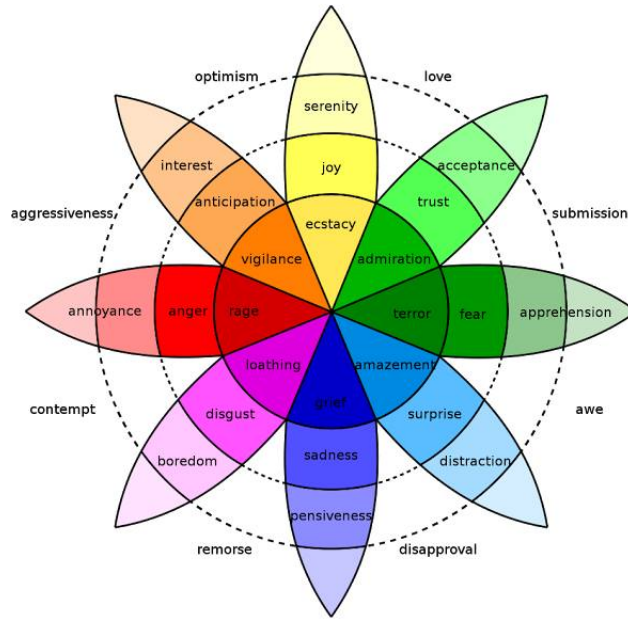


Figure 1: Plutchik’s Wheel of Emotions

### 3.3 Crowdsourcing platform configuration

In implementing our annotation task, two of the most popular crowdsourcing platforms, Amazon Mechanical Turk (AMT) and CrowdFlower, were considered and compared to each other in terms of supporting functionalities. We eventually selected CrowdFlower as our platform due to its in-built measures for ensuring that only high-quality judgements are collected. For instance, it allows for the incorporation of hidden test questions (with corresponding gold standard answers) that could help distinguish hasty annotators from those who are more serious about the task. In this way, only the more conscientious annotators can proceed with the task and contribute their judgements, thus helping us to automatically eliminate ones performing at a low level of accuracy.

After signing up for a trial account in CrowdFlower, we created a task (termed as “job”) and uploaded our data set of 1348 tweets in the form of a spreadsheet. For the purpose of presenting the data to the annotators in a more intuitive manner, a user-interactive web-based form was designed using the CrowdFlower Markup Language (CML). This resulted in the interface depicted in Figure 2, which presented each tweet as well as the nine possible emotion types that an annotator can choose from (Anger, Anticipation, Joy, Sadness, Trust, Surprise, Disgust, Fear and Other) in the form of radio buttons. In order to make the choices more graphical, corresponding illustrative icons were also displayed. Only five tweets (termed in CrowdFlower as “rows”) per page were presented to the annotator at a time, together with the guidelines described above.

Various measures were taken to ensure that only high-quality annotations have been included in our corpus. Firstly, we configured the job to require that each row is assigned independent judgements from at least three different annotators, thus enabling us to assess the level of inter-annotator agreement for each tweet. Furthermore, we took advantage of CrowdFlower’s functionality for including hidden test questions in order to disallow annotators who were performing at a low accuracy, to proceed with the task. To this end, we randomly selected a small set of 50 tweets and manually categorised each of them according to our scheme. Out of these, 28 tweets representative of the emotion types of interest were handpicked as hidden test questions which were interspersed with the rest of the tweets. In defining these test questions, we were allowed by CrowdFlower to specify multiple gold standard answers, e.g., in cases where determining a tweet’s conveyed emotion is not straightforward, i.e., where more than one emotion type could potentially apply. Judgements from annotators eliminated based on our test questions (i.e.,

Choose the most predominant emotion expressed in the Tweet

RT @ogjediaz: Hay, nako, Ulan, Ulan, Ulan...Tandaan mo to: May Araw ka rin.

Anger Fear Anticipation Trust Surprise Joy Sadness Disgust Other

Choose the most predominant emotion expressed in the Tweet

RT @ABSC1213: One of the best things we can do now is stay indoors. Lets not add ourselves to the list of flood victims that need to be...

Anger Fear Anticipation Trust Surprise Joy Sadness Disgust Other

Choose the most predominant emotion expressed in the Tweet

@peydayana bongga ng birthday mo. pa swimming sa buong metro manila. :D happy birthday! :) stay safe. sobra lakas nmn ng ulan.tss..

Anger Fear Anticipation Trust Surprise Joy Sadness Disgust Other

Figure 2: Interface for the annotation of emotions in tweets

those whose accuracy was computed to be less than 70%) were automatically marked by CrowdFlower as untrusted.

In CrowdFlower, task proponents are allowed to specify which performance measure determines task completion. On the one hand, choosing optimal speed (performance = 1) defines the job as complete once the required number of judgements has been obtained, regardless of whether they come from trusted or non-trusted contributors. Choosing optimal quality (performance = 3), on the other hand, makes the job accessible to only the platform’s handful of most trusted contributors, thus potentially taking a longer time to obtain the required number of judgements. For our task, we opted for a compromise between speed and quality (performance = 2), thus allowing us to obtain judgements in a timely manner without sacrificing quality.

Whilst our task is aimed at gathering annotations on Filipino tweets, CrowdFlower does not as yet offer a Filipino language crowd, in the same way that it does for Spanish, French, German, Italian, Hindi, Arabic, Indonesian, Turkish, Italian, Russian, Vietnamese and Chinese (Josephy, 2014). As a workaround, to maximise the exposure of our task to Filipino speakers, we configured our job’s geographical location settings to specify that only contributors from the Philippines are allowed to access the job. However, before launching the job officially, we first gathered feedback on the task from invited contributors (termed in CrowdFlower as “internal workforce”). After making changes according to their suggested revisions on the web-based form and wording of the guidelines, we finally launched our first CrowdFlower job to external contributors with the maximum allowed 999 rows. Upon its completion, we launched a similarly configured second job, this time with the remaining 349 unannotated tweets. While the first job took 26 hours to complete, the second one finished in less than six hours.

Emotion type	Filipino	English	Overall
Anger	67 (10.36%)	11 (2.20%)	78 (6.81%)
Anticipation	37 (5.72%)	14 (2.81%)	51 (4.45%)
Disgust	20 (3.09%)	3 (0.60%)	23 (2.01%)
Fear	20 (3.09%)	5 (1.00%)	25 (2.18%)
Joy	165 (25.50%)	43 (8.62%)	208 (18.15%)
Sadness	72 (11.13%)	22 (4.41%)	94 (8.20%)
Surprise	10 (1.55%)	7 (1.40%)	17 (1.48%)
Trust	33 (5.10%)	20 (4.01%)	53 (4.62%)
other	223 (34.47%)	374 (74.95%)	597 (52.09%)
TOTAL	647 (100.00%)	499 (100.00%)	1146 (100.00%)

Table 1: Distribution of Filipino and English tweets according to emotion

Emotion type	With consensus from 3 annotators	With consensus from 2 annotators	Overall
Anger	19 (28.36%)	48 (71.64%)	67
Anticipation	8 (21.62%)	29 (78.38%)	37
Disgust	5 (25.00%)	15 (75.00%)	20
Fear	5 (25.00%)	15 (75.00%)	20
Joy	94 (56.97%)	71 (43.03%)	165
Sadness	39 (54.17%)	33 (45.83%)	72
Surprise	2 (20.00%)	8 (80.00%)	10
Trust	9 (27.27%)	24 (72.73%)	33
other	95 (42.60%)	128 (57.40%)	223
TOTAL	276 (42.66%)	371 (57.34%)	647

Table 2: Inter-annotator agreement on Filipino tweets

## 4 Results and analysis

A total of 1,348 tweets were manually assigned emotion labels according to the methods described above. However, only judgements on which at least two annotators agreed were retained in order to keep the annotations in our corpus reliable and of high quality. Specifically, an annotated tweet was included in our corpus only if at least two out of three contributors labelled it with the same emotion type. Upon applying this filter, 202 annotations were discarded, leaving a total of 1,146 annotated tweets in our corpus, of which 647 are in Filipino and 499 are in English. Table 1 presents the distribution of these tweets according to the emotion labels assigned to them.

Overall, more than half (52.09%) of the tweets were categorised under the catch-all type Other, many of which were labelled as such for not conveying any emotion, e.g., containing only informative news or announcements. The distribution of such emotion-empty tweets is different though, when the number of annotations is analysed while taking into account the tweets’ language. While most of the English tweets (74.95%) do not express any emotion, in the case of Filipino tweets, majority do convey some emotion (with emotion-empty ones accounting for only 34.47% of the total). This pattern is consistent with previously reported findings that Filipinos tend to tweet in the Filipino language when expressing their feelings, whereas English is mostly used for sharing news and announcements (De Leon and Estuar, 2013). The predominant emotion in both Filipino and English tweets is Joy, having a relative frequency of 25.50% and 18.15%, respectively. For both sets of tweets, the emotion which is least observed is Surprise, which comprises only 1.48% of the tweets.

To aid in our analysis of inter-annotator agreement on the crowdsourced judgements, we compared the number of Filipino tweets that were annotated with perfect agreement (i.e., obtaining consensus from all contributors) against those with majority agreement (i.e., with consensus from only two out of three

Emotion type	With consensus from 3 annotators	With consensus from 2 annotators	Overall
Anger	4 (36.36%)	7 (63.64%)	11
Anticipation	0 (0.00%)	14 (100.00%)	14
Disgust	0 (0.00%)	3 (100.00%)	3
Fear	1 (20.00%)	4 (80.00%)	5
Joy	20 (46.51%)	23 (53.49%)	43
Sadness	12 (54.55%)	10 (45.45%)	22
Surprise	2 (28.57%)	5 (71.43%)	7
Trust	5 (25.00%)	15 (75.00%)	20
Other	243 (64.97%)	131 (35.03%)	374
TOTAL	287 (57.52%)	212 (42.48%)	499

Table 3: Inter-annotator agreement on English tweets

contributors), shown in Table 2. It can be observed that out of the eight primary emotions, Joy and Sadness are the two categories that contributors have assigned to Filipino tweets with perfect agreement more often than not, i.e., at the rates of 56.97% and 54.17%, respectively. In contrast, perfect agreement was much more difficult to obtain in the case of other emotion categories. For instance, 80% and 78.38% of the tweets assigned the labels of Surprise and Anticipation, respectively, were placed under these categories based on majority agreement. Meanwhile, based on inter-annotator agreement on English tweets (Table 3), it can be observed that perfect agreement is difficult to achieve on tweets categorised under Anticipation and Disgust, with all of such annotations resulting from majority agreement only. As in the case with Filipino tweets, many of the English tweets under Joy and Sadness (46.51% and 54.55%, respectively) were obtained based on perfect agreement.

The lack of perfect agreement on many of the annotations indicate that the task of categorising Philippine-generated tweets according to the emotion they convey is non-trivial. This thus confirms our motivation for undertaking this manual annotation task: that the complex language used in tweets necessitates the development of more language resources and advanced NLP methods. To further verify that currently available off-the-shelf tools and resources are not sufficient for accurately categorising tweets according to emotion, we attempted to automatically reproduce the labels manually assigned to our corpus’ Filipino and English tweets by leveraging existing resources. Specifically, we made use of the Hashtag Emotion Lexicon (Mohammad, 2012), a dictionary of 16,862 words frequently appearing in tweets<sup>4</sup>. In this resource, the association of each word with any of Plutchik’s eight primary emotions is specified using a real-valued score, with bigger values indicating stronger associations. We thus predicted the predominant emotion in each of our corpus’ 1,146 tweets by matching words against this lexicon. This allowed us to calculate a cumulative score for each of the eight emotions per tweet; based on this, we took the highest scoring emotion as the tweet’s predominant emotion. The predictions obtained in this manner were then compared against the emotion labels manually assigned to the tweets through crowdsourcing. Shown in Table 4 are the results per emotion category in terms of precision, recall and F-score. Overall, a very low F-score of 13.18% was obtained, although for the Joy and Surprise categories, individual F-scores are higher, i.e., 32.77% and 16.00%, respectively. There are, however, categories (e.g., Anticipation and Disgust) for which no correct predictions were obtained. These findings confirm that further language resources, e.g., gold standard corpora such as the one being proposed in this work, need to be built in order to support the development of accurate methods for identifying emotions in Filipino and English tweets.

<sup>4</sup><http://saifmohammad.com/WebPages/lexicons.html>

Emotion type	True positives	False positives	False negatives	Precision	Recall	F-score
Anger	7	30	71	18.92%	8.97%	12.17%
Anticipation	0	126	51	0.00%	0.00%	0.00%
Disgust	0	42	23	0.00%	0.00%	0.00%
Fear	5	164	20	2.96%	20.00%	5.15%
Joy	106	333	102	24.15%	50.96%	32.77%
Sadness	7	76	87	8.43%	7.45%	7.91%
Surprise	4	29	13	12.12%	23.53%	16.00%
Trust	14	186	39	7.00%	26.42%	11.07%

Table 4: Performance of lexicon-based prediction of emotions against crowdsourced annotations

## 5 Future work and Conclusions

Through crowdsourcing, we were able to build an emotion-annotated corpus of 1,146 disaster-relevant tweets from the Philippines. Our results demonstrate that with appropriate measures for quality control, crowdsourcing can indeed facilitate the efficient collection of emotion-annotated Filipino and English tweets. This was evidenced by the short turnaround time and satisfactory level of inter-annotator agreement on the obtained annotations. We investigated if the human-provided emotion labels of our tweets can be automatically predicted based on a publicly available word-association lexicon. Results from this experiment were not favourable, thus confirming the need for language resources that can facilitate the development of automatic emotion detection methods which obtain better accuracy. One of our immediate future steps involves increasing the number of emotion-annotated tweets in our corpus, especially for categories which currently have low frequencies, e.g., Surprise, Disgust, Fear. Nevertheless, we have made the current version of our newly constructed resource, the EMOTERA (Emotion-annotated Tweets for Disaster Risk Assessment) Corpus, available to the NLP community (<http://tinyurl.com/emoteracorporus>).

## Acknowledgements

We thank all of our reviewers for their invaluable feedback. We are also grateful to the Ateneo de Manila University’s Social Computing Science Laboratory, headed by Dr. Regina Estuar, for sharing their collection of tweets with us. Furthermore, the first author acknowledges the funders of his PhD studies: the Accelerated Science and Technology Human Resource Development Program (ASTHRDP) of the Philippines’ Department of Science and Technology-Science Education Institute (DOST-SEI).

## References

- [Cameron2016] Nathan Cameron. 2016. The Social Media Capital of the World. Online: <http://www.godinternational.org/god-intl-blog/2016/7/the-social-media-capital-of-the-world>. Accessed: 2016-09-20.
- [De Leon and Estuar2013] Marlene M. De Leon and Ma. Regina E. Estuar. 2013. Disaster Emotions: A Bilingual Sentiment and Affect Analysis of Disaster Tweets. In *Proceedings of the 2013 Annual International Conference on Computer Games and Multimedia*, page 70.
- [Esplanada2015] Jerry Esplanada. 2015. PH on UN list of top 5 disaster-prone areas. Online: <http://globalnation.inquirer.net/132796/ph-on-un-list-of-top-5-disaster-prone-areas>. Accessed: 2016-09-20.
- [Josephy2014] Tatiana Josephy. 2014. CrowdFlower Now Offering Twelve Language Crowds. Online: <https://www.crowdflower.com/crowdflower-now-offering-twelve-language-skill-groups/>. Accessed: 2016-09-20.
- [Kemp2016] Simon Kemp. 2016. Digital in 2016. Online: <http://wearesocial.com/uk/special-reports/digital-in-2016>. Accessed: 2016-09-20.



- [Lagmay2012] AMF Lagmay. 2012. Disseminating near-real time hazards information and flood maps in the philippines through web-gis. *Project NOAH Open File Reports*, 1:21–36.
- [Mohammad and Turney2013] Saif M Mohammad and Peter D Turney. 2013. Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, 29(3):436–465.
- [Mohammad2012] Saif Mohammad. 2012. #Emotional Tweets. In *Proceedings of the Sixth International Workshop on Semantic Evaluation (SemEval 2012)*, pages 246–255, Montréal, Canada, 7-8 June. Association for Computational Linguistics.
- [Pang and Lee2008] Bo Pang and Lillian Lee. 2008. Opinion mining and sentiment analysis. *Foundations and trends in information retrieval*, 2(1-2):1–135.
- [Plutchik2001] Robert Plutchik. 2001. The Nature of Emotions. *American Scientist*, 89(4):344–350.
- [Strapparava and Mihalcea2007] Carlo Strapparava and Rada Mihalcea. 2007. SemEval-2007 Task 14: Affective text. In *Proceedings of the 4th International Workshop on Semantic Evaluations*, pages 70–74. Association for Computational Linguistics.
- [Wen and Wan2014] Shiyang Wen and Xiaojun Wan. 2014. Emotion Classification in Microblog Texts Using Class Sequential Rules. In *AAAI*, pages 187–193.