

Automatic detection of stance towards vaccination in online discussion forums

Maria Skeppstedt^{1,2}, Andreas Kerren¹, Manfred Stede²

¹Computer Science Department, Linnaeus University, Växjö, Sweden

{maria.skeppstedt, andreas.kerren}@lnu.se

²Applied Computational Linguistics, University of Potsdam, Potsdam, Germany

stede@uni-potsdam.de

Abstract

A classifier for automatic detection of stance towards vaccination in online forums was trained and evaluated. Debate posts from six discussion threads on the British parental website Mumsnet were manually annotated for stance *against* or *for* vaccination, or as *undecided*. A support vector machine, trained to detect the three classes, achieved a macro F-score of 0.44, while a macro F-score of 0.62 was obtained by the same type of classifier on the binary classification task of distinguishing stance *against* vaccination from stance *for* vaccination. These results show that vaccine stance detection in online forums is a difficult task, at least for the type of model investigated and for the relatively small training corpus that was used. Future work will therefore include an expansion of the training data and an evaluation of other types of classifiers and features.

1 Introduction

There have been outbreaks of vaccine-preventable diseases that were caused by decreased vaccination rates, which in turn were due to negative attitudes towards vaccination. Two examples are an outbreak of polio in 2003-2004, which started in northern Nigeria and spread to 15 other countries (Larson and Ghinai, 2011), and an outbreak of measles in Minnesota in 2017 (Modarressy-Tehrani, 2017).

Information on vaccination can be gathered from many different types of sources. A survey among British parents showed that 34% consulted web-based resources for vaccination information (Campbell et al., 2017). The survey also showed that 31% of the parents that had consulted

chat rooms or discussion forums had seen information that “would make them doubt having their child(ren) immunised or persuade them not to immunise”, compared to 23% for parents consulting Twitter and 8% among all parents included in the survey.

Discussion forums thus form an important outlet for vaccine hesitancy, and this genre might therefore be relevant to automatically monitor for an increase in posts that express a negative stance towards vaccination. Most previous work on training and evaluation of classifiers for automatic detection of vaccination stance has, however, been carried out on tweets. In this study, we therefore take on the task of automatic vaccine stance detection of debate posts in online discussion forums.

2 Background

Mohammad et al. (2017) define *stance detection* as “[...] the task of automatically determining from text whether the author of the text is in favor of, against, or neutral toward a proposition or target”. They distinguish stance detection from the better known task of sentiment analysis by that “in stance detection, systems are to determine favorability toward a given (pre-chosen) target of interest”, whereas sentiment analysis is the task of “determining whether a piece of text is positive, negative, or neutral, or determining from text the speakers opinion and the target of the opinion”. For instance, the utterance “The diseases that vaccination can protect you from are horrible” expresses a stance *for* the pre-chosen target “vaccination”, while expressing a *negative* sentiment towards the sentiment-target “diseases”.

This definition of stance is used in several stance detection studies. For instance, in studies performed on the text genres web debate forums (Somasundaran and Wiebe, 2010; Anand et al.,

2011; Walker et al., 2012; Hasan and Ng, 2013), news paper text (Ferreira and Vlachos, 2016; Fake News Challenge, 2017) and tweets (Augenstein et al., 2016; Mohammad et al., 2017).

Stance detection is generally considered more difficult than sentiment analysis and thereby a task for which currently available methods achieve lower results. This was, for instance, shown by a recent shared task on three-category stance classification of tweets, where an F-score of 0.59 was achieved by a classifier that outperformed submissions from 19 shared task teams (Mohammad et al., 2017). For the task of stance classification of posts of two-sided discussion threads, an F-score of 0.70 is the best result we have been able to find in previous research (Hasan and Ng, 2013)¹.

Previous studies on attitudes towards vaccination do not make use of the term stance, but discuss negative/positive sentiment towards vaccination. There are a number of such sentiment detection studies conducted on tweets, while studies on online forums, to the best of our knowledge, are limited to the task of topic modelling (Tangherlini et al., 2016).

Most vaccination sentiment studies have been conducted on tweets that contain keywords related to HPV (human papillomavirus) vaccination. In one of these, 1,470 HPV-related tweets were manually categorised according to sentiment towards the HPV vaccine (positive, negative, neutral, or no mention). A decision tree was thereafter trained to perform the sentiment classification, and a leave-one-out evaluation resulted in an AUC score of 0.92 (Massey et al., 2016). Another HPV study applied a three-level hierarchical classification scheme and 6,000 tweets were manually categorised as (i) related or unrelated to HPV vaccination, (ii) positive, negative or neutral towards vaccination and (iii) whether they concerned safety, efficacy, emotional resistance, cost and others (Du et al., 2017). A hierarchical classification scheme corresponding to the hierarchy of the categories was applied in the form of support vector machine classifiers, which resulted in a micro-average F-score of 0.74 and a macro-average F-score of 0.59.

¹For these two studies, F-score refers to macro F-score calculated over five and four different stance targets, respectively. This figure was not reported in the paper by Hasan and Ng (2013), but was calculated here from the best results reported for each target. From experiments by Hasan and Ng (2013) that included non-textual features and information from other posts from the same debater, better results than these have been reported.

In a third HPV study, tweets were annotated into the binary category of whether they expressed an anti-vaccine opinion or not (1,050 tweets as training data and 1,100 as test data). A support vector machine trained on bigrams achieved an F-score of 0.82 for detecting negative tweets.

Tweets on the A(H1N1) influenza vaccine have also been automatically classified (Salathé and Khandelwal, 2011). 47,143 tweets that contained keywords related to vaccination were manually classified into four categories; positive/negative/neutral sentiment towards vaccination, or not concerning the A(H1N1) vaccine. The 630 tweets that had been classified by at least 44 annotators, and for which more than 50% of these annotators had selected the same category were used as evaluation data. The rest of the annotated data was used for training a machine learning model in the form of an ensemble classifier built on a Naive Bayes and a Maximum Entropy classifier. This resulted in a classifier accuracy of 0.84.

There are also a number of studies in which purely manual analyses of opinions on vaccination have been carried out. For instance, analyses of blog posts (Brien et al., 2013), online forums (Skea et al., 2008), and of reports on vaccination in many different types of online materials (Larson et al., 2013). Data from the latter analysis has been incorporated into a disease surveillance tool and used for comparing sentiment towards vaccination in different parts of North America (Powell et al., 2016b). However, as pointed out by Powell et al. (2016a), to be able to use this kind of surveillance tool for vaccine-preventable diseases on a larger scale, manual analyses are not enough. Instead, the functionality that we investigate here is required, that is to be able to automatically detect stance towards vaccination. While previous studies on detection of stance/sentiment towards specific types of vaccines in tweets have been carried out, we here aim to investigate the possibility of automatic vaccine stance detection in the important genre of online discussion forums.

3 Method

A corpus was first compiled and pre-processed, and thereafter annotated for stance towards vaccination. The annotated corpus was then used to train models to detect the stance categories.²

²The code used for the experiments, as well as the annotation and post meta-data (Mumsnet ID, debater, debate thread)

3.1 Corpus selection

The experiment was carried out on discussion threads from the British parental website Mumsnet, which is a site that hosts online forums where users discuss and share information on parenting and other topics.³ We based the choice of forum on the reasoning provided by Skea et al. (2008). They chose Mumsnet for manual analysis of vaccine stance, as the site contains discussion threads on many different topics and therefore is likely to attract a more diverse set of debaters than e.g., an anti-vaccination site. In addition, the discussion threads are publicly available without a login and the debaters are asked to anonymise their postings, e.g., by using a chat nickname. This makes it less likely that the posts include content that debaters would like to keep private.

Mumsnet lists a large number of main discussion topics, of which vaccination is one. The debaters can either choose a main topic and start a new discussion thread on a more specific topic, e.g., “Refusing to vaccinate your child”, or submit a post to an existing thread. We extracted posts from the six discussion threads, which, given their name, we assessed as most likely to spur a debate *against/for* vaccination. The topics of all threads with more than 80 posts and for which the latest post was written between the years 2011 and 2017 were considered (68 threads). Only thread names that encouraged discussions on child vaccination in general were included, while debates on more specific aspects on vaccination or vaccination for specific diseases were excluded. Examples of topics excluded for these reasons were “Vaccinations and nursery schools”, “Staggering Vaccinations?” or “HPV gardasil⁴”. Other types of threads excluded were those with a yes/no question as thread name, as answers to these might be more difficult to understand without context, and threads asking for explanations for an opposite view — e.g., “Please explain, succinctly, the anti vac argument.” — as such threads might prompt debaters to list opinions of opponents rather than to express their own arguments.

3.2 Corpus pre-processing and filtering

The text and meta-data of the discussion posts could be extracted from the html-pages based on

is available at: github.com/mariask2/vaccination_stance

³www.mumsnet.com/Talk

⁴The name of a vaccine.

their *div class*. Html tags in the text were removed and the text was segmented into paragraphs using jusText (Pomikálek, 2011).

Texts previously written by other debaters are sometimes copied into new posts in order to indicate that a comment to this previous post is made (the posts are all posted on the same level, and there is no functionality for posting an answer to a specific previous post). Although the debaters do not use a uniform approach to indicate that text has been copied from another debater, we devised a simple method for removing as many instances as possible of copied text. Paragraphs that were exclusively constructed of sentences that had occurred in previous posts were removed, using the standard sentence segmentation included in NLTK (Natural Language Toolkit) for sentence matching (Bird, 2002). In addition, text chunks longer than three words that were marked in bold or by double quotation were removed, in order to exclude citations in general, from other debaters as well as from external sources. Also names of opponents are sometimes mentioned in the posts, as a means to indicate that the content of the post is addressed to a specific opponent debater. These names were also automatically removed.

Similar to previous vaccination studies on tweets, we considered the content of each debate post without the context of surrounding posts. To increase the likelihood that the debater’s stance towards vaccination would be interpretable without context, only posts containing at least one of the following character combinations were included in the experiment: *vacc / vax / jab / immunis / immuniz*. The removal of posts that did not contain these character combinations resulted in that the original set of 2,225 debate posts included in the six extracted threads was reduced to a set of 1,190 posts. These 1,190 posts (written by 136 different authors) were manually annotated and used in the machine learning experiment.

3.3 Annotation

The 1,190 posts to include in the experiment were presented for manual classification in a random order, without revealing who the debater was or which thread the post belonged to. The annotation was performed by one of the authors of the paper.

Following the principle of the guidelines by Mohammad et al. (2017), we classified the posts as taking a stance *against* or *for* vaccination, or

to be *undecided*. The third category was applied to posts in which the debater explicitly declared to be undecided, as well as to posts for which the debater’s stance towards vaccination could not be determined. The post did not have to explicitly support or oppose vaccination to be classified according to the categories *against* or *for*, but it was enough if an opposition or support could be inferred from the post. The stance taken could, for instance, be conveyed without actually mentioning vaccination, as in “You are very lucky to live in the west, and you are free to make that decision because the majority are giving you herd immunity.” It could also be conveyed through an agreement or disagreement with a known proponent/opponent of vaccination, as the statement “Andrew Wakefield is a proven liar and a profiteer — therefore his “research” is irrelevant to any sane, rational discussion [...]” Web links to external resources were, however, not included in the classification decision, even when a stance towards vaccination could be inferred from the name of the URL.

Mohammad et al. (2017) did not specify in detail how to distinguish between stance *against* and *for* the targets included in the study. However, several of the posts in our data did not express a clear positive or clear negative stance towards vaccination in general, and we therefore needed more detailed guidelines for how to draw the line between *against* and *for*. We adopted the basic rule of classifying the post as *against* vaccination when the debater expressed a stance that opposed an official vaccination policy, e.g., as recommended in the health care system. This included, for instance, posts that expressed criticism against some of the recommended vaccines but an acceptance of others, or an acceptance of vaccination in general but not of the officially recommended vaccination scheme. The post “I challenge my DD vaccine schedule all the time. Last time I refused to allow her have MMR with yet another vaccine just because a government quango says so [...]” was thereby classified as *against* vaccination, although the debater is not negative towards all forms of vaccination. Posts that contained a concern over waning immunity from vaccination were classified as *undecided*, except when this concern was used as an argument against vaccination. Posts that expressed a stance against compulsory vaccination, without revealing a stance on vaccination in general, were also classified as *undecided*.

3.4 Machine learning experiments

A standard text classification approach, in the form of a linear support vector machine model, was applied to the task of automatically classifying the debate posts. This follows the approach of Mohammad et al. (2017), as well as of many of the previously performed vaccine sentiment studies. The model was trained on all tokens in the training data, as well as on 2-, 3- and 4-grams that occurred at least twice in the data. The standard NLTK stop word list for English was used for removing non-content words when constructing one set of n-grams. An additional set of n-grams was generated with a reduced version of this stop word list, which mainly consisted of articles, forms of copula, and forms of “it”, “have” and “do”. The reason for using a reduced list was that negations, pronouns etc. that were included in the standard NLTK stop word list can be important cues for classifying argumentative text.

Two types of classifiers were trained: one to perform the task of classifying posts into all three categories annotated, and the other one to perform the task of distinguishing posts annotated as *against* vaccination from those annotated as *for* vaccination. The classifiers were implemented using scikit-learn’s LinearSVC class with the default settings. For training/evaluation, we applied cross-validation on the 1,190 annotated posts. Due to the relatively small data size, we used 30 folds, instead of the more standard approach of 10 folds.

4 Results

Average F-scores for the classifiers and the confusion matrix was calculated using the standard functionality in scikit-learn⁵. Macro average F-scores of 0.44 and 0.62 were achieved for the three-class classifier and for the binary classifier, respectively (Table 1). The confusion matrix and the precision/recall scores for the three-class classifier show that there were frequent misclassifications between all three categories (Table 2). It can also be derived from this table that there was an even distribution between posts annotated as taking a stance *against* vaccination (41%) and those taking a stance *for* vaccination (38%).

⁵sklearn.metrics.f1_score

Classifier	Micro	Macro
<i>against / for / undecided</i>	0.48	0.44
<i>against / for</i>	0.62	0.62

Table 1: Macro and micro F-score for the two experiments, i.e., (i) a classifier that classifies posts as taking a stance *against* and *for* vaccination or being *undecided* and (ii) a binary classifier that classifies posts as *against* or *for* vaccination.

5 Discussion

Similar to what as been shown in previous stance detection studies, the detection of stance towards vaccination was proven to be a difficult task, at least for the type of model investigated and for the relatively small training corpus used. Results cannot be directly compared to previous studies, as there are a number factors that vary between the studies. For instance, the number of training samples used, evaluation measures applied, as well as criteria in some previous studies for excluding samples from the data set that were difficult to classify. There is, however, a large difference between the results achieved here, and some of the previous studies on detection of sentiment towards vaccination, which probably cannot be accounted for by these variations. Instead, it is likely that these differences are due to that the previous studies i) were conducted on tweets, and ii) used a more precise stance target. As tweets are short, they are likely to be more to the point than the more elaborate and longer discussions of the debate forums that we used, and therefore easier for stance detection. The more precise stance target is likely to result in that a more limited set of topics are discussed, which are easier to learn compared to the more wide stance target of vaccination in general that we applied. In future work, the training corpus will be expanded in order to explore if this can improve results. In addition, an evaluation of a wider range of machine learning methods and features will be conducted.

For constructing the corpus used here, a number of decisions had to be made. In the following sections we reflect on these decisions and make a number of suggestions for future studies.

5.1 Annotation decisions

The decision to treat each debate post as one independent unit without taking its context into account is not self-evident, as a post is more meaningful to interpret in the context of its discussion

		Classified as:			total
		<i>against</i>	<i>for</i>	<i>undec.</i>	
Annotated as:	<i>against</i>	275	148	70	493
	<i>for</i>	137	240	73	450
	<i>undec.</i>	101	89	57	247
		<i>against</i>	<i>for</i>	<i>undec.</i>	
	Precision	0.54	0.53	0.29	
	Recall	0.56	0.50	0.23	

Table 2: Confusion matrix and precision/recall-scores for the three-class classifier. The table also shows the total number of posts annotated as *against*, *for* or *undecided*.

thread. Taking the context into account for determining the stance would, however, also entail a more complex classification task. At least in a first step in future work, we will therefore concentrate on the types of debate posts that can be interpreted without context. Research on online forums by Anand et al. (2011) has shown that the incorporation of features from the previous post can improve the performance of a stance classifier for posts that express rebuttals. This work, however, used meta-data of the debaters’ stance for training the classifiers and did not provide the option to classify a post as undecided when its stance could not be determined without its debate context.

Another possible approach to take would be to classify *debaters* according to the stance they take, instead of classifying each individual post into a stance. The set of debate posts written by one debater would then be treated as one unit of text, and the assumption that debaters do not change their stance within a discussion thread would have to be made. Previous research has shown that stance classification of posts in online forums can be improved by also taking classifier output from other posts by the same author into account (Hasan and Ng, 2013).

On the same assumption that debaters do not change their stance, it might also be possible to give some kind of measure of the validity of the stance annotations. If the annotations are to be considered valid, posts from the same debater should ideally always take the same stance, or at least only exceptionally take the opposite stance. Such a measure could be used as a complement to annotator agreement that measures reliability rather than validity (Artstein and Poesio, 2008). Annotator agreement should, however, also be measured.

Previous studies on stance do not reason to a

large extent on where to draw the line between the categories *against* and *for* the target. In the case of previous vaccination stance studies, this might be explained by that these studies have focused on attitudes towards one specific type of vaccine; whereas stance towards vaccination in general is a larger issue with a wider range of possible nuances among debaters' opinions. [Mohammad et al. \(2017\)](#), on the other hand, used broader targets, e.g., "Feminist movement" and "Atheism", but left the task of drawing the line between *against* and *for* to the annotators. Eight annotators classified each tweet, and only the subset of tweets for which at least 60% of the annotators agreed on the classification were retained in the data set.

Other studies have circumvented the problem of giving an exact definition of stance towards a target by using data sets from debate portals, where meta-data on the debaters' stance is provided. Our decision, to classify debate posts as *against* vaccination when they opposed an official vaccination policy, was based on that debaters often implicitly argue against such a policy. In addition, a system for surveilling increases in vaccine hesitancy is likely to take an official policy as its point of departure.

The eight annotators that classified each tweet in the study by [Mohammad et al. \(2017\)](#) were employed through a crowdsourcing platform, which was made possible by that the stance targets were chosen with the criterion that they should be commonly known in the United States. For annotating stance on vaccination, however, annotators with some amount of prior knowledge of vaccine debate topics and vaccine controversies might be preferred. Crowdsourcing might therefore not be a viable option for this annotation task.

40 randomly selected posts that did not fulfil the criterion of containing any of the selected vaccine-related filter terms, and which therefore had been excluded from the study, were also annotated. Although this set is too small to make any definite conclusions, the relatively large proportion of posts in the set that expressed a stance *against* or *for* vaccination (25%) indicates that the filtering criterion used was too crude and led to the exclusion of relevant posts. Future studies should therefore either apply a better filtering criterion, or include all discussion thread posts in an annotation and machine learning study.

5.2 Machine learning decisions

The choice of machine learning model was primarily based on that a linear support vector machine was successful on data from the previously mentioned shared task of stance detection of tweets ([Mohammad et al., 2017](#)). This model outperformed submissions from teams that used methods which might intuitively be better adapted to the task, i.e., an LSTM classifier ([Augenstein et al., 2016](#)). Support vector machines have also been used in many of the previous vaccine sentiment studies. In addition, a linear support vector machine classifier is also a standard method that is often used for different types of text classification tasks ([Manning et al., 2008](#), pp. 335-337). The model is for these reasons suitable to use as a baseline against which to compare future experiments on the vaccine stance classification task.

Despite being the most successful method for stance detection of tweets, it is likely that a support vector machine, trained on word and character n-grams in the entire text is not the optimal method for stance detection of discussion posts. First of all, discussion thread posts are typically longer than a tweet and consist of several sentences, which often each of them form an own argument with a relatively independent content. A classifier that operates on the level of a sentence, or other shorter parts of the text, and combines the stance classification of each of these segments into a post-level classification might be better suited to the task. For instance, [Hasan and Ng \(2013\)](#) improved their stance classification results by expanding their feature set to also include an unsupervised estimation of the stance of each sentence in the debate post.

In addition, it is likely that even if applied on a sentence-level, the use of n-grams would not capture the full complexity of the argumentative text genre. Instead, the structure of the words in the sentences might need to be incorporated in the feature set. A classifier trained on a token-level and where neighbouring tokens in a large context window are incorporated as features could be one such approach. Another possibility would be the previously mentioned approach of stance detection using an LSTM classifier ([Augenstein et al., 2016](#)).

Previous studies have also been able to improve results, at least for some targets, by incorporating other features than n-grams. For instance, features constructed using an arguing lexicon ([Somasun-](#)

garan and Wiebe, 2010), or word embeddings constructed in an unsupervised fashion using a large corpus from the same text genre as the text to classify (Mohammad et al., 2017).

Apart from making a decision on what type of classifier and features to use, it must also be decided on how to gather more training data. Strategies for reducing the manual labelling effort should be investigated, in particular since the annotation task, as discussed above, might not be suitable for crowdsourcing. One possible approach would be to use weakly supervised data. Posts from vaccine-related discussion threads contrasted with posts from other discussion threads might be used as weakly supervised data for classifying posts as either taking a stance on vaccination or being undecided. Weakly supervised data for classifying into stance *against* or *for* vaccination could be gathered by using the assumption that debaters do not change their stance within a thread. A few posts from each of the debaters would then be manually annotated and the rest of the posts from this debater would automatically be assigned the same stance category. Hasan and Ng (2013) improved results by incorporating weakly labelled data that was gathered through harvesting text adjacent to phrases that, with a large confidence, are indicators of stance *against* or *for* the target in question.

It can be noted that the number of contributors to each discussion thread seems to be rather small, as the posts annotated for the experiment had been written by only 136 debaters. Future experiments on this data set and its extensions should therefore evaluate to what extent a classifier trained on the Mumsnet data is able to detect vaccination stance in discussion threads in general. That is, to make sure the models have avoided overfitting on the language typical to a small set of authors and to arguments typical to Mumsnet. This could be carried out by evaluating the classifier on vaccine-related debate posts from other forums. The experimental design would then consist of using the Mumsnet data for the parameter setting and for training the models, and posts from other debate forums for evaluation.

6 Conclusion

A macro F-score of 0.62 was achieved for a binary classifier that distinguished debate posts taking a stance *against* vaccination from those tak-

ing a stance *for* vaccination. When also including the category *undecided*, an F-score of 0.44 was achieved for the three-class classifier. The detection of stance towards vaccination in online forums was proven to be a difficult task, at least for the support vector machine model trained on n-grams that was used as classifier and for the relatively small training corpus used. Future work will therefore include the expansion of the training data set, as well as an evaluation of other types of machine learning models and feature sets.

Acknowledgements

We would like to thank the reviewers for their valuable input. We would also like to thank the Swedish Research Council (Vetenskapsrådet) that funded this study, mainly through the project “Navigating in streams of opinions: Extracting and visualising arguments in opinionated texts” (No. 2016-06681) and partly through the StaViCTA project, framework grant “the Digitized Society – Past, Present, and Future” (No. 2012-5659).

References

- Pranav Anand, Marilyn Walker, Rob Abbott, Jean E Fox Tree, Robeson Bowmani, and Michael Minor. 2011. Cats rule and dogs drool!: Classifying stance in online debate. In *Proceedings of the 2nd workshop on computational approaches to subjectivity and sentiment analysis*, pages 1–9, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Ron Artstein and Massimo Poesio. 2008. Inter-coder agreement for computational linguistics. *Comput. Linguist.*, 34(4):555–596.
- Isabelle Augenstein, Tim Rocktäschel, Andreas Vlachos, and Kalina Bontcheva. 2016. Stance detection with bidirectional conditional encoding. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 876–885, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Steven Bird. 2002. Nltk: The natural language toolkit. In *Proceedings of the ACL Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics*, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Stephanie Brien, Nona Naderi, Arash Shaban-Nejad, Luke Mondor, Doerthe Kroemker, and David L. Buckeridge. 2013. Vaccine attitude surveillance using semantic analysis: constructing a semantically annotated corpus. In *Proceedings of International*

- World Wide Web Conference, IW3C2, New York, NY, USA. ACM.
- Helen Campbell, Angela Edwards, Louise Letley, Helen Bedford, Mary Ramsay, and Joanne Yarwood. 2017. Changing attitudes to childhood immunisation in english parents. *Vaccine*, 35(22):2979–2985.
- Jingcheng Du, Jun Xu, Hsingyi Song, Xiangyu Liu, and Cui Tao. 2017. Optimization on machine learning based approaches for sentiment analysis on hpv vaccines related tweets. *J Biomed Semantics*, 8(1):9.
- Fake News Challenge. 2017. Exploring how artificial intelligence technologies could be leveraged to combat fake news. <http://www.fakenewschallenge.org>, Accessed 2017-08-24.
- William Ferreira and Andreas Vlachos. 2016. Emergent: a novel data-set for stance classification. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL HLT*, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Kazi Saidul Hasan and Vincent Ng. 2013. Stance classification of ideological debates: Data, models, features, and constraints. In *International Joint Conference on Natural Language Processing, IJCNLP*, pages 1348–1356, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Heidi J Larson and Isaac Ghinai. 2011. Lessons from polio eradication. *Nature*, 473(7348):446–7.
- Heidi J Larson, David M D Smith, Pauline Paterson, Melissa Cumming, Elisabeth Eckersberger, Clark C Freifeld, Isaac Ghinai, Caitlin Jarrett, Louisa Paushter, John S Brownstein, and Lawrence C Madoff. 2013. Measuring vaccine confidence: analysis of data obtained by a media surveillance system used to analyse public concerns about vaccines. *The Lancet Infectious Diseases*, 13(7).
- Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze. 2008. *Introduction to information retrieval*. Cambridge University Press, Cambridge.
- Philip M Massey, Amy Leader, Elad Yom-Tov, Alexandra Budenz, Kara Fisher, and Ann C Klassen. 2016. Applying multiple data collection tools to quantify human papillomavirus vaccine communication on twitter. *J Med Internet Res*, 18(12):e318.
- Caroline Modarressy-Tehrani. 2017. Anti-vaxxers and fear caused Minnesota’s massive measles outbreak. *Vice news*.
- Saif Mohammad, Parinaz Sobhani, and Svetlana Kiritchenko. 2017. Stance and sentiment in tweets. *ACM Transactions on Internet Technology (TOIT)*, 17(3):1–23.
- Jan Pomikálek. 2011. *Removing boilerplate and duplicate content from web corpora*. Ph.D. thesis, Masaryk university, Faculty of informatics, Brno, Czech Republic.
- Guido Powell, Kate Zinszer, Jahnvi Dhananjay, Chi Bahk, Lawrence C Madoff, John S Brownstein, Sabine Bergler, and David L Buckeridge. 2016a. Monitoring discussion of vaccine adverse events in the media: Opportunities from the vaccine sentiment. In *AAAI Workshop: WWW and Population Health Intelligence*, Palo Alto, California, USA. Association for the Advancement of Artificial Intelligence.
- Guido Antonio Powell, Kate Zinszer, Aman Verma, Chi Bahk, Lawrence Madoff, John Brownstein, and David Buckeridge. 2016b. Media content about vaccines in the united states and canada, 2012–2014: An analysis using data from the vaccine sentiment. *Vaccine*, 34(50):6229–6235.
- Marcel Salathé and Shashank Khandelwal. 2011. Assessing vaccination sentiments with online social media: implications for infectious disease dynamics and control. *PLoS Comput Biol*, 7(10):e1002199.
- Zoë C Skea, Vikki A Entwistle, Ian Watt, and Elizabeth Russell. 2008. ‘Avoiding harm to others’ considerations in relation to parental measles, mumps and rubella (MMR) vaccination discussions - an analysis of an online chat forum. *Soc Sci Med*, 67(9):1382–90.
- Swapna Somasundaran and Janyce Wiebe. 2010. Recognizing stances in ideological on-line debates. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, CAAGET ’10*, pages 116–124, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Timothy R Tangherlini, Vwani Roychowdhury, Beth Glenn, Catherine M Crespi, Roja Bandari, Akshay Wadia, Misagh Falahi, Ehsan Ebrahimzadeh, and Roshan Bastani. 2016. “Mommy blogs” and the vaccination exemption narrative: Results from a machine-learning approach for story aggregation on parenting social media sites. *JMIR Public Health Surveill*, 2(2):e166.
- Marilyn A. Walker, Pranav Anand, Robert Abbott, and Ricky Grant. 2012. Stance classification using dialogic properties of persuasion. In *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL HLT ’12*, pages 592–596, Stroudsburg, PA, USA. Association for Computational Linguistics.