# The autonomous vehicle social network: Analyzing tweets after a recent Tesla autopilot crash

Jacelyn Jefferson, Anthony D. McDonald
Texas A&M University

Automated vehicle technologies offer a potentially safer alternative than manually driven vehicles, but only if they are accepted and used appropriately. Social media platforms may offer an opportunity to assess peoples' willingness to accept and use automated vehicle technology, but questions remain on the structure and content of the social media conversation. To answer these questions, we performed an analysis of tweets surrounding a recent Tesla Autopilot incident. Tweets were analyzed at three levels: term frequency, account tweet and retweet frequency, and sentiment. The most frequent terms of the conversation shifted from "amazon" and "startup" to "autopilot" and "vehicle" following the crash, however, the specific tweet content referenced an earlier event. A small portion of accounts were responsible for the majority of the tweets in the dataset, and were rarely retweeted. Positive and negative sentiment decreased following the crash, suggesting that a more complex sentiment analysis is needed to gauge changes in public opinion of automated vehicles.

## INTRODUCTION

In 2016, the National Highway Traffic Safety Administration (NHTSA) reported that there were approximately 7.3 million crashes in the United States (National Center for Statistics and Analysis, 2017). Automated vehicle technologies could reduce these crashes but they also introduce new challenges. One such challenge is bridging the gap between the driver's perception of how automated vehicles should behave and their actual behavior. Research suggests that mismatches in these expectations and the actions of automation may lead to crashes (Victor et al., 2018).

Most of the current research on automated vehicle perceptions and driver expectations has been conducted through surveys. Studies have assessed willingness to pay for autonomous vehicles (Bansal, Kockelman, & Singh, 2016), perceptions of automation benefits (Bansal et al., 2016), trust and willingness to accept automation (Zhang et al., 2019), and experiences with automated technologies such as Tesla Autopilot and Summon (Dikmen & Burns, 2016). These studies have identified mixed opinions regarding automated vehicle expectations. Many individuals are optimistic about autonomous vehicles, yet others have significant safety concerns related to failures and errors (Bazilinskyy, Kyriakidis, & de Winter, 2015; Dikmen & Burns, 2016; Hulse, Xie, & Galea, 2018). While surveys are beneficial for assessing perceptions, they are limited by their sampling and response frequency.

Social media platforms such as Twitter are less controlled than surveys, but represent a much broader sample and provide an almost real-time index of perception. Nearly 45% of adults aged 18-24 and 30% of adults aged 25-49 use Twitter in the United States and the platform hosts over 300 million active accounts worldwide (Shearer & Gottfried, 2017). Recent studies have used social media to understand trends in politics, natural disasters and emergency management (Hughes & Palen, 2012; Mellon & Prosser, 2017; Reddick, Chatfield, & Ojo, 2017). One common thread aross these studies is that Twitter use escalates during significant events. In the automated vehicle domain, crashes are currently viewed as a significant event, often involving significant media and public attention. Thus, analyzing Twitter conversations immediately after crashes may provide direct insights into public perceptions of automated vehicles.

Despite the potential of Twitter as a data source, few transportation researchers have investigated its use for analyzing automated vehicles. Prior work has investigated Twitter perceptions of distracted driving (Roberts & Lee, 2014), and automated vehicle perceptions on YouTube (Li et al., 2018). Some initial work is underway analyze tweets about automated vehicles (Kohl, Mostafa, Böhm, & Krcmar, 2017; Penmetsa, Sheinidashtegol, Musaev, Adanu, & Hudnall, 2019), however more holistic analyses of conversations after crashes are needed. Specifically, there is a need to understand who is talking, who is leading the conversation, and what is being said. The goal of this study is to understand these factors through a case study of tweets surrounding a recent Tesla Autopilot-related crash.

## METHODS

The crash explored in this analysis occurred on February 10th, 2019 (Silverstein, 2019). The driver involved in the incident described the vehicle as "confused" when the Autopilot sensor misinterpreted lines on the roadway as another lane, and collided with a curb and several road signs. This incident was chosen because it was the most recent publicly reported crash in automated mode at the time of data collection. Data was extracted from Twitter and analyzed to understand the change in conversations about autonomous vehicles, three days before the event, on the day the event was first reported, and three days after the event. These time periods were selected based on prior observations that crash communication begins and fades within a period of approximately 3 days following an incident (Penmetsa et. al, 2019).

### Data collection

The dataset used in this analysis was collected as part of a larger effort to continuously monitor tweets on automated vehicles. This larger effort consists of a script that collects tweets twice per days using the twitteR library in R (Gentry, 2017). The search was guided by keywords including: "automated

Copyright 2019 by Human Factors and Ergonomics Society. DOI 10.1177/1071181319631510

vehicle," "autonomous vehicle," "automated car," "autonomous car," "driverless car," "self-driving car," "robotic car," "Mercedes ACC," "Tesla Autopilot" and, "Google self-driving car." These search terms were informally assembled based on a review of terms used in news articles, terms used in academic articles (e.g., McDonald et. al, 2019), and terms defined by the Society of Automotive Engineers (SAE, 2018). The exclusion of current technologies (e.g., Volvo Pilot Assist) is a limitation to be addressed in future work. The full dataset included tweet texts, username, retweet count, and favorite count.

### Data preprocessing

The tweet texts were preprocessed to remove punctuation, URLs, and twitter username references, then transformed to lowercase. Following these steps, common words that are likely to be uninformative for analysis such as "and" and "the" (i.e. stop words) were removed from the dataset. The stop words consisted of a list of standard English words and common domain words including "car", "autonomous", "self-driving," "automated," and "vehicle". The final preprocessing step consisted of stemming the words to their root. The preprocessing steps were completed in R using the qdapRegex package (Rinker, 2017).

Following initial preprocessing an initial manual review was conducted to validate the search criteria. This review identified that few of the retrieved tweets referred to the Tesla crash. To facilitate understanding of the whole dataset, and the responses to the Tesla crash, a second dataset was created containing only Tweets that referenced the Tesla crash. The Tesla crash tweets were identified though a keyword search and manual evaluation. The keywords used to identify the Tesla crash included: "New Jersey," "NJ," and "Tesla." These terms were identified through manual analysis of the data, which showed that the tweets discussing the Tesla crash most often mentioned the location, and the vehicle type.

### Frequency analysis

The datasets were analyzed across several dimensions including the frequency of words within the tweets, frequency of tweets by account, and the frequency and ratio of retweets to new tweets. The tweets by account provides an approximate index for communication drivers, and the type of communication surrounding tweets. The retweets—a Twitter feature that allows users to repost other's Tweets on their own account—provide an index for the amount of new information compared to repeated information.

### Sentiment analysis

In addition to the broad frequency analyses, the data were processed with a sentiment analysis. Sentiment analysis is a method of analyzing a collection of text to measure its positivity or negativity. The analysis consists of looking up words or phrases in a sentiment dictionary and calculating a sentiment score. Sentiment dictionaries map words and phrases to emotional scores, with higher valanced emotions leading to larger absolute scores, and neutral words leading to zero scores. In this study, we employed a sentiment analysis approach developed by Mohammad and Turney (2010). This approach has been used by others to analyze short informal texts and track emotions (Kiritchenko, Zhu, & Mohammad, 2014; Mohammad, Zhu, Kiritchenko, & Martin, 2015). The 'set_nrc_sentiments' function in the 'syuzhet' library was used to attain the sentiment scores of tweets. The function calculates the presence of 8 different emotions: "anger", "anticipation", "disgust", "fear", "joy", "sadness", "surprise", "trust" and overall "negative", and "positive" sentiment scores.

## RESULTS

The searches identified 11,164 tweets overall and 103 tweets referencing the Tesla crash. The number of tweets over the time period varied considerably with 3,569 tweets occurring three days before the crash, 3,039 tweets occurring on the day the crash was first reported, and 4,556 tweets on the third day following the crash.

### Twitter conversation analysis

The analyses of frequently used terms are shown in Figure 1. Figure 1 shows the word "Crash" does not appear prior to the crash, but is a common term after the crash. However, the most common terms on three days prior are "Amazon," "startup," and "invest." These terms suggest that a significant portion of the twitter discussion focused on technology development rather than the crash. On the day the crash was first reported, the most frequent terms were "vehicle", "autopilot", and "drive". Three days after the crash was reported, the most frequent terms were "vehicle", "crash", and "autopilot". Indeed, a keyword search suggested that only 10% of the uncovered tweets specifically mentioned the term "crash" or other variants (e.g., "crashed"). The tweets that did use the term "crash" had substantial variance in their subject matter. For example, several tweets discussed the NHTSA analysis of Tesla safety and several tweets discussed ethical considerations of automated vehicle use. Still others discussed an event captured on a dashboard camera, where an autopilot responded to an unexpected cut-in scenario and avoided a crash.

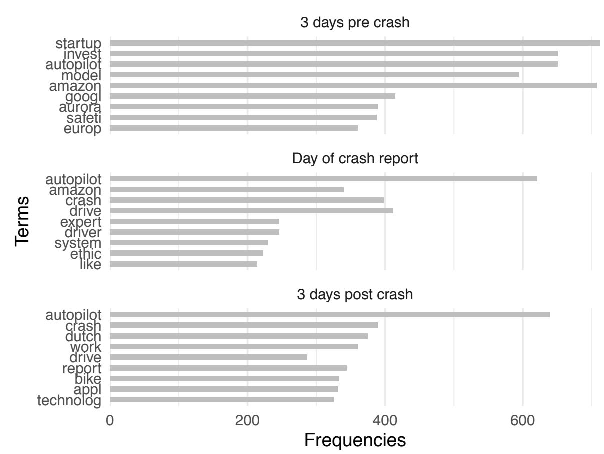


**Figure 1 Bar graph displaying terms frequently tweeted prior to the crash (left) and on the day the incident was first reported (right).**

The tweets surrounding the Tesla crash were considerably more focused. Notable themes in these tweets include confusion and blame. The tweets including the term "confused," most often discussed the fact that the driver involved in the crash suggested that the vehicle became "confused" due to missing lane lines and subsequently crashed. Tweets mentioning "blame" focused on the fact that the driver attributed the blame of the crash to the vehicle.

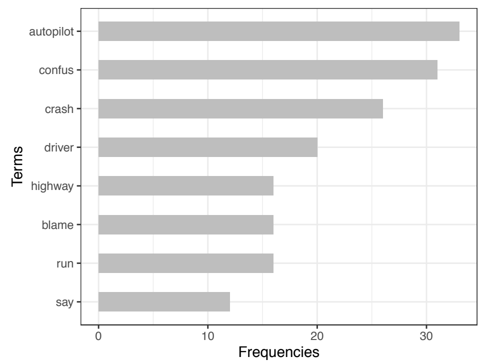


**Figure 2 Bar graph depicting the Twitter conversation surrounding the Tesla crash**

### Retweets and tweet frequency

The tweets by account analysis suggested that a small portion of accounts dominate the conversation on automated vehicles. Figure 3 shows a bar chart of the number of tweets and retweets by account. The chart shows a Pareto-like distribution in which the majority of accounts tweeted three or fewer times. The five accounts that tweeted the most are shown in Table 1 along with the number of tweets they included referencing crashes in general and the Tesla crash. It is notable that at least one of these accounts, TrackBots, is an automated bot account. Interestingly, only one of these accounts discussed the Tesla crash. The tweets referencing the Tesla crash showed a similar exponential trend to the overall tweets, except that the conversation was more balanced across accounts—the maximum number of tweets by a single account was three. The three accounts that tweeted most about the crash also were either representative of stakeholder groups (e.g., auto repair technicians) or individual accounts.

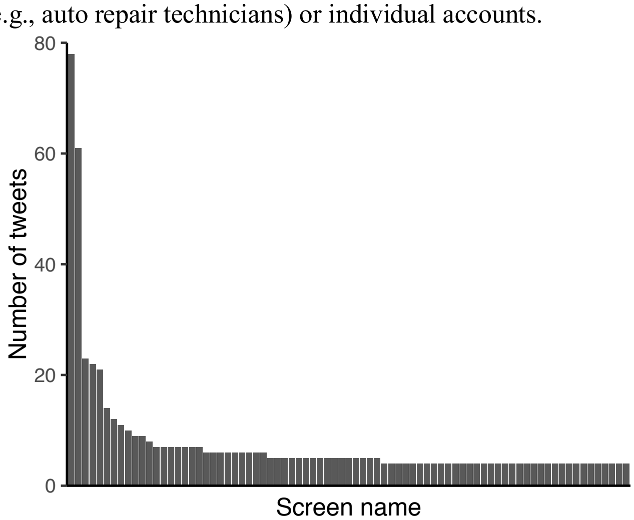


**Figure 3 Bar graph of the frequency of tweets by screen name immediately following the crash.**

**Table 1 Summary of tweet content from the 5 most frequent tweeting accounts. Note that all of the TrackBots tweets are retweets, the other accounts have zero retweets.**

| Screen Name | Total tweets and retweets | Tweets referencing crashes | Tweets referencing the Tesla crash |
|---|---|---|---|
| LidarMonkey | 78 | 8 | 1 |
| TrackBots | 61 | 7 | 0 |
| innovative_usjp | 23 | 1 | 0 |
| wiomax_cn | 22 | 1 | 0 |
| startupcrunch | 21 | 0 | 0 |

Beyond the frequency analysis, there were several notable trends in both datasets. Retweets accounted for 45% of the overall conversation and 41% of the Tesla discussion. Accounts typically generated exclusively original tweets or retweets rather than a mix. Furthermore, there was little overlap between tweet volume and popularity (as measured by retweets). Figure 4 shows the retweets and total tweets for the accounts included in the complete dataset. The figure shows that the most popular tweets (the bottom right portion of the chart) originated from accounts that tweeted fewer than ten times, whereas frequently tweeting accounts were less retweeted. The frequently retweeted tweets typically discussed new technology advancements, or contained humor. For example, several discussed newly launched Tesla hardware.

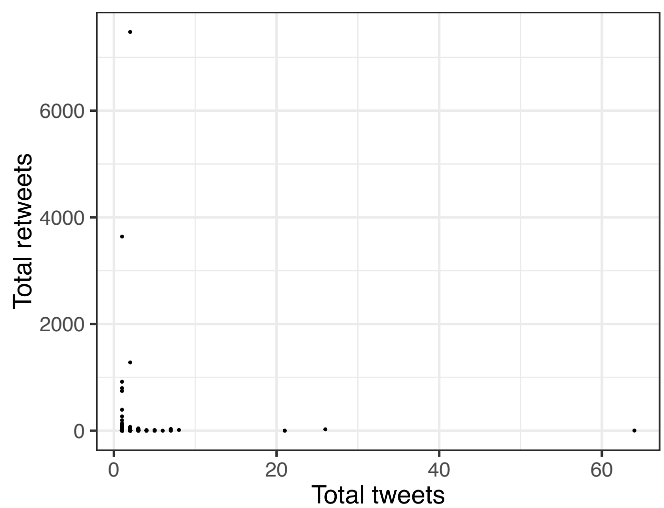


**Figure 4 Scatter plot of total retweets by total tweets for the accounts included in this analysis. Each point represents an account.**

### Sentiment analysis

Figure 5 shows the sentiment scores for the complete dataset across the three days of data collection. The figure illustrates that across time, both positive and negative sentiment decreased following the crash. However, positive sentiment was higher than negative at each time. The drop in positive sentiment was much larger than the decrease in negative sentiment, and the positive sentiment also increased more than the negative sentiment on the 3rd day following the crash.

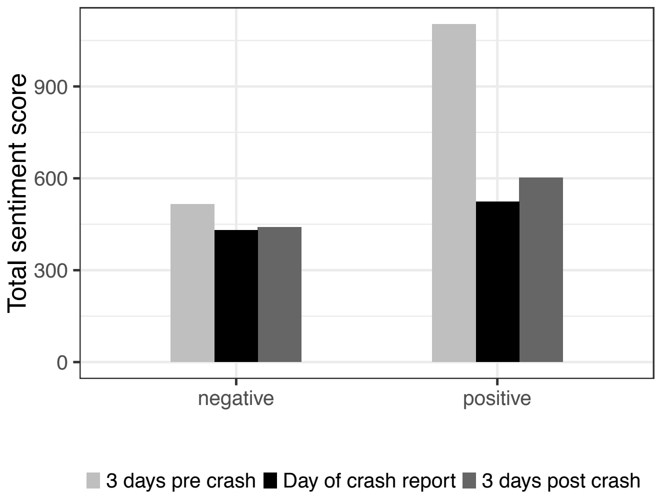


**Figure 5 Positive and negative sentiment scores shown across the three data collection times.**

More clarity on the sentiment surrounding the Tesla crash can be gathered through sentiment analysis of the Tesla crash-related tweets, which is shown in Figure 6. While the tweets have a higher negative sentiment, there is still substantial positive sentiment. This finding should be investigated in future work to examine if it is related to the sentiment dictionary or other factors (e.g., sarcasm being detected as positive sentiment).

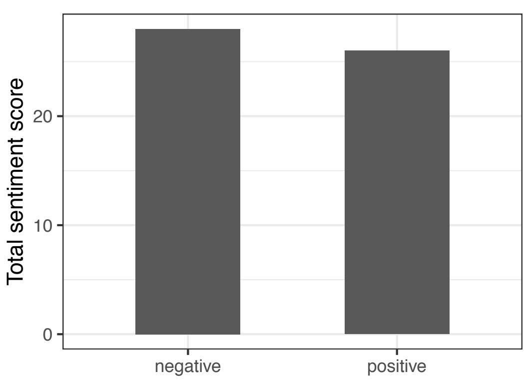


**Figure 6 Positive and negative sentiment totals for the tweets discussing the Tesla crash.**

## DISCUSSION

There are a variety of available methods to assess driver expectations such as surveys and focus groups, however these methods are limited by their ability to sample large groups. Social media platforms, such as Twitter, offer a potential method for overcoming this limitation but they have been under explored in prior research. This study performed an initial assessment of the utility of Twitter for understanding driver expectations and changes in those expectations following an automated vehicle crash.

### Twitter conversation analysis

The analysis of the Twitter conversation was conducted through term frequencies. The results suggested that even following the crash, the majority of the Twitter conversation focused on the development and release of new technologies and company acquisitions (e.g., Amazon's acquisition of Aurora). Following the crash, the term "crash" entered the discussion, however, it represented a variety of topics. Conversations including "crash" focused on NHTSA reports, ethical considerations, overall safety of Tesla vehicles, and the crash of interest. The complexity of these topics is concerning because it may undermine broad analyses such as sentiment or word frequencies. For example, these conflicting subjects may explain the variations in the sentiment analysis reported here.

The tweets discussing the Tesla crash in this analysis were more focused than the complete dataset. Notably many tweets attached to the concept that the driver blamed the Tesla for the crash, and the notion that the Tesla became confused prior to the crash. The repetition of "blame" and "confusion" is interesting as it suggests that Twitter users may be drawn to discussions of crashes involving compelling and concise narratives. The emergence of these terms may highlight a limitation in the keyword search, where new terms emerge and become attached to a crash. This suggests a need for more advanced search methods.

### Retweets and tweet frequencies

The retweet and tweet frequency analysis illustrated that a few Twitter accounts were responsible for the majority of tweets on automated vehicles. Interestingly there was a distinct separation between these accounts and the accounts that were most often retweeted. This suggests that frequently tweeting accounts are not conversation leaders in this context. The conversation surrounding the Tesla crash significantly differed from the overall discussion and involved participation from more individual accounts and special interest groups. This is promising, as it suggests that individuals, and potential or current users, participate in the conversation surrounding automated vehicles. However, as with the larger conversation, significant additional analysis is needed to understand the connection between these individual users and why they felt compelled to enter the discussion.

### Sentiment analysis

The sentiment analysis here identified some expected trends, such as a more negative than positive sentiment in tweets discussing the Tesla crash, however, it also produced several puzzling results. Notably many of the tweets following the Tesla crash were positive, and overall both positive and negative sentiment decreased following the crash. With the current level of analysis, it is unclear if these puzzling findings were due to the sentiment dictionary, the variance in topics discussed in the conversation of crashes, or other variables not reported here. More detailed analyses are needed at each of these levels in future work.

### Limitations and future work

While this study offered a promising initial view of the Twitter conversation regarding automated vehicles and automated vehicle crashes, it is limited in several respects. First, the keyword search here likely missed several tweets involving both automated vehicles and the Tesla crash, as evidenced by the drift of the Tesla conversation to the terms confused and blame. Beyond these factors the sentiment analysis may be limited in several respects. The sentiment dictionary used here is based off the English dictionary and thus it may not align with subjective sentiments associated with transportation domain words. It is also important to note that analysis focused solely on tweet written in English, which limits the extension of these findings to non-English speaking countries. Future work may address these limitations through a more sophisticated search algorithm. Beyond the current limitations, the complexities in the dataset suggest the need for more advanced analysis techniques in future work such as topic modeling (e.g., Lee & Kolodge, 2018).

## CONCLUSION

This analysis provided several insights to the Twitter conversation regarding automated vehicles. Crashes are a topic of conversation, however, most of the focus in the current dataset was on a prior fatal crash. The conversation was dominated by a small set of accounts, however the tweets from these accounts were rarely retweeted. The overall sentiment analysis results were inconclusive. The results provide a promising platform for future work, but illustrate the complexity associated with the use of social media to evaluate public perceptions.

## ACKNOWLEDGEMENTS

Support for this research was provided in part by a grant from the U.S. Department of Transportation, University Transportation Centers Program to the Safety through Disruption University Transportation Center (451453-19C36).

## REFERENCES

Bansal, P., Kockelman, K. M., & Singh, A. (2016). Assessing public opinions of and interest in new vehicle technologies: An Austin perspective. *Transportation Research Part C: Emerging Technologies*, *67*, 1–14.

Bazilinskyy, P., Kyriakidis, M., & de Winter, J. (2015). An International Crowdsourcing Study into People's Statements on Fully Automated Driving. *Procedia Manufacturing*, *3*(Ahfe), 2534–2542.

Dikmen, M., & Burns, C. M. (2016). Autonomous Driving in the Real World: Experiences with Tesla Autopilot and Summon. *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '16), Ann Arbor, MI, USA.*, (May), 225–228.

Gentry, J. (2015). twitteR: R Based Twitter Client. Retrieved from https://cran.r-project.org/package=twitteR

Hughes, A. L., & Palen, L. (2012). The Evolving Role of the Public Information Officer: An Examination of Social Media in Emergency Management. *Journal of Homeland Security and Emergency Management*, *9*(1).

Hulse, L. M., Xie, H., & Galea, E. R. (2018). Perceptions of autonomous vehicles: Relationships with road users, risk, gender and age. *Safety Science*, *102*(April), 1–13.

Kiritchenko, S., Zhu, X., & Mohammad, S. M. (2014). Sentiment Analysis of Short Informal Texts. *Journal of Artificial Intelligence Research*, *50*, 723–762.

Kohl, C., Mostafa, D., Böhm, M., & Krcmar, H. (2017). Disruption of Individual Mobility Ahead? A Longitudinal Study of Risk and Benefit Perceptions of Self-Driving Cars on Twitter. In *Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017)* (pp. 1220–1234).

Lee, J. D., & Kolodge, K. (2018). Understanding Attitudes Towards Self-Driving Vehicles: Quantitative Analysis of Qualitative Data. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *62*(1), 1399–1403.

Li, T., Lin, L., Choi, M., Fu, K., Gong, S., & Wang, J. (2018). YouTube AV 50K: an Annotated Corpus for Comments in Autonomous Vehicles. In *Proceedings of the Thirteenth International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP 2018)*.

McDonald, A. D., Alambeigi, H., Engström, J., Markkula, G., Vogelpohl, T., Dunne, J., & Yuma, N. (2019). Toward Computational Simulations of Behavior During Automated Driving Takeovers: A Review of the Empirical and Modeling Literatures. Human Factors, 61(4), 642–688.

Mellon, J., & Prosser, C. (2017). Twitter and Facebook are not representative of the general population: Political attitudes and demographics of british social media users. *Research and Politics*, *4*(3), 1–9.

Mohammad, S. M., & Turney, P. D. (2010). Emotions Evoked by Common Words and Phrases: Using Mechanical Turk to Create an Emotion Lexicon. *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, Los Angeles, California, June 2010*.

Mohammad, S. M., Zhu, X., Kiritchenko, S., & Martin, J. (2015). Sentiment, emotion, purpose, and style in electoral tweets. *Information Processing and Management*, *51*(4), 480–499.

National Center for Statistics and Analysis. (2017). *2016 fatal motor vehicle crashes: Overview (Traffic Safety Facts Research Note. Report No. DOT HS 812 456)*. Washington, D.C.

Penmetsa, P., Sheinidashtegol, P., Musaev, A., Adanu, E. K., & Hudnall, M. (2019). Did the Recent Autonomous Vehicle Crashes Change Public Perceptions on the Technology? *Journal of Intelligent Transportation Systems (In Review)*.

Rinker, T. W. (2017). qdapRegex: Regular Expression Removal, Extraction, and Replacement Tools. Buffalo, New York. Retrieved from http://github.com/trinker/qdapRegex

Reddick, C. G., Chatfield, A. T., & Ojo, A. (2017). A social media text analytics framework for double-loop learning for citizen-centric public services: A case study of a local government Facebook use. *Government Information Quarterly*, *34*(1), 110–125.

Roberts, S. C., & Lee, J. D. (2014). Deciphering 140 characters: Text mining tweets on #DriverDistraction. In *Proceedings of the Human Factors and Ergonomics Society* (pp. 2195–2199).

SAE International. (2018). *Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles* (Technical report No. J3016). Retrieved from https://saemobilus.sae.org/content/j3016_201401

Shearer, E., & Gottfried, J. (2017). *News Use Across Social Media Platforms 2017 | Pew Research Center*. Retrieved from http://www.journalism.org/2017/09/07/news-use-across-social-media-platforms-2017/

Silverstein, J. (2019, February 19). Driver says Tesla car gets "confused" and crashes on highway. *CBS News*. Retrieved from https://www.cbsnews.com/news/tesla-autopilot-car-gets-confused-and-crashes-on-highway/

Victor, T. W., Tivesten, E., Gustavsson, P. P., Johansson, J., Sangberg, F., & Ljung Aust, M. (2018). Automation Expectation Mismatch: Incorrect Prediction Despite Eyes on Threat and Hands on Wheel. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *60*(8), 1095–1116.

Zhang, T., Tao, D., Qu, X., Zhang, X., Lin, R., & Zhang, W. (2019). The roles of initial trust and perceived risk in public's acceptance of automated vehicles. *Transportation Research Part C: Emerging Technologies*, *98*(November), 207–220.