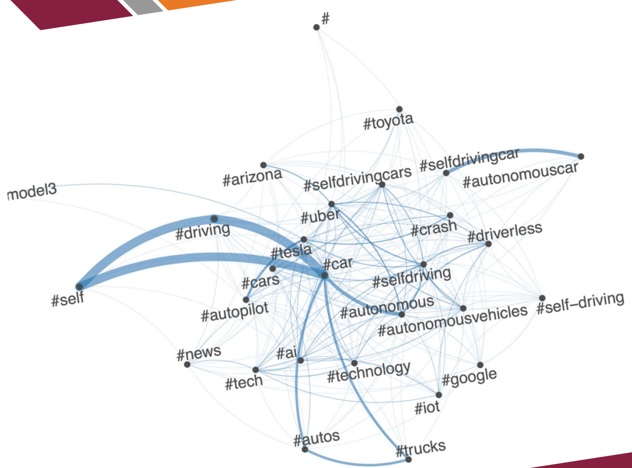
Data mining Twitter to improve automated vehicle safety

February 2021 | Final Report



SAFE-D
SAFETY THROUGH DISRUPTION

VIRGINIA TECH TRANSPORTATION INSTITUTE
VIRGINIA TECH.

Texas A&M Transportation Institute

San Diego State University
Leadership Starts Here

Disclaimer

TECHNICAL REPORT DOCUMENTATION PAGE

<table>
<tr><td>1. Report No.
04-098</td><td>2. Government Accession No.</td><td>3. Recipient's Catalog No.</td></tr>
<tr><td colspan="2" rowspan="2">4. Title and Subtitle
Data mining Twitter to Improve Automated Vehicle Safety</td><td>5. Report Date
February 2021</td></tr>
<tr><td>6. Performing Organization Code:</td></tr>
<tr><td colspan="2">7. Author(s)
Anthony D. McDonald (TTI/TAMU)
Bert Huang (VTTI/VT)
Ran Wei (TTI/TAMU)
Hananeh Alambeigi (VTTI/VT)
Chidubem Arachie (VTTI/VT)
Alec Smith (VTTI/VT)
Jacelyn Jefferson (TTI/TAMU)</td><td>8. Performing Organization Report No.
Report 04-098</td></tr>
<tr><td colspan="2" rowspan="2">9. Performing Organization Name and Address:
Safe-D National UTC
Texas A&M University
Texas A&M Transportation Institute
3135 TAMU
College Station, Texas 77843-3135

Virginia Polytechnic Institute and State University
Virginia Tech Transportation Institute
3500 Transportation Research Plaza
Blacksburg, Virginia 24061</td><td>10. Work Unit No.</td></tr>
<tr><td>11. Contract or Grant No.
69A3551747115/Project 04-098</td></tr>
<tr><td colspan="2" rowspan="2">12. Sponsoring Agency Name and Address
Office of the Secretary of Transportation (OST)
U.S. Department of Transportation (US DOT)</td><td>13. Type of Report and Period
Final Research Report</td></tr>
<tr><td>14. Sponsoring Agency Code</td></tr>
<tr><td colspan="3">15. Supplementary Notes
This project was funded by the Safety through Disruption (Safe-D) National University Transportation Center, a grant from the U.S. Department of Transportation – Office of the Assistant Secretary for Research and Technology, University Transportation Centers Program.</td></tr>
<tr><td colspan="3">16. Abstract
Automated vehicle (AV) technologies may significantly improve driving safety, but only if they are widely adopted and used appropriately. Adoption and appropriate use are influenced by user expectations, which are increasingly being driven by social media. In the context of AVs, prior studies have observed that major news events such as crashes and technology announcements influence user responses to AVs; however, the exact impact and dynamics of this influence are not well understood. The goals of this project were to develop a novel search method to identify AV-relevant user comments on Twitter, mine these tweets to understand the influence of crashes and news events on user sentiment about AVs, and finally translate these findings into a set of guidelines for reporting about AV crashes. In service of these goals, we developed a novel semi-supervised constrained-level learning machine search approach to identify relevant tweets and demonstrated that it outperformed alternative methods. We used the relevant tweets identified to develop a topic model of AV events which illustrated that crashes, fault and safety, and technology companies were the most discussed topics following major events. While the sentiment among these topics was mostly neutral, tweets about crashes and fault and safety were negatively biased. We combined these findings with a series of interviews with Public Information Officers to develop a set of five basic guidelines for AV communication. These guidelines should aid proper public calibration and subsequent acceptance and use of AVs.</td></tr>
</table>

17. Key Words automated vehicles, communication, marketing, social media, Twitter		18. Distribution Statement No restrictions. This document is available to the public through the Safe-D National UTC website, as well as the following repositories: VTechWorks, The National Transportation Library, The Transportation Library, Volpe National Transportation Systems Center, Federal Highway Administration Research Library, and the National Technical Reports Library.	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 38	22. Price $0

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Abstract

Automated vehicle (AV) technologies may significantly improve driving safety, but only if they are widely adopted and used appropriately. Adoption and appropriate use are influenced by user expectations, which are increasingly being driven by social media. In the context of AVs, prior studies have observed that major news events such as crashes and technology announcements influence user responses to AVs; however, the exact impact and dynamics of this influence are not well understood. The goals of this project were to develop a novel search method to identify AV-relevant user comments on Twitter, mine these tweets to understand the influence of crashes and news events on user sentiment about AVs, and finally translate these findings into a set of guidelines for reporting about AV crashes. In service of these goals, we developed a novel semi-supervised constrained-level learning machine search approach to identify relevant tweets and demonstrated that it outperformed alternative methods. We used the relevant tweets identified to develop a topic model of AV events which illustrated that crashes, fault and safety, and technology companies were the most discussed topics following major events. While the sentiment among these topics was mostly neutral, tweets about crashes and fault and safety were negatively biased. We combined these findings with a series of interviews with Public Information Officers to develop a set of five basic guidelines for AV communication. These guidelines should aid proper public calibration and subsequent acceptance and use of AVs

Acknowledgements

This project was funded by the Safety through Disruption (Safe-D) National University Transportation Center, a grant from the U.S. Department of Transportation – Office of the Assistant Secretary for Research and Technology, University Transportation Centers Program.

Table of Contents

List of Figures

List of Tables

Introduction

Automated vehicle (AV) technologies promise to improve driving safety; however, the true impact of AV technologies will be limited by their adoption and proper use. Recent studies suggest that over 50% of drivers still have significant concerns about AVs (Hulse et al., 2018; J. D. Lee & Kolodge, 2019; Schoettle & Sivak, 2015). Although there are several sources of concern, safety is one of the most significant and most frequently cited (Kyriakidis et al., 2015). One method of reducing concern is providing prospective drivers with information regarding the capabilities of AVs. Studies have found that providing drivers with narratives that clarify safety issues (Hohenberger et al., 2016) or provide idealized views of AVs (Nees, 2016) can increase a driver’s intent to adopt AV technology. Conversely, widely reported AV crashes have been shown to negatively bias reported AV acceptance (C. Lee et al., 2019). These findings highlight that the safe adoption and use of AV technologies depends on providing clear information on AV capabilities (Hohenberger et al., 2016).

The proper use of AVs will also have a significant impact on safety. Failures of the human driver to effectively monitor the driving environment have been a contributing factor in several recent crashes (Banks et al., 2017; Seppelt & Victor, 2016). Monitoring failures are often caused by automation misuse (Parasuraman & Riley, 1997), which is a result of overtrust in and overreliance on the automation (J. D. Lee & See, 2004). Decisions to trust and rely on automation are driven by a complex relationship of many factors; however, driver expectations play a significant role (Victor et al., 2018). Guiding users through accessible media should lead to more appropriate expectations, trust, and use of automation, which should provide direct safety benefits.

Social media platforms offer a unique opportunity to both detect expectations and distribute guidance to assuage safety concerns because of the volume of users and their ability to facilitate the spread of information. Twitter is an ideal platform for this analysis as the Twitter demographics are aligned with the demographics of likely AV adopters (Shearer & Gottfried, 2017). Recent surveys suggest that in the United States 45% of adults age 18 to 24 and approximately 30% of adults age 25 to 49 use Twitter. Furthermore, nearly three-quarters of Twitter users depend on Twitter as a source of news (Shearer & Gottfried, 2017). Prior analysis on different media platforms suggests that crashes and safety initiatives are common topics oi discussions of AVs, but little is understood about the network surrounding these communications and how the discussion is impacted by AV crashes (Li et al., 2018). While other studies have performed initial analyses of analogous media platforms (Li et al., 2018), there has not been a comprehensive analysis of Twitter discussions of AVs. The two most significant barriers to such an analysis are the lack of established comprehensive Twitter search methods and the large volume of data returned by such methods. Therefore, our first two goals in this project were to develop a comprehensive Twitter search algorithm and to conduct an analysis to understand the Twitter conversation about AVs.

While understanding the conversation on Twitter about AVs will provide useful insights, additional work is needed to translate the findings into actionable safety benefits. Insight into this translation can be gained through a parallel literature on emergency management. In a recent review on the use of Twitter for emergency management, Luna and Pennock (2018) found that Twitter is an effective method of improving safety following an accident or disaster, specifically if it is used for providing recommendations regarding safety and guidance to the general public. However, the effectiveness of Twitter may be undermined by the spread of misinformation and by a lack of consistent, structured messaging from regulatory agencies (Reuter et al., 2016). One method of providing structured messaging is by defining a set of guidelines that regulatory agencies can use in their communications. While guidelines exist for emergency response from groups such as law enforcement (Community Oriented Policing Services & Police Executive Research Forum, 2013), there are no publicly available guidelines for responses to AV crashes. Therefore, the second goal of this project was to address this gap by developing a set of specific guidelines for AV crash response through feedback from Public Information Officers (PIOs) working in government and transportation.

The goals of this project were accomplished in three phases: (1) develop a search algorithm to identify tweets about AVs and AV crashes; (2) identify themes and trends in the tweets returned by the algorithm; and (3) design a user-centered list of guidelines to guide PIOs when communicating about AVs. The remaining sections of this report describe these steps in detail and illustrate their contributions.

Twitter Search Methodology

The Twitter search methodology developed here builds on initial work on cyberbullying (Raisi & Huang, 2018a, 2018b). A challenge in translating that work to AV crashes is that activities like cyberbullying are person (or account)-centered. In contrast, AV crashes are event centered, meaning that they are more distributed. We addressed this challenge by developing a three-phase process of tweet identification and filtering, which is illustrated in Figure 1. First, we used a keyword search using a curated list of keywords derived from scientific articles, technology announcements, media reports, and SAE automation documentation (SAE International, 2018) indicative of the topic. The keywords were accident, automated, autonomous, autopilot, car, crash, Cruise, driverless, fatality, Google, hit, kill, killed, Lexus, Mercedes, robotic, self-driving, Tesla, Uber, vehicle, Waymo, and Zoox. Second, we followed online social interactions to expand the search. Finally, we filtered the expanded data collection using weakly supervised machine learning. The rationale behind this method is that the initial keyword search is very focused, so it is likely to have high precision (i.e., each tweet containing the keywords is highly likely to be relevant to the AV topic). The social expansion phase significantly increases recall (i.e., the number of relevant tweets that we collect but which reduces the precision of the overall collection). We implemented the keyword search and social expansion phase using a web scraper to facilitate searching of historical tweets outside Twitter's limits on its developer application program interface. The final phase uses low-cost machine learning to filter this collection to recover the high level of precision while retaining the broad recall. The specific procedures for this search are discussed below along with results when applied to the AV dataset.

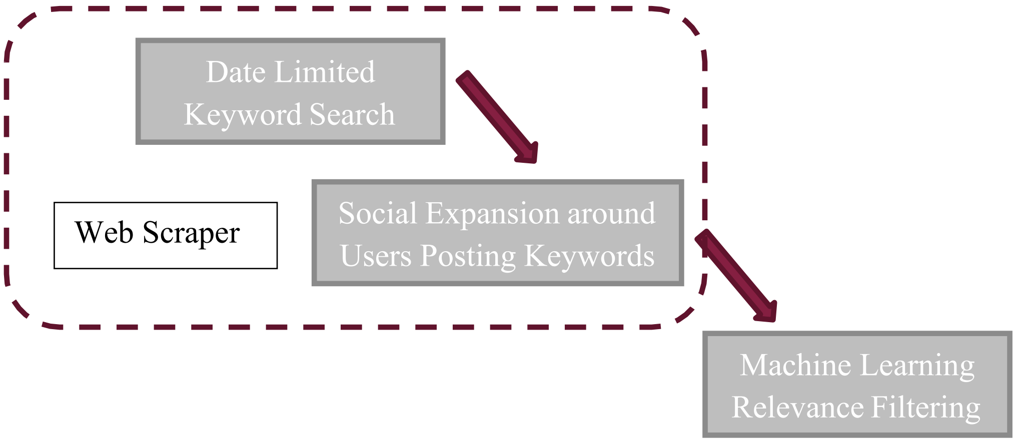


Figure 1. Flowchart. The data curation process.

Weakly Supervised Machine Learning

While machine learning methods are now well known for being powerful tools for automated data analysis, many successful methods require painstaking data annotation of individual examples to train effective models. For the analysis of text data such as detection of topical relevance, one typically must label tens of thousands of examples. Weakly supervised machine learning is a machine learning paradigm in which a learning algorithm trains from approximate indicators of relevance (Arachie & Huang, 2019; Ratner et al., 2017, 2018). The effort required to design these approximate indicators can be orders of magnitude less than that needed for labeling individual examples for fully supervised machine learning. We thus use weakly supervised machine learning for our relevance filtering.

Constrained Label Learning for AV Crashes

We developed a novel weakly supervised algorithm, Constrained Label Learning (CLL), for the project. We assessed the CLL on seven significant AV events: (1) the first AV fatal crash, (2) the introduction of the Tesla autopilot technology, (3) a Tesla crash in New Jersey where the driver suggested that the technology became confused by lane markings, (4) a crash in California involving a Waymo vehicle, (5 and 6) two announcements from the University of Michigan's Mcity connected and automated vehicle environment, and (7) Audi's announcement of the first SAE Level 3 vehicle. These events were selected because they represent a sampling of stakeholders (i.e., private companies, the public, universities), polarity (i.e., positive technology announcements, slightly negative minor crashes, fatal crashes), and times (i.e., 2014–2019).

Separate search processes were conducted for each event. First, we collected keyword-relevant tweets posted in the time window beginning two days before the event and ending two days after. We then expanded the search by collecting tweets sent by users who replied to these original tweets. As noted previously, this expansion increases the breadth of our search by following the online social communities whose members have interests in AV news. However, with this expansion, our data collection process gathered many social media posts that were not relevant to AVs. Thus, we used weakly supervised machine learning to filter the expanded data.

The weakly supervised learning process was guided by four types of approximate indicators: (1) the presence of one of our original keywords, (2) the presence of a URL in the tweet, (3) whether the tweet is detected to be relevant by an unsupervised topic model, and (4) whether the tweet is detected to be in a relevant cluster by an unsupervised clustering model. Specifically, for indicator (3), we trained a topic model and manually examined the topics to identify whether they seemed relevant to AVs. The topic model estimates a probability that each tweet is about each topic, so we used the probability that each tweet is in one of the relevant topics as the relevance indicator. For indicator (4), we converted each tweet into a semantic vector representation using the GloVe method (Pennington et al., 2014), then clustered each event's collection into topics. We then manually examined a sample of tweets from each topic and identified relevant topics. Finally, we used weakly supervised methods to train logistic regression classifiers using GloVe vector representations of the tweets. Our CLL method requires an estimate of the error rate of each approximate indicator. To obtain this estimate, we randomly sampled 50 tweets marked as relevant by each indicator and annotated whether they were in fact relevant. We then used the ratio of irrelevant tweets as the error rate.

To evaluate the CLL method output, we randomly sampled 200 tweets or 10% of the tweets—whichever was smaller—and measured the performance of our weakly supervised learning. For each event, we measured the precision, recall, and F-measure for models trained using majority vote (MV) and CLL. The precision is the proportion of identified tweets that are confirmed to be relevant. The recall is the proportion of confirmed relevant tweets that are identified. The F-measure is the arithmetic average of the precision and recall. The MV baseline outputs as label the majority vote from the weak signals; we used a hard voting scheme by rounding the signals to binary labels. MV assigns equal weights to all the weak signals, unlike CLL, which tries to learn a model for the labels using the weak signals. We also compared CLL and MV to a random selection baseline.

Results

The evaluation results, shown in Table 1, indicate that both MV and CLL significantly improved the quality of the search results over random selection, and that CLL generally outperformed MV. Notably, CLL's precision is comparable to the MV method, but the CLL recall is consistently higher. This improvement in recall is critical for AV crash analyses due to the need to comprehensively capture conversations.

Table 1. Precision, Recall, and F-scores for the CLL and MV Methods Compared to a Random Baseline. Bold indicates the best method on each crash and metric.

Dataset	Precision -- CLL	Precision -- MV	Precision -- Random	Recall– CLL	Recall– MV	F-score– CLL	F-score– MV
First AV fatality	0.907	**0.911**	0.740	**0.993**	0.966	**0.948**	0.938
Minor crash between Waymo vehicle and a Lexus SUV	**0.992**	0.990	0.720	**0.854**	0.701	**0.918**	0.821
Introduction of the Tesla Autopilot	**0.907**	0.878	0.659	**0.967**	0.945	**0.936**	0.910
Introduction of the Mcity driverless shuttle program	**0.938**	0.920	0.610	**0.869**	0.656	**0.902**	0.766
New Jersey "confused" Tesla crash	0.684	**0.733**	0.330	**0.818**	0.500	**0.745**	0.595
Announcement of the Mcity AV testing environment	0.943	**0.969**	0.805	**0.826**	0.782	**0.881**	0.866
Audi announcement of the world's first L3 vehicle	**0.880**	0.846	0.675	**0.978**	**0.978**	**0.926**	0.907

Discussion

We developed a new general-purpose method, called *Constrained Label Learning* (CLL), for weakly supervised machine learning that is robust to coincidences in the weak supervision. While we developed this method to improve our search methodology for this project, the method is widely applicable. Because the approximate indicators can make related mistakes, naively trusting them can lead to the learned model amplifying these mistakes. Simple methods such as training using the majority vote of the weak supervision signals can suffer exactly from this problem. Our new method instead formulates the learning task as a constrained optimization where the weak supervision forms constraints on the space of possible labels of the data. We then sample a random labeling of the full data from this constrained space. We showed with experiments on various benchmark datasets that our new method outperforms other traditional methods and random search.

AV Tweet Analysis

After validating the CLL search method, the next phase of the project consisted of deploying the method to collect data on a larger set of events and to analyze the data to identify broad user sentiments. This phase of the project analyzed 12 significant AV events that occurred between 2014 and 2019. As in the CLL validation, the events—summarized in Table 2—were selected to reflect a distributed sample of positive and negative events originating from multiple stakeholders. The data from each event were analyzed in a two-phase process of probabilistic topic modeling (PTM) and sentiment analysis. PTM was used to filter tweets and then categorize them by emergent themes in the dataset, and sentiment analysis was used to measure the emotional polarity of the conversations over time.

Methods

Dataset

The 12 AV events analyzed here occurred between 2014 and 2019. For each event, a separate CLL search was conducted for the 10 days before and after each crash event. The broader time period compared to the CLL validation was selected to identify changes associated with each event. The events, dates, description, and corresponding tweets are summarized in Table 2.

Table 2. Summary of the Events and the Number of Tweets Included in the Analysis

Event	Date	Description	Number of Tweets
2014 Tesla Advancement	10/09/2014	Tesla software delivered hands-free driving capability on highways and freeways.	32,522
2015 Mcity Advancement	07/20/2015	World-class test facility designed specifically for AV technology testing.	42,769
2016 China Tesla Fatal Crash	01/20/2016	Vehicle crashed into the back of a cleaning vehicle. Vehicle was in autopilot mode and no braking attempt was made.	36,942
2016 California Google Minor Crash	09/23/2017	Google's self-driving vehicle was rear-ended during testing.	51,749
2016 Florida Tesla Fatal Crash	06/30/2016	Vehicle in autopilot mode failed to detect a white semitrailer truck and attempted to drive under it.	92,818
2017 Audi Advancement	07/27/2017	Audi announced 2018 Audi A8 will have Level 3 automation capabilities.	43,104
2017 Las Vegas Shuttle Minor Crash	11/08/2017	A delivery truck backed into the stopping autonomous shuttle.	44,652
2018 Arizona Uber Fatal Crash	03/18/2018	Vehicle in autonomous mode with emergency braking disabled ran into a pedestrian while the driver was watching Hulu TV.	117,841
2018 California Tesla Fatal Crash	03/23/2018	Vehicle in autopilot mode collided into a concrete divider after giving driver warning but received no intervention.	129,714
2018 Mcity Advancement	06/04/2018	Mcity tested how passengers react to driverless shuttles as a way to gauge consumer acceptance.	85,513
2019 New Jersey Tesla Minor Crash	02/11/2019	Vehicle ran into the curb and traffic sign supports after mistaking diagonal white lines for new line.	67,690
2019 Florida Tesla Fatal Crash	03/01/2019	Vehicle ran into the side of a tractor-trailer in autopilot mode without driver's hands on the wheel.	93,358

Data Preprocessing

The datasets were preprocessed with six steps: (1) remove hyperlinks, tags and hashtags, pictures and emojis; (2) extract word tokens and retain nouns, verbs, and adjectives; (3) stem words into morphological roots; (4) remove stop words (i.e., words of common high frequency but low contextual meaning); (5) remove words less than three letters and tweets that contained less than three words; and (6) identify and remove duplicate tweets. All preprocessing steps were conducted in Python. The spaCy library (Honnibal, M., & Montani, 2017) was used for word token extraction and part of speech extraction. The natural language toolkit's (Loper & Bird, 2002) wordnet was used to stem words. All other analyses were conducted with standard Python packages.

Probabilistic Topic Model Fitting and Hyperparameter Selection

The probabilistic topic models used here were implemented with the Latent Dirichlet Allocation (LDA) method. This method was selected due to its success in other domains and its ease of implementation. LDA assumes that a distribution of topics exists among a collection of documents, also called a corpus, which is generated by a distribution of words from each topic (Blei et al., 2002). Fitting an LDA model consists of calibrating three hyperparameters: topic sampling rate (α), word sampling rate (β), and number of topics (n). The sampling rates can be optimized along with the training algorithm; however, the number of topics must be determined separately.

We optimized hyperparameters across two metrics: topic coherence and topic diversity. Topic coherence is a measure of the interpretability of a topic. It is the average pointwise mutual information between topic words in the corpus (Röder et al., 2015). Topic diversity is the percentage of different words in the top words of all topics. The most recent recommendations suggest that the best practice for choosing the optimal number of topics includes comparing multiple metrics and using human judgement (Hagen, 2018). Thus, the number of topics in the current models were identified through an iterative process of calculating the product of normalized topic coherence and diversity, then using human judgement to confirm. For each iteration, the hyperparameters were updated at a fixed interval until model convergence. The process was conducted using the Mallet LDA implementation in Python, which uses a Gibbs sampling method for parameter estimation. This process has been shown to provide more accurate density estimation compared to other commonly used methods (McCallum, 2002).

Two-phase Topic Modeling Approach

While the CLL method substantially reduces irrelevant tweets, initial analyses suggested a need for an additional tweet filter. To address this, we conducted a two-stage topic modeling process consisting of topic filtering and topic identification. In both phases, separate topic models were fit to each individual event. In the topic filtering phase, a topic model was fit to data for the event and then the words most strongly associated with a topic were analyzed for relevance. Relevant and irrelevant topics were identified based on the presence of select words summarized in Appendix A. These words were either strongly associated with AVs (e.g., autopilot) or strongly disassociated with AVs (e.g., Trump). Following this identification, tweets that were strongly associated (more than 0.4 likelihood) with AV relevant topics were retained in the dataset. The remaining tweets in the filtered dataset are shown in Table 3. Following the topic filtering step, the tweets were combined into a single dataset. This combined dataset was used to train a final topic model. The remaining analyses were performed on the results of this topic model. Note that although the final topic model was trained using the filtered datasets, only tweets after each event were used in the remaining analysis.

Table 3. Summary of Tweets Remaining in Each Dataset Following the Pre-processing and Initial Topic Model Filtering

Event	Tweets Before Event	Tweets After Event	Total Number of Tweets
2014 Tesla Advancement	658	2,489	3,147
2015 Mcity Advancement	1,867	4,000	5,867
2016 China Tesla Fatal Crash	1,364	3,429	4,793
2016 California Google Minor Crash	1,688	6,337	8,025
2016 Florida Tesla Fatal Crash	2,214	12,804	15,018
2017 Audi Advancement	2,378	4,437	6,815
2017 Las Vegas Shuttle Minor Crash	761	3,650	4,411
2018 Arizona Uber Fatal Crash	454	10,363	10,817
2018 California Tesla Fatal Crash	1,393	10,564	11,957
2018 Mcity Advancement	3,948	6,872	10,820
2019 New Jersey Tesla Minor Crash	2,009	7,493	9,502
2019 Florida Tesla Fatal Crash	1,688	10,018	11,706

Sentiment Analysis

After the tweets were organized into topics, a sentiment analysis was used to understand the overall attitudes associated with the conversations in each topic. Sentiment analysis is the process of identifying the emotional tone associated with a text through a valence dictionary. Each valence dictionary contains a set of words and associated sentiment values. The sentiment score for a tweet is calculated by first looking up the words in the tweet in the dictionary then adding their individual sentiment scores together. We used the "sentimentr" package in R (Rinker, 2019) to calculate sentiment for this analysis. The package has been employed in a number of contexts, such as analyzing tweets in healthcare (Deng et al., 2020) or the energy (Ikoro et al., 2018) domain. Sentimentr augments the typical additive calculation of sentiment by incorporating weights to valence shifters to calculate the sentiment scores at the sentence level (Naldi, 2019). The valence shifters comprise negators (e.g., not), which change the size of the polarized word; amplifiers (e.g., very), which intensify the impact of the polarized word; de-amplifiers (e.g., barely), which decrease the impact of the polarized word; and adversative conjunctions (e.g., but), which overrule the impact of the polarized word. Sentiment scores were calculated for the tweet text across all topics and crashes.

Results

The topic modeling and sentiment analysis produced a set of topics present in the conversations after AV events indexed by time and sentiment. Below we present these results in distinct sections.

Topic Modeling

The final topic modeling analysis optimization identified six topics. The 20 most strongly associated words with each topic are illustrated in Figure 2 and repeated in Table 4. In addition to the strength of word association, Table 5 shows three word rankings: FREX, Lift, and Score. FREX gives preference to words with both high frequency and exclusivity in a topic by taking the harmonic average of word-conditional-topic probabilities and the sum of conditional probabilities in other topics. Lift attempts to penalize words that are frequent in the corpus by dividing word-topic probabilities by the empirical frequency of words. Score is similar to FREX but instead divides the logarithm of word-topic probability by the sum of the logarithm of word-topic probabilities in other topics. Based on these results, the topics can be categorized under six themes: (1) *Crashes*, (2) *Fault & Safety,* (3) *Market & Sales*, (4) *Tech Companies*, (5) *Electric Vehicles*, and (6) *Public Transit*. These topics can be further understood through example tweets presented in Appendix B.

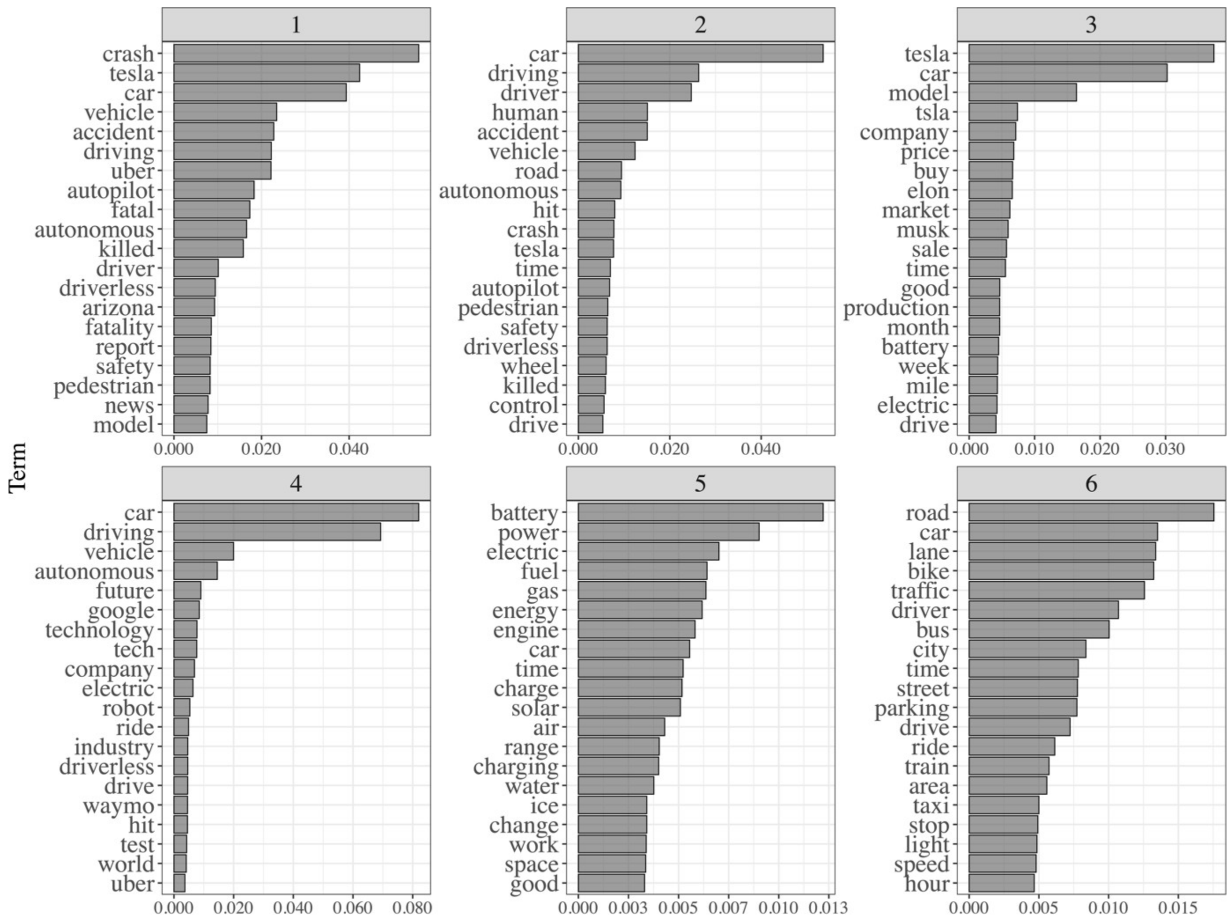


Figure 2. Graphs. Top 20 terms that are most strongly associated with each topic. These topics can be interpreted as (1) *Crashes*, (2) *Fault & Safety*, (3) *Market & Sales*, (4) *Tech Companies*, (5) *Electric Vehicles*, and (6) *Public Transit*, respectively.

Table 4. Topics and Strongly Associated Words by Probability, FREX, Lift, and Score

Topic Interpretation	Probability	FREX	Lift	Score
1 – Crashes	crash, tesla, car, vehicle, accident, driving, uber, autopilot, fatal,	crash, tesla, car, vehicle, accident, uber, driving, autopilot, fatal,	potter, jensen, predicts, ducey, administration, avoids, davie, palo, attenuator, joshua,	medium, mode, man, breaking, revealed, role, vehicle, comment,

	autonomous, killed, driver, driverless, arizona, fatality, report, safety, pedestrian, news, model	autonomous, killed, driver, driverless, arizona, fatality, report, safety, pedestrian, news, model	investigating, fiery, awan, nvidias, huang, condolence, confirms, compilation, deceased, prompt	operating, story, public, program, state, question, suit, update, click, photo, impact, car
2 – Fault & Safety	car, driving, driver, human, accident, vehicle, road, autonomous, hit, crash, tesla, time, autopilot, pedestrian, safety, driverless, wheel, killed, control, drive	car, driver, driving, human, accident, vehicle, road, autonomous, hit, crash, tesla, time, autopilot, pedestrian, safety, driverless, wheel, killed, control, drive	anticipate, attention, trusted, booze, protocol, guinea, carnage, handover, risking, punished, infrared, background, elephant, deserves, encounter, overcome, blood, distracted, attentive, jaywalking	step, problem, aware, owner, happen, matter, missed, fact, feel, car, aid, question, rule, perfect, understand, ability, mode, hold, level, man
3 – Market & Sales	tesla, car, model, tesla, company, price, buy, elon, market, musk, sale, time, good, production, month, battery, week, mile, electric, drive	tesla, car, model, tesla, company, price, buy, elon, market, musk, sale, time, production, month, good, battery, week, mile, electric, demand	engineered, resale, lemon, quarter, edition, bankruptcy, gigafactory, etron, import, powerwall, induced, hatchback, maserati, secured, auction, purchasing, sarcasm, purchase, fraudulent, tourist	owner, month, money, germany, option, list, worth, problem, long, maker, hope, step, update, incredible, big, amazing, small, roll, stick, industry
4 – Tech Companies	car, driving, vehicle, autonomous, future, google, technology, tech, company, electric, robot, ride, driverless, industry, drive, waymo, hit, test, world, uber	car, driving, vehicle, autonomous, future, google, technology, tech, company, electric, robot, ride, industry, driverless, waymo, drive, hit, test, world, uber	alphabet, googl, advancement, inspired, goog, udacity, intro, mapping, sunset, invests, aurora, microsoft, blackberry, mit, pokemon, smartphones, robert, reflection, expo, ibm	hack, industry, maker, closer, team, car, program, roll, engineer, handful, motor, government, vehicle, world, rule, public, begin, amazing, stuff, set
5 – Electric Vehicles	battery, power, electric, fuel, gas, energy, engine, car, time, charge, solar, air, range, charging, water, change, ice, work, space, good	battery, power, electric, fuel, gas, energy, engine, charge, time, solar, car, air, range, charging, water, ice, change, work, space, good	aero, falcon, polar, band, servicing, rotation, washer, checkout, spill, thruster, abort, warmer, capsule, reusable, blade, turbine, coal, carbon, liter, gem	air, handle, factor, small, problem, high, long, depends, germany, skip, matter, ensure, amazing, hold, work, change, motor, big, minute, process
6 –Public Transit	road, car, lane, bike, traffic, driver, bus, city, time, street, parking, drive, ride, train, area, taxi, stop, light, speed, hour	road, lane, bike, traffic, car, driver, bus, city, street, parking, time, drive, ride, train, area, taxi, stop, light, speed, hour	ridership, avenue, metro, streetcar, cabby, oakland, donut, luggage, denver, alternate, merging, licensed, mot, suburb, parra, upfront, caput, calming, placard, calm	public, slow, turn, minute, photo, place, rule, fair, problem, work, regular, provide, money, crazy, live,

				stick, location, free, long, feel

The topics can be partially validated and understood by comparing topic frequency across the events. Figure 3 shows a heat map of topic frequency by event organized by event type (i.e., fatal crash, minor crash, technology advancements). The figure shows that the *Crashes* topic consistently appears with high percentage in fatal crash events, with the 2016 Florida Tesla crash, the 2018 Arizona Uber crash, and the 2018 California crash having the highest frequency of this topic. Tweets related to the *Tech Companies* topic tend to appear with higher relative frequency in advancement events. Interestingly, *Faulty & Safety* topic occurs in between one-quarter to one-third of all tweets across the events. Topics related to the *Public Transit*, *Electric Vehicles*, and *Market & Sales* topics are typically generally less frequent, except in the 2019 New Jersey Tesla crash, where the percentage of *Market & Sales* is around 32%. These results agree with the categorization and highlight the frequency of safety discussions regardless of events.

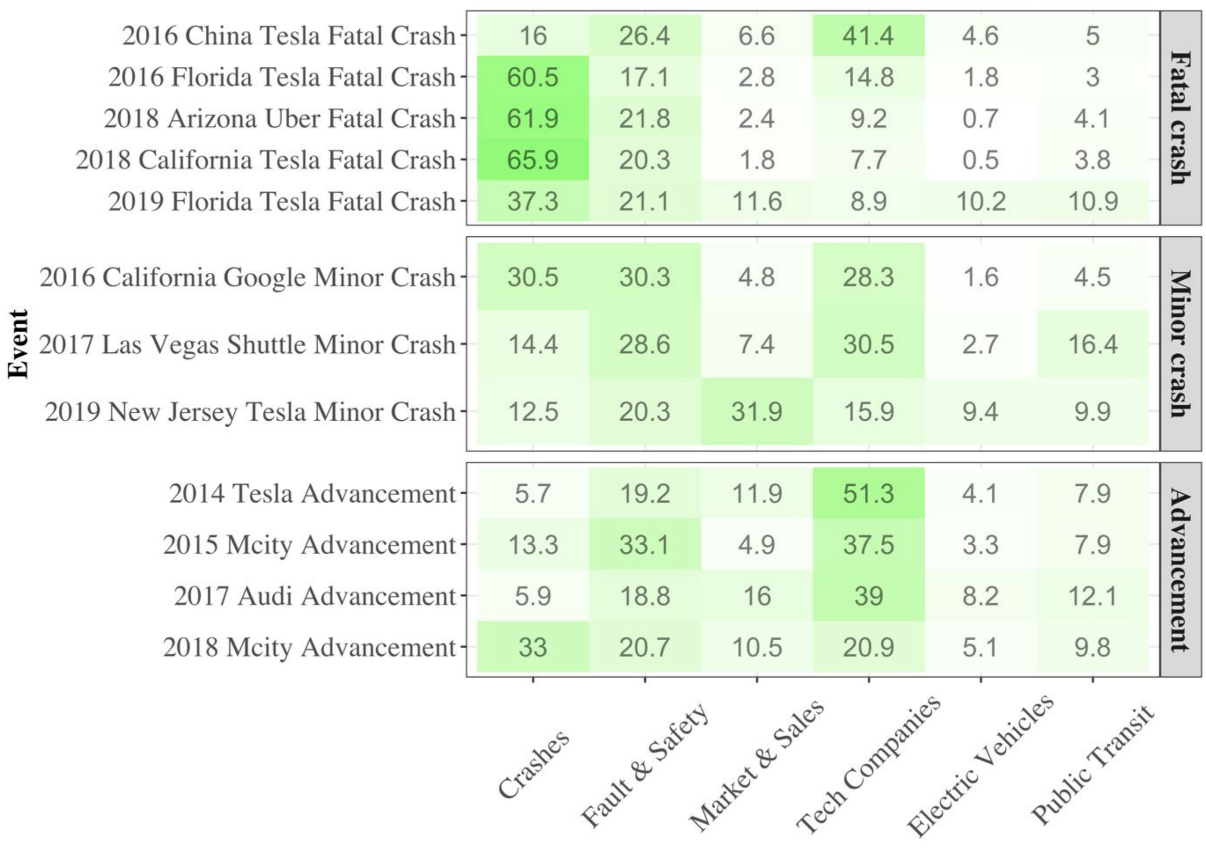


Figure 3. Heat map of the percentage of topics in each event. The transparency of the shading indicates lower (more transparent) or higher (less transparent) percentage of tweets.

Post-event Responses

Additional context into responses to the tweets can be gained by analyzing the frequency of tweets associated with each topic over time. Figure 4 shows the post-event tweet volumes for each topic and event; note that Day 0 in the figure corresponds to the date of the event. The figure clearly shows a dramatic increase in the *Crashes* discussions (the pink line and points in the figure) in the days following each crash event. Perhaps more interestingly, the discussions of the *Crashes* topic quickly dissipate within 3 days after their peak. However, at their peaks there are more tweets in the *Crashes* topic than any other topic. In a similar trend, the *Tech Companies* topic (blue line and dots in the figure) also peaks in the days following the announcement. It is notable that several shifts in the *Tech Companies* topic coincide with shifts in the *Fault & Safety* topic—for example, in the 2016 California Google crash and the 2017 Las Vegas Shuttle crash. While these trends corroborate expectations, there are some unexpected findings. Notably there is a spike in the *Crashes* and *Fault & Safety* topics in the 2015 Mcity announcement (the second chart in the top row). Among the 1,068 tweets in this category, 478 contain the words "NTSB" or "investigate." Thus, this spike is likely related to the release of a National Transportation Safety Board (NTSB) report on a prior California Tesla Crash. The *Market & Sales* topic (the green line and dots in Figure 4) was not dominant, but an increase was seen starting the sixth day after the 2019 New Jersey crash (the last chart in the bottom row), which is likely explained by coinciding announcements regarding changes in Tesla's leadership and their effects on Tesla's stock price.

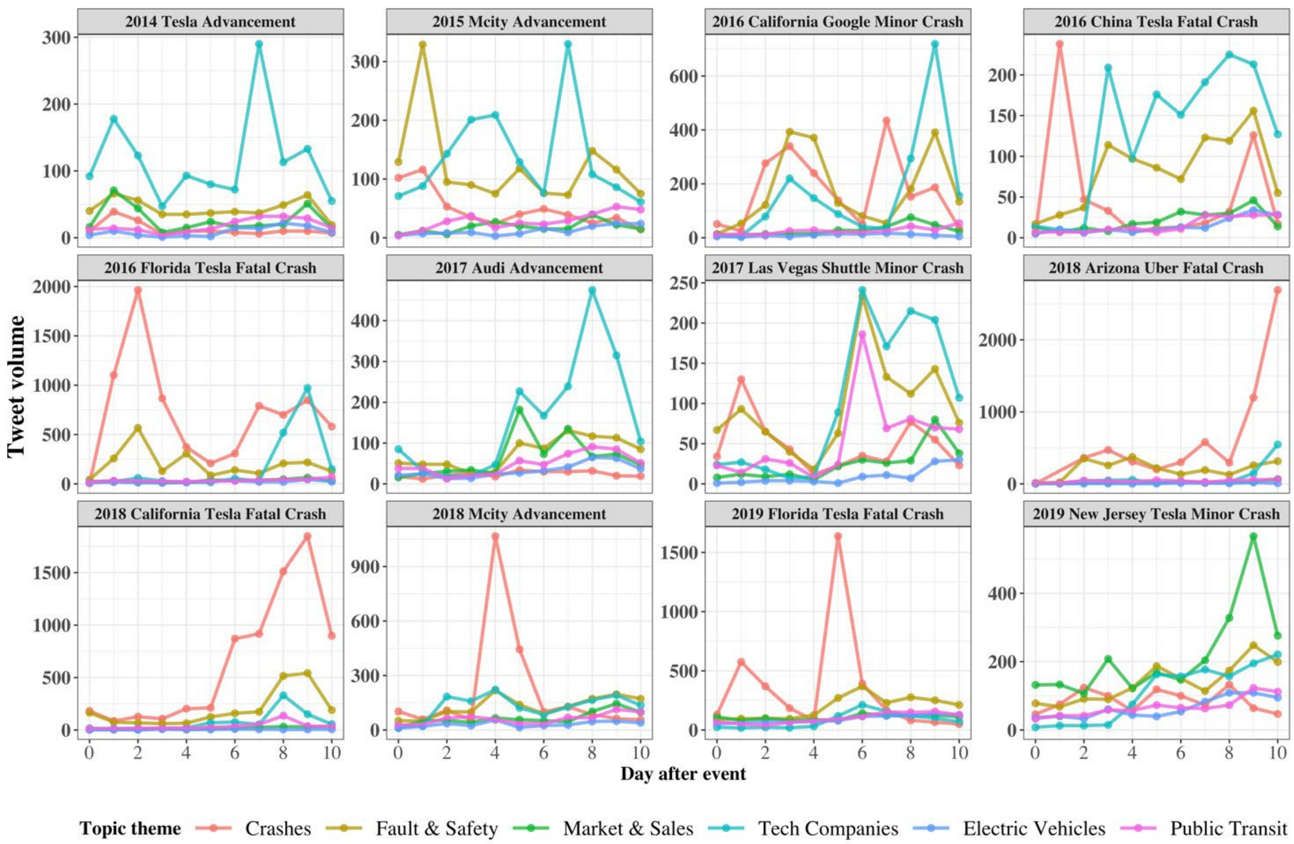


Figure 4. Graphs. Number of tweets after each event.

Sentiment Analysis

The sentiment analysis, summarized by topic in Figure 5, showed that most tweets identified by the search were neutral in sentiment. However, tweets associated with the *Crashes* and *Fault & Safety* topics were more negative than the other topics, especially when these topics were discussed after a fatal crash (see the top chart in the figure). Additional insight into these findings can be gleaned from analyzing the sentiments of the most strongly associated words with each topic. Figure 6 shows an example of this analysis with the 80 most strongly associated tweets by each topic. In the figure, a word's vertical position is the frequency of use in tweets associated with the topic, the *x* position represents the sentiment with some variability to prevent word over-plotting, and the color of the word represents its sentiment without variability. The figure illustrates that neutral words (e.g., car, driving, self) are the most frequent across all topics. However, the *Crashes* topic and, to a lesser extent, the *Fault & Safety* topic also have words with negative sentiment (e.g., crash, fatality, accident) among their most frequent words. The frequency of these words explains the negative sentiment distribution in Figure 5. One notable observation of this analysis is that there are some mismatches between the general sentiment dictionary labels and domain terms. For example, the term *autonomous* has a positive sentiment in the dictionary, but given its broad use in this context it is likely a neutral word. These findings suggest that there is a need to adjust the sentiment dictionary to the context in future work.

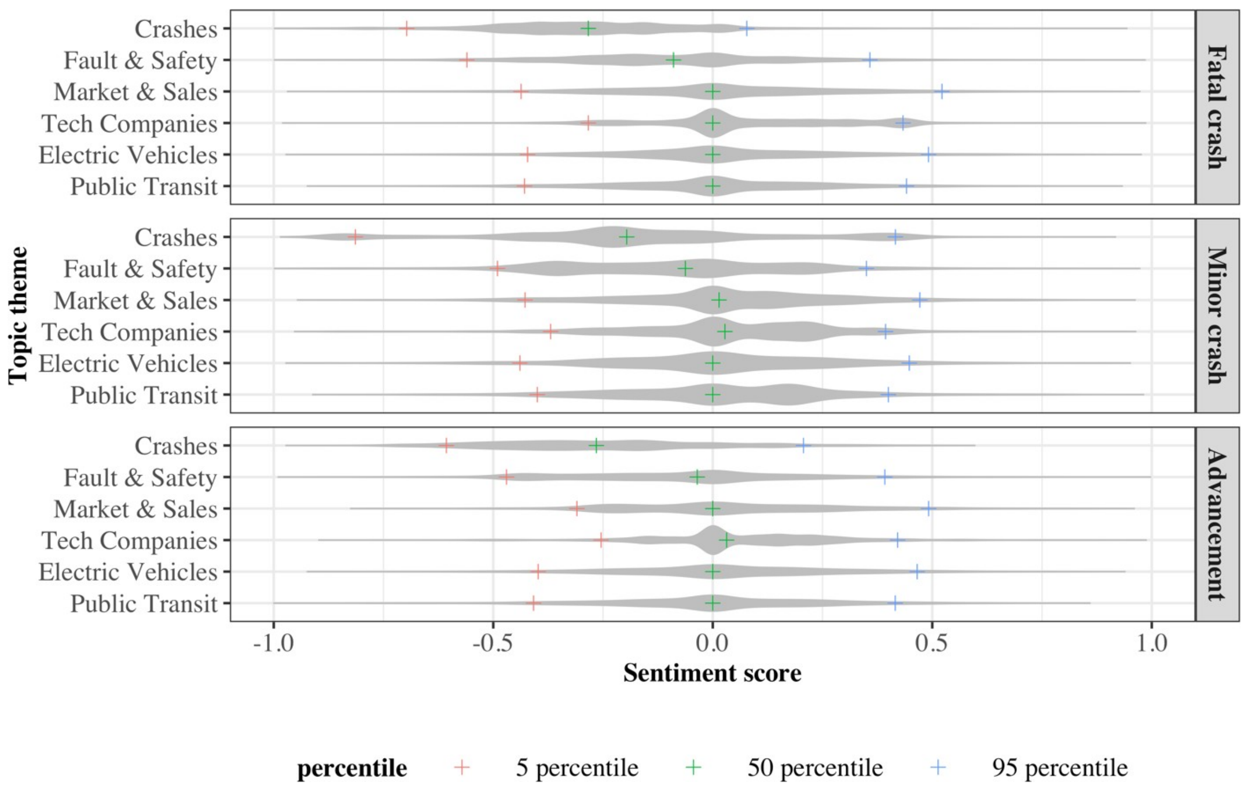


Figure 5. Chart. Sentiment density. Markers represent 5th, 50th, and 95th percentiles of sentiments for each topic and event type.

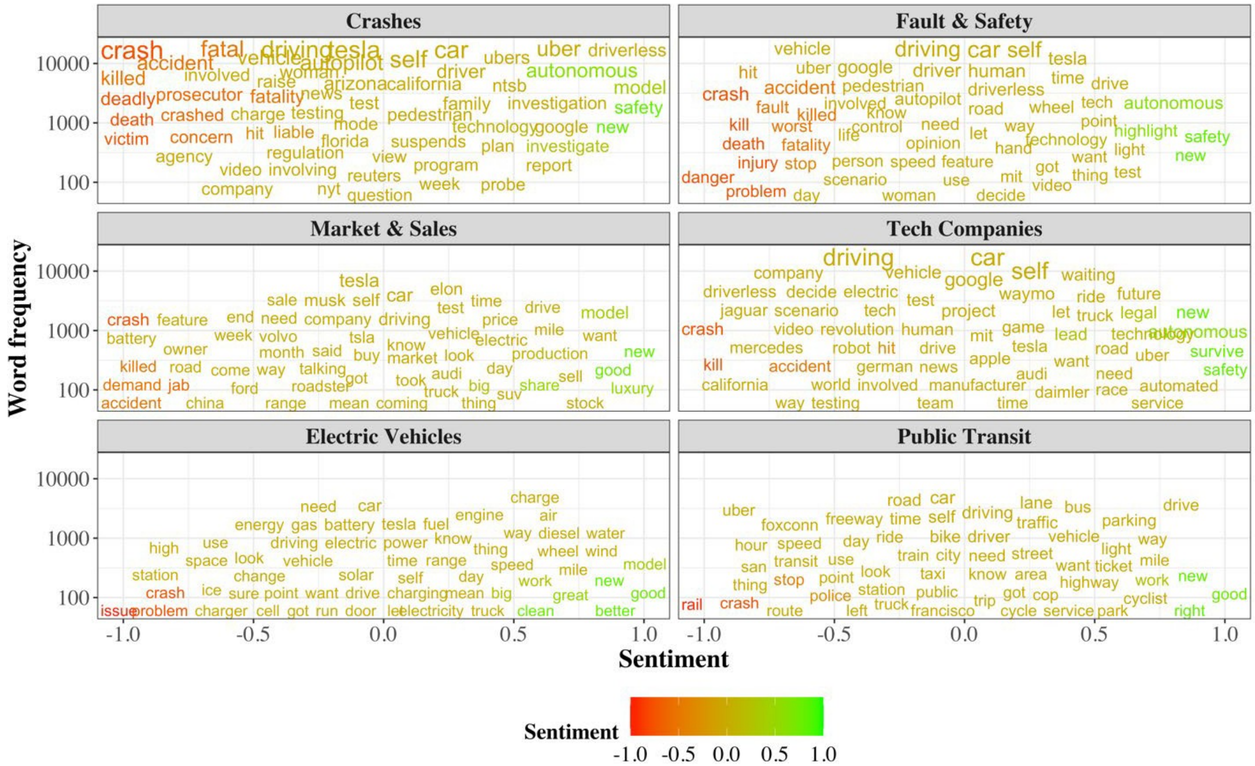


Figure 6. Charts. Top 80 most frequent words in each topic colored by sentiment score.

Discussion

The results illustrate that the most common themes in Twitter discussions of AV crashes and technology advancements center on *Crashes*, *Fault & Safety, Market & Sales*, *Tech Companies*, *Electric Vehicles*, and *Public Transit*. The frequency of these themes is heavily dependent on the type of event. Namely, fatal crashes increase discussions of in the *Crashes* and *Fault & Safety* topics, whereas technology announcements increase discussions in the *Tech Companies* topic. While this may be expected, the analysis also illustrates that discussions of these topics typically peak shortly after an event and decline steadily within a matter of days. Finally, while discussions are mostly neutral in sentiment, *Crashes* and *Fault & Safety* discussions are generally negative in sentiment. This negative sentiment is driven by words to describe crashes, including "crash," "accident," and "fatality."

While these findings are informative, they provide less information than expected regarding users' expectations and reactions to AV events. However, there are some important findings that may guide future communication on AV events. First, the timing of tweet frequency following fatal crashes suggests that it is critical to provide information in the first 10 days following a crash, most likely in the first 5 days. This timing will ensure that the tweet enters the discussion when it is most active. Second, there is considerable variance in the terms used to discuss AV technology. While some of this terminology is related to specific technology (e.g., "autopilot"), it is important to standardize terms. Finally, the prevalence of the *Fault & Safety* topic suggests that users are attracted to narratives that assign or discuss fault in AV crashes. Notably, this interest is especially peaked by reports from federal agencies, such as the NTSB report observed in the Mcity announcement event dataset. This suggests that officials communicating about crashes should be judicious about the use of the terms fault and safety.

These results highlighted several deficiencies in the application of topic modeling and sentiment analysis in the AV crash domain. First, there is a mismatch between sentiment assigned to words in a traditional sentiment dictionary and the likely sentiment of words in the AV domain. This suggests that there is a need to develop a domain-specific sentiment dictionary for AV crashes. Second, the event-based approach was effective for understanding topics of interest but it is limited in its ability to capture the dynamics of topics over time. Thus, there is a need to develop CLL extensions to provide more real-time streaming data from Twitter. Third, the analysis revealed that many of the identified tweets referenced or linked to web links—of which many were news articles. It is likely that many users communicated through the use of these links and that the sentiment of the articles would be a better measure of the user's sentiment. Finally, this analysis was limited in its exploration of the Twitter network as it relates to information transmission. Future work should explore the network in more detail to understand who is transferring information regarding AVs.

User-centered Design of AV Reporting Guidelines

The final phase of the project was to translate the findings from the Twitter analysis to a set of guidelines for stakeholders to communicate about AV crashes. After initial exploration, it was determined that the most relevant stakeholders were PIOs. PIOs are generally responsible for managing Twitter and other social media accounts for organizations. Our exploration also found that the PIOs most often engaged in transportation discussions on Twitter were local law enforcement and state departments of transportation (DOTs). Given these findings, we used a user-centered design process to identify the needs of PIOs and to develop an initial set of guidelines. User-centered design is a design process in which end users have significant input in the format of the final design (Abras et al., 2004). We employed user-centered design here through a semi-structured interview study where we interviewed PIOs on their role and needs for AV communication guidelines.

Methods

Eight PIOs with reported average related experience of 13 years (range 1 to 20 years) participated in the study. Five participants were from DOTs, and three participants were first responders (e.g., police officers). Participants were not required to have experience on reporting crashes with AVs, as we were interested in how they currently respond to crashes with non-AVs and how they would respond in the future to crashes involving AVs. Two of the participants had previous experience working with the publicity and rollout of AVs in their respective areas. Participants were interviewed either individually or in groups of two as their schedules permitted. After email consent, the PIOs were asked a series of questions from an interview protocol (see Appendix C). The interview protocol asked questions on the role of the PIO in their organization, their use of social media platforms, their experience with social media guidelines, their interactions with AVs, and their role in traffic safety communication. Throughout the discussion, interviewers asked clarifying and follow-up questions to gather additional details. Following the data collection, the interviews were transcribed and verified for accuracy. The transcriptions were analyzed by a single expert coder with an iterative thematic coding process to identify themes in the responses. These themes were consolidated with the findings from the AV Twitter analysis to develop a set of guidelines for reporting about AV crashes and discussing AVs.

Results

Thematic Analysis

Three themes were identified during the thematic analysis: guidelines for general social media use, guidelines for discussing AVs on social media, and guidelines for building trust with the public through social media. Each theme is comprised of multiple subthemes which create a holistic definition of each theme. There are some contextual differences between PIOs from different agencies. The PIOs from DOTs were primarily focused on relaying information concerning the roadways, closures, and traffic. The first responders were primarily concerned about the safety of road users. This difference also impacts the role that an individual's organization played in the subthemes and definitions. Finally, the communication allowed during crashes and the information each organization is permitted to share differ: DOT PIOs share information related to road closures, while PIO first responders share information about the crash and those involved. The complete thematic coding results are presented in Table 5, Table 6, and Table 7. Visualization of the themes and their respective subthemes and supporting quotes are presented in Appendices D, E, and F, respectively.

Table 5. Guidelines for General Social Media Use: Subthemes and Definitions from the Interviews with PIOs

Subtheme	Definition
Interaction with the public	Respond to questions from the public in a timely manner.
Working within organizational purview	Share only permitted information while allowing other organizations to communicate their messages.
Maintain relevant knowledge	Ensure knowledge of relevant trends or technologies is up-to-date.
Authority	Follow the mandates of leadership or acquire required sign-offs before posting information.
Presentation of information	Present factual accounts of situations or correct misinformation as necessary.
Etiquette	Treat all interactions as a professional encounter.

Subtheme	Definition
Brevity	Create short messages to deliver the main point more quickly.
Use of humor	Use humor sparingly; humor has its place in lighthearted themes but should not be used in serious matters (i.e., injuries, crashes, major safety concerns).
Consistency	Post content and interact with the public with regularity.
Social media persona	Ensure the use of a single persona across a variety of users on a social media account.
Filtering content	Filter out negative or profane content in replies or comments.
Firmness of guidelines	Create guidelines with clear expectations and do not allow deviations.

Table 6. Guidelines for Discussing AVs on Social Media: Subthemes and Definitions from the Interviews with PIOs

Subtheme	Definition
AV crashes presented like typical vehicle crashes	Do not present an AV crash differently from typical vehicle crashes as agency is typically assigned to the driver; await further information before sharing that the crash was with an AV.
Educate the public on AVs	Create a program with regular informative posts on AVs for the public.
Educate the organization on AVs	Create an educational program within the organization on AVs.
Partner with AV companies	Build relationship with companies that produce AVs to help promote information to the public.
Clarity in terms	Use clear terms that the public will understand with consistency to describe AVs.
The public already has preconceived notions of AVs	The public has a negative perception or distrusting attitude toward AVs that may be difficult to change.

Table 7. Guidelines for building trust with the public**: Subthemes and Definitions from the Interviews with PIOs**

Subtheme	Definition
The importance of locality	Place the local areas of interest over national (or other) interests in terms of information dissemination.
Interaction with the public	Create interactive posts and respond to the public's questions in a timely manner.
Presentation of information	Present information in an honest and factual manner.

Subtheme	Definition
Image	Present a professional, transparent and honest depiction of the organization.
Stability	Maintain consistency in posts and content across time as a trusted source.

Reporting Guidelines

When combined with the findings from the topic modeling analysis, the thematic coding results highlight the need for timely and accurate communication from first responders following a crash. The findings also highlight the importance of consistency in communication and educating the public. It is notable that in Figure 6, the terms "self," "driving," and "autonomous" are used nearly interchangeably (i.e., they have similar frequencies), yet these terms have substantially different meanings and are a poor reflection of current on-road technology (e.g., Tesla Autopilot), which is better described as "automated." This consistency will be especially important when it contrasts with brand names of technologies that may be misaligned with technological capability. Beyond the timing and terminology, the frequency of the theme *Fault & Safety* after crashes suggests that users are attracted to narratives that place fault on human drivers or automation. Given the need to present AV crashes like typical crashes and to educate the public with factual evidence, these findings suggest that discussions of fault should be limited following a crash. Finally, the need to educate the public suggests that there is a need to connect the public with standards of definitions of automated technology. While there are some limitations to categorizing automation into levels, the explanations provided by the Society of Automotive Engineers (SAE) offer an opportunity to educate the public on the maximal capability of various technologies. With these findings, we propose the following guidelines:

1. Communication within the first 5 days of a crash will be the most effective to get public attention.
2. Consistent terminology should be used when communicating about AV technologies. The most common terms referring to these technologies and their definitions are represented in Table 8.

Table 8. Most Common Terminologies Used for AVs and Their Associated Definitions

Name	Definition
Automated	A vehicle in which at least some aspects of a safety-critical control function (e.g., steering, throttle, or braking) occur with little or without direct driver input.
Self-driving	A vehicle having the ability to drive by itself using onboard sensors, without the need of any intervention from a human driver.
Driverless	A driverless car is a robotic vehicle that is designed to travel between destinations without a human operator.
Autonomous	A vehicle that is capable of sensing its environment and navigating without human input. A human may select a destination but is not required to mechanically operate the vehicle.
Connected	Connected vehicles can communicate with other vehicles, infrastructure, and devices through wireless network technology, such as Wi-Fi, GPS, or radio frequencies. A vehicle can be connected but not automated, automated but not connected, neither, or both.

3. In the event that technology brand names contrast with the capabilities of the technology, the term from the table that best describes the technology should be used.
4. The words “fault” and “blame” in reference to drivers or automation being “at fault” in a crash should be used sparingly.
5. Public education efforts should focus on clarifying technology capabilities and assuaging fear. The National Highway Traffic Safety Administration and SAE standards (National Highway Traffic Safety Administration, 2021) are good resources for this information.

Conclusions

The three phases of this project lead us to the following conclusions:

- CLL is an effective method of identifying relevant tweets centered on significant events.
- Crashes are a substantial theme in Twitter discussions of AVs.
- Crash discussions begin and end shortly after crashes or major news events (e.g., the release of an NTSB report).
- Sentiments of these discussions on Twitter are relatively neutral except for words describing crashes.
- PIOs should focus on efficiently conveying accurate information about AVs while maintaining a persistent interaction to educate the public on AV use and capabilities.

These findings also illustrated the limits of traditional Twitter analysis techniques such as sentiment analysis for AV crashes. The findings highlighted the need for additional investigations of domain-specific sentiment dictionaries and analyses that investigate the sentiment of news articles distributed through Twitter. Despite these limitations, the analyses here provide an important contribution to the literature on social media analysis, public opinion of AVs, and PIO interactions with the public.

Additional Products

The complete details on the project can be found at the project website located at: https://safed.vtti.vt.edu/projects/data-mining-twitter-to-improve-automated-vehicle-safety.

Education and Workforce Development Products

Dr. McDonald and Dr. Huang have developed a curated dataset of AV tweets that will be used for an activity in the Spring of 2021 in Dr. McDonald's course at TAMU, ISEN 413 Advanced Data Analytics. Dr. McDonald and his student, Ms. Jaycelyn Jefferson, developed an activity and application to introduce Twitter analysis to high school students. The app can be accessed at http://tamuhfml.shinyapps.io/twitter_activity_final (note that if the link is inactive, users should request access from mcdonald@tamu.edu). In addition, Dr. McDonald mentored three Ph.D. students, one M.S. student, and one undergraduate student's work on the project, and Dr. Huang mentored two Ph.D. students on the project. These students published two conference proceedings papers (Jefferson & McDonald, 2019; Wei et al., 2020) and submitted three additional contributions to highly competitive computer science conferences (NeurIPS, ICML, PAKDD).

Technology Transfer Products

The project has produced two accepted conference proceedings and one pending conference paper. The conference proceedings include an initial update on the progress of the project (Jefferson & McDonald, 2019), and a report on the final AV Twitter analysis (Wei et al., 2020). The submitted paper discusses the CLL method (Arachie & Huang, 2020) and was submitted to the Pacific Asia Conference on Knowledge Discovery and Data Mining, a popular conference on data mining. We plan to submit one additional conference proceedings and one journal article describing this work. Beyond these contributions, we will publish the guidelines we developed and release them to PIOs to guide their communication.

Data Products

The Twitter data gathered by the project is the property of Twitter and not directly publishable; however, we provide the tweet IDs for our collected tweets around each event at https://www.vtti.vt.edu/utc/safe-d/index.php/projects/[tbd]. Provided the original tweets have not been deleted by the users or Twitter, the data can be downloaded using the tweet IDs for later study.

References

Abras, C., Maloney-krichmar, D., & Preece, J. (2004). User Centered Design. In *Encyclopedia of Human-Computer Interaction*. Sage Publications. https://doi.org/10.3233/WOR-2010-1109

Arachie, C., & Huang, B. (2020). Constrained Labeling for Weakly Supervised Learning. *Preprint*. http://arxiv.org/abs/2009.07360

Arachie, C., & Huang, B. (2019). Adversarial label learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, *33*, 3183–3190.

Banks, V. A., Plant, K. L., & Stanton, N. A. (2017). Driver error or designer error: Using the Perceptual Cycle Model to explore the circumstances surrounding the fatal Tesla crash on 7th May 2016. *Safety Science*, *108*, 278–285. https://doi.org/10.1016/j.ssci.2017.12.023

Blei, D. M., Ng, A. Y., & Jordan, M. T. (2002). Latent dirichlet allocation. *Advances in Neural Information Processing Systems*, *3*, 993–1022.

Community Oriented Policing Services, & Police Executive Research Forum. (2013). *Social Media and Tactical Considerations For Law Enforcement*.

Deng, H., Wang, Q., Turner, D. P., Sexton, K. E., Burns, S. M., Eikermann, M., Liu, D., Cheng, D., & Houle, T. T. (2020). Sentiment analysis of real-world migraine tweets for population research. *Cephalalgia Reports*, *3*, 251581631989886. https://doi.org/10.1177/2515816319898867

Hagen, L. (2018). Content analysis of e-petitions with topic modeling: How to train and evaluate LDA models? *Information Processing and Management*, *54*(6), 1292–1307. https://doi.org/10.1016/j.ipm.2018.05.006

Hohenberger, C., Spörrle, M., & Welpe, I. M. (2016). How and why do men and women differ in their willingness to use automated cars? The influence of emotions across different age groups. *Transportation Research Part A: Policy and Practice*, *94*, 374–385. https://doi.org/10.1016/j.tra.2016.09.022

Honnibal, M., & Montani, I. (2017). *spacy 2: Natural language understanding with bloom embeddings, convolutional neural networks and incremental parsing*. To appear, 7(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. https://doi.org/10.1016/j.ssci.2017.10.001

Ikoro, V., Sharmina, M., Malik, K., & Batista-Navarro, R. (2018). Analyzing Sentiments Expressed on Twitter by UK Energy Company Consumers. *2018 5th International Conference on Social Networks Analysis, Management and Security, SNAMS 2018*, *October 2018*, 95–98. https://doi.org/10.1109/SNAMS.2018.8554619

Jefferson, J., & McDonald, A. D. (2019). The autonomous vehicle social network: Analyzing tweets after a recent Tesla autopilot crash. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *63*(1), 2071–2075. https://doi.org/10.1177/1071181319631510

Kyriakidis, M., Happee, R., & de Winter, J. C. F. (2015). Public opinion on automated driving: Results of an international questionnaire among 5000 respondents. *Transportation Research Part F: Traffic Psychology and Behaviour*, *32*, 127–140. https://doi.org/10.1016/j.trf.2015.04.014

Lee, C., Seppelt, B., Reimer, B., Mehler, B., & Coughlin, J. F. (2019). Acceptance of Vehicle Automation: Effects of Demographic Traits, Technology Experience and Media Exposure. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *63*(1), 2066–2070. https://doi.org/10.1177/1071181319631425

Lee, J. D., & Kolodge, K. (2019). Exploring Trust in Self-Driving Vehicles Through Text Analysis. *Human Factors*. https://doi.org/10.1177/0018720819872672

Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

Li, T., Lin, L., Choi, M., Fu, K., Gong, S., & Wang, J. (2018). YouTube AV 50K: an Annotated Corpus for Comments in Autonomous Vehicles. *Proceedings of the Thirteenth International Joint Symposium on Artificial Intelligence and Natural Language Processing (ISAI-NLP 2018)*. http://arxiv.org/abs/1807.11227

Loper, E., & Bird, S. (2002). *NLTK: The Natural Language Toolkit*. http://arxiv.org/abs/cs/0205028

Luna, S., & Pennock, M. J. (2018). Social media applications and emergency management: A literature review and research agenda. *International Journal of Disaster Risk Reduction*, *28*(October 2017), 565–577. https://doi.org/10.1016/j.ijdrr.2018.01.006

McCallum, A. K. (2002). *MALLET: a machine learning for language toolkit.*

Naldi, M. (2019). *A review of sentiment computation methods with R packages*. 1–11. http://arxiv.org/abs/1901.08319

National Highway Traffic Safety Administration. (2021). *The road to full automation*. https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety#topic-road-self-driving

Nees, M. A. (2016). Acceptance of Self-driving Cars. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *60*(1), 1449–1453. https://doi.org/10.1177/1541931213601332

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *39*(2), 230–253. https://doi.org/10.1518/001872097778543886

Pennington, J., Socher, R., & Manning, C. D. (2014). Glove: Global vectors for word representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1532–1543.

Raisi, E., & Huang, B. (2018a). Weakly supervised cyberbullying detection with participant-vocabulary consistency. *Social Network Analysis and Mining*, *8*(1), 38.

https://doi.org/10.1007/s13278-018-0517-y

Raisi, E., & Huang, B. (2018b). Weakly Supervised Cyberbullying Detection Using Co-Trained Ensembles of Embedding Models. *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 479–486. https://doi.org/10.1109/ASONAM.2018.8508240

Ratner, A., Bach, S. H., Ehrenberg, H., Fries, J., Wu, S., & Ré, C. (2017). Snorkel: Rapid training data creation with weak supervision. *Proceedings of the VLDB Endowment. International Conference on Very Large Data Bases*, *11*(3), 269.

Ratner, A., Hancock, B., Dunnmon, J., Goldman, R., & Ré, C. (2018). Snorkel metal: Weak supervision for multi-task learning. *Proceedings of the Second Workshop on Data Management for End-To-End Machine Learning*, 1–4.

Reuter, C., Ludwig, T., Kaufhold, M. A., & Spielhofer, T. (2016). Emergency services′ attitudes towards social media: A quantitative and qualitative survey across Europe. *International Journal of Human Computer Studies*, *95*, 96–111. https://doi.org/10.1016/j.ijhcs.2016.03.005

Rinker, T. (2019). *Package "sentimentr": Calculate Text Polarity Sentiment*. 59. https://cran.r-project.org/web/packages/sentimentr/sentimentr.pdf

Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the space of topic coherence measures. *WSDM 2015 - Proceedings of the 8th ACM International Conference on Web Search and Data Mining*, 399–408. https://doi.org/10.1145/2684822.2685324

SAE International. (2018). *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (Technical Report No. J3016)*. https://saemobilus.sae.org/content/j3016_201401

Schoettle, B., & Sivak, M. (2015). *A Preliminary Analysis of Real-World Crashes Involving Self-Driving Vehicles* (Issue October). https://doi.org/UMTRI-2015-34

Seppelt, B. D., & Victor, T. W. (2016). Potential solutions to human factors challenges in road vehicle automation. In G. Meyer & S. Beiker (Eds.), *Road Vehicle Automation 3* (Issue August, pp. 131–148). Springer International Publishing. https://doi.org/10.1007/978-3-319-40503-2

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

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. https://doi.org/10.1177/0018720818788164

Wei, R., Alambeigi, H., & McDonald, A. D. (2020). Topic modeling social media data after fatal automated vehicle crashes. *Presented at the Human Factors and Ergonomics Society's 2020 International Annual Meeting*.

Appendices

Appendix A: Words used to filter topics

Topic type	Keywords
AV relevant	Tesla, autopilot, autonomous, crash
AV irrelevant	Trump, liberal, election
Other	Government, tax, market, road

Appendix B: Topic modeling metric definitions

Metric	**Definition**
FREX	The harmonic average of word-conditional-topic probabilities (frequency) in a given topic and the sum of word-conditional-topic probabilities in other topics (exclusivity). This measure ranks words that are both strongly associated with a topic and exclusive to a topic high.
Lift	The ratio between word-conditional-topic probabilities and empirical counts of a given word. This measure ranks words that are both strongly associated with a topic and infrequent in the corpus high.
Score	The logarithm of the ratio between word-conditional-topic probabilities in a given topic and the average word-conditional-topic probabilities in other topics. This measure ranks words that are both strongly associated with a topic and exclusive to a topic high.

Appendix C: Topic model example tweets by topic

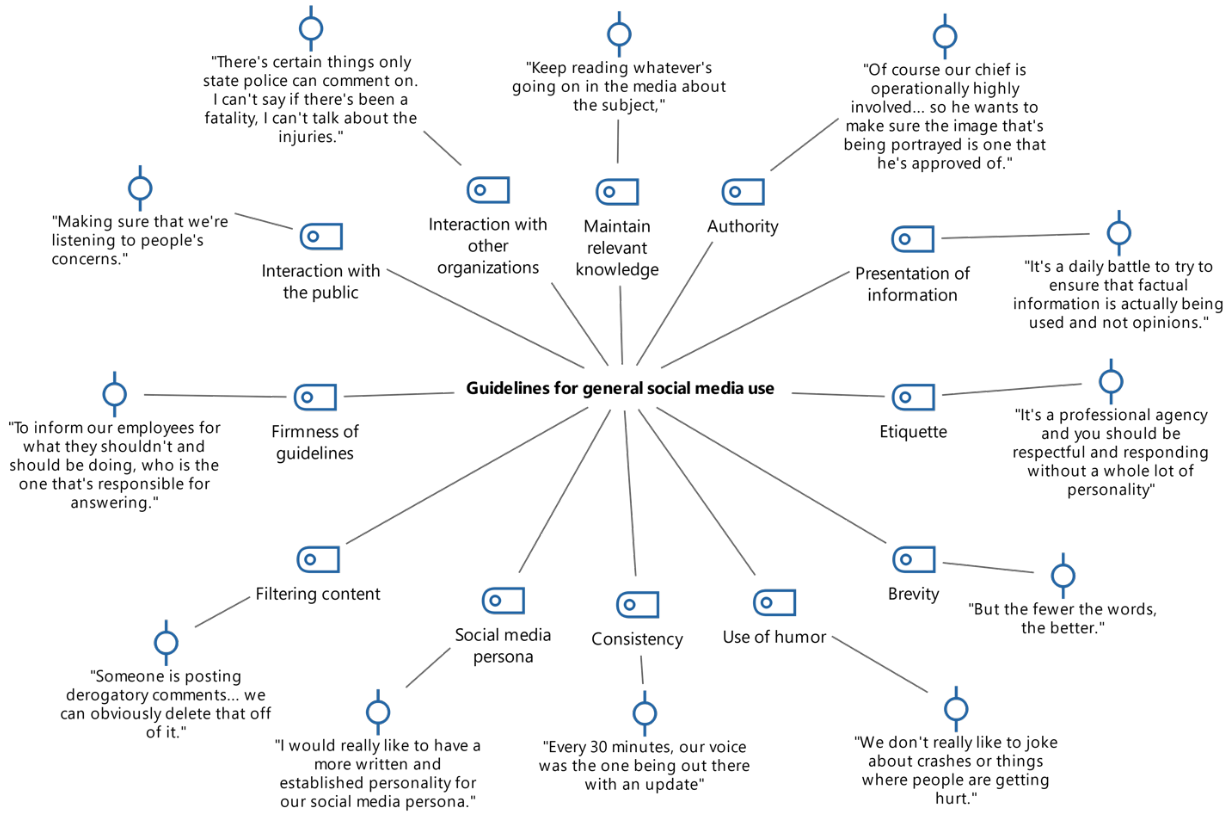
Topic	Sample tweets
Topic 1: *Crashes*	Arizona governor suspends Uber's self-driving car tests As the investigation into last week's fatal crash where an autonomous Uber SUV struck and killed a pedestrian in Tempe, AZ, the state's governor has suspended Uber's permission to test its cars there… https://ift.tt/2DXDEY2 Video of deadly Uber autonomous car crash raises more questions than it answers Uber has put the brakes on its experimental autonomous vehicles in Phoenix, Pittsburgh, San Francisco, and Toronto following a deadly crash between an Uber vehicle and a woman https://goo.gl/Dg3QBm 2 Teens Killed in Fiery Tesla Crash in Fort Lauderdale: Two Fort Lauderdale teens were killed and a third was injured after a Tesla Model S crashed and caught fire, the Fort Lauderdale Police Department said. FLPD said the single-car crash that happened… http://dlvr.it/Qy6GJKpic.twitter.com/eoQeLt6u6F US authorities are investigating a fatal Tesla Inc Model S crash in Florida last Sunday that killed the driver and caused a massive fire, the second fatal Tesla crash in the state this week being probed, the National Highway Traffic Safety Administration (NHTSA) said on Saturday
Topic 2: *Fault & safety*	Speed limiters feature on trucks. Cars & trucks have seatbelts. Modern cars have airbags. What safety measures are on your bike? Lights? Helmets an essential tool, helps keep u safer. Bikes aren't cars. You can't be vulnerable road users & then argue helmets not compulsory. In this tragic case, the driverless car had a human assistant who also failed to hit the breaks. The unfortunate victim was crossing the middle of the road at night. If the sober human assistant failed to hit the breaks, I'm almost sure a drunk driver would also have failed to. Uber autonomous car (but with a safety driver in car) driving in an autonomous zone (read very controlled area specifically designed and equipped) killed a pedestrian crossing road. Many legal and technical issues until common place usage This is a great article! Many news reports fail to discuss the technology and suggest blaming the "human" non-driver. Radar, Laser, Lidar and other sensors failed. The car did not break or change trajectory before striking the pedestrian. BIG AI FAIL
Topic 3: *Market & sales*	The big EV push in Europe will come in 2020 when CO2 super credits count. Merc and Audi EV pdn plans for China hardly insignificant. Porsche wages to fund Taycan? Not true. Low wages across the German economy since Schroder was chancellor. Worker share dividend is generous. Top graph is unit sales. The grand tour had a good episode this week on how absurdly expensive imports are in China. The Chinese Tesla is a bigger deal than I thought. With EV incentive pressures nearly forbidding ICE sales and domestic production it will crush the Germans. Tesla's entire range is now on competitive pricing compared to similar ICE - remember conv about Model SP100D vs BMW M5? Tesla was more expensive but now with almost 35% cut on perf model it is lower prices than M5 and this not including any gas saving. Last nail in ICE coffin Norway fiscal incentives (compared to ICEs) leaves every other country in its shadows. Diesel trade-in in Germany (extended twice by OEMs) should end June 30th. €4k will continue though. Germany on graph should flatten out from around June on since boost started then last year
Topic 4: *Tech companies*	Apple poaches senior self-driving engineer from Waymo(Reuters) - Apple Inc has hired a senior self-driving car engineer from Alphabet Inc's Waymo unit, Apple said on Friday, a sign that the iPhone maker maintains autonomous vehicle ambitions. copyright © 2016Mar… Waymo Adds 20,000 Autonomous Jaguar I-Pace SUVs to Test Fleet - Long-term strategic partnership puts driverless I-Pace EVs on Phoenix streets Google's self-driving car division Waymo and Jaguar announced a new long-term stra... pic.twitter.com/IJtowj3UOm Waymo will expand self-driving services to Europe, CEO says: Waymo aims to expand beyond the U.S. by entering Europe, potentially offering a mobility service with a fully driverless car fleet

	in cooperation with a local partner, CEO John… http://dlvr.it/QWTt53 Automotive pic.twitter.com/tu7f4pVmcd Ford returns to Detroit with its dedicated global electric vehicle organization, Ford Team Edison, plus its autonomous vehicle business team. This will accelerate Ford's push into electrification and strengthen its development of self-driving vehicles. https://www.rmagroup.net/all-news-update/historic-corktown-becomes-home-to-ford-motor-company/
Topic 5: *Electric vehicles*	Water can be heated in many ways, including solar thermal collectors or heat pumps run on wind or solar PV, where the heat source is either the air, ground, water, or waste heat. The water is then piped underground to heat soil or stored directly in water pits or aquifers. And lower Co2 and nearly zero NoX highest thermal efficient engine and Gen 4 hybrid system charges faster and gives more output, Diesel MPG from cheaper cleaner petrol compared to Diesel, brake pad emissions reduced as charging system slows vehicle down. Just some of the benefits Should be fine then. Assume 25%+ cold weather range loss depending on factors including battery temp, outside temp, heater temp, seat heaters, etc. 50 Km one way with ability to charge at both ends means you can start with a full charge (90%) within a couple of hours. A mild hybrid recovers kinetic energy into its battery, a plugin hybrid recovers KE into its battery, a pure BEV recovers KE into its battery, so all three self-charge in the same limited way thus not a point of difference. All need external primary energy from fuel & electricity
Topic 6: *Public transit*	"at rush hour some station dwells stretch over minutes due to the railroad's uniquely high number of passengers with bicycles" Stupid policy. Famously bike-friendly Netherlands bans bikes at peak and charges extra out-of-peak (except folded bikes = hand luggage). Use bike P+R. Yellow taxis are a garbage service, just the other night a yellow taxi threw out my 9 month pregnant wife and I because traffic was too heavy on FDR going to Inwood. Took the nearest exit and said he had car trouble. Garbage service and I hope Uber & Lyft continue to expand. She was likely crossing over to use the Northbound bike lane. Also note that the dotted white line defines the zone where right-turning traffic crosses over the bike lane (chicane). There is a sign that warns drivers to Yield to Bikes about 40ft uproad from the red spot. Municipal taxes fund municipal roads. And income taxes fund regional roads. Every taxpayer, including people who walk, take transit, and ride bikes, pays. But, unlike people in cars, they impose none of the other societal and maintenance costs. IOW, they\u2019re subsidizing drivers.

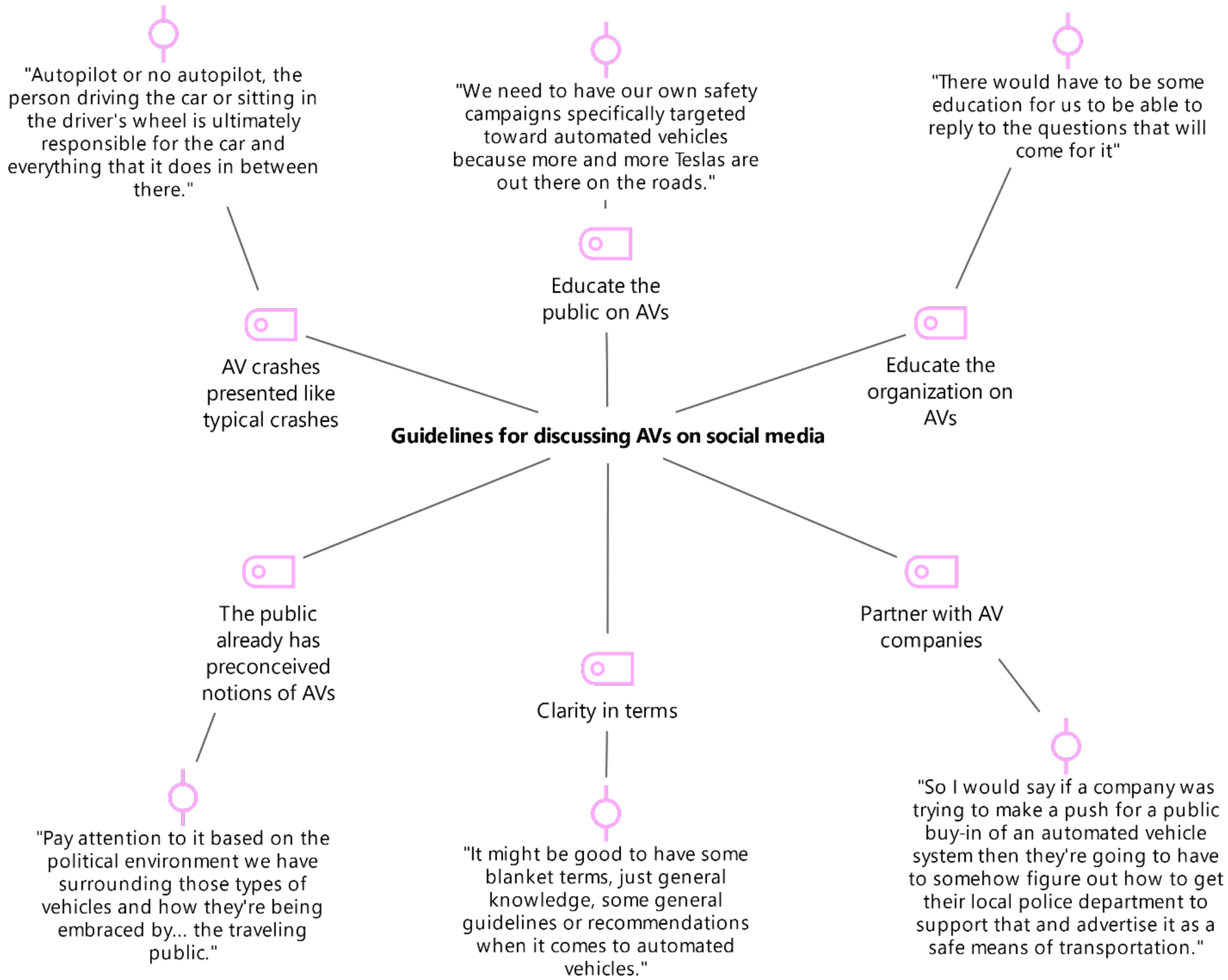
Appendix D: Semi-structured interview guide

1. What is your role and responsibility as a Public Information Officer?
2. What does a typical day look like for a Public Information Officer at your organization?
3. Tell us about your experience with different social media platforms as a Public Information Officer.
4. In what ways do you think social media communication can help build trust between you and your audience?
5. How do you think social media will impact your job as a Public Information Officer in the next few years?
6. Does your organization have specific guidelines for using social media? If so, what is your level of interaction with those guidelines?
7. Are there any aspects of those guidelines you wish you could change?
8. Explain why not having a proper system structure for sharing critical information via social media can be detrimental to an organization's credibility?
9. Does your organization have a role in communicating about vehicle crashes? If so, what is that role?
10. What do you think about the way your organization responds to serious vehicle crashes?
11. What images come to mind when I mention the term "automated vehicles"?
12. Do you see your organization being involved in social media communication about automated vehicles currently or in the future?
13. Do you feel like there is a need for guidelines on using social media for reporting vehicle crashes?
14. If so, what are some of the features you think those guidelines should include?
15. Do you have any closing comments?

Appendix E: Visualization of guidelines for general social media use theme with supporting quotes



Appendix F: Visualization of guidelines for discussing AVs on social media theme with supporting quotes



Appendix G: Visualization of guidelines for building trust with the public theme with supporting quotes