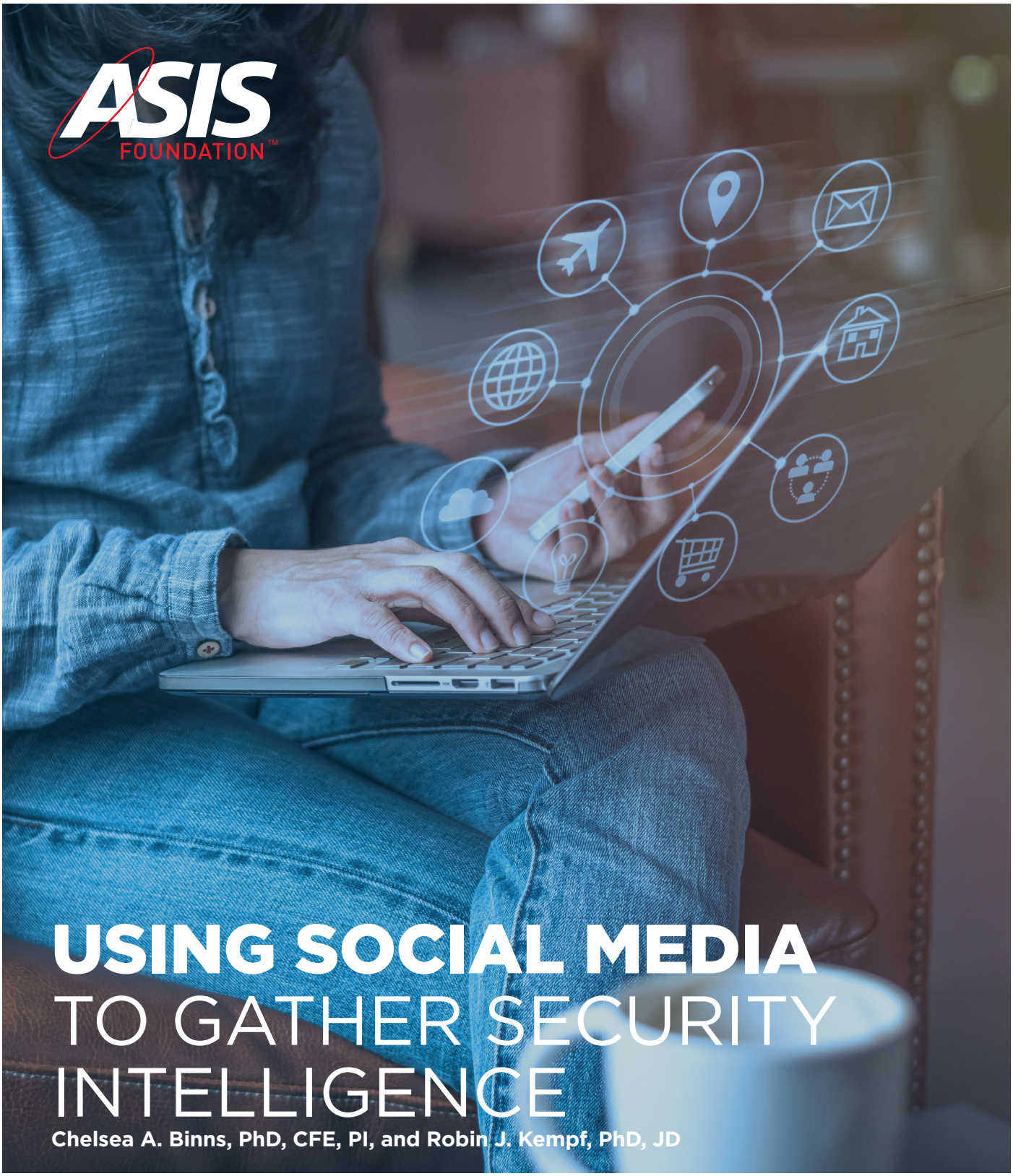


USING SOCIAL MEDIA TO GATHER SECURITY INTELLIGENCE

Chelsea A. Binns, PhD, CFE, PI, and Robin J. Kempf, PhD, JD

Table of Contents

Authors

Chelsea A. Binns, PhD, CFE, PI
Assistant Professor
John Jay College of Criminal Justice
Department of Security, Fire, and Emergency Management
524 W 59th Street
New York, NY 10019
cbinns@jjay.cuny.edu

Robin J. Kempf, PhD, JD
Assistant Professor
University of Colorado Colorado Springs
School of Public Affairs
1420 Austin Bluffs Parkway
Colorado Springs, CO 80918
rkempf@uccs.edu

Acknowledgment

The authors wish to acknowledge the hard work and contributions of our research assistants, Julien Roussel and Shea Hastings Connors, and our proofreader, Peter Haxton. Thanks also to Asher Fergusson with https://www.asherfergusson.com/ for his expert consultation and support, and to Germán Sanchís with Sciling (https://sciling.com/), who supported the artificial intelligence for this project.

This work was supported by funding from the ASIS Foundation and an Emergency Funding grant from the Office for the Advancement of Research at John Jay College of Criminal Justice.

ASIS International | 1625 Prince Street | Alexandria, Virginia, USA 22314

EXECUTIVE SUMMARY

Today, social media data are ubiquitous. This study asked whether security professionals could use these data to increase their understanding of security risks related to their industry.

Here, Twitter data relating to two companies in different industries, homesharing and ridesharing, were analyzed using qualitative methodologies and artificial intelligence. Indeed, this study determined that serious security concerns, such as hacking extortion, theft, and sexual harassment are prevalent in the services these organizations promote. The results demonstrate that social media data can provide organizations with actionable security information. The information provided in this report serves as a blueprint for organizations to replicate and conduct their own analyses.

For this research, the targets of the analysis are referred to as Stay, Inc., a well-known homesharing company, and Ride, Inc., a popular ridesharing company.

SUMMARY OF RESEARCH AND TAKEAWAYS FOR SECURITY PROFESSIONALS

This research study asked: Can social media data, analyzed using AI, provide the security practitioner with unique, actionable information that can improve safety and security? To answer this question, the following sub questions were explored:

- Does social media data contain valuable security-related data?
- Can AI technologies enhance the analytical process?
- Can information gleaned from AI improve security practices?

This study revealed that using artificial intelligence to analyze social media can provide a wealth of information to security professionals. One example relates to reports of Stay, Inc., hacking. Tweeters who were hacked reported having money stolen from their linked accounts. They were also locked out of their own accounts, where they had other trips booked. This means their future stays were also security-compromised, because now the hackers know where the tweeter is staying.

Without analyzing Twitter, a security professional would not likely have known the extent of this problem. Despite the volume of tweets from users complaining their accounts were hacked, there is little public information discussing this issue. Stay, Inc.'s website does not discuss the issue of account hacking, and it was found that very limited news articles exposed the issues. In fact, in those articles, much of the information the authors used in their story came from Twitter. Consequently, social media appears to be the primary source of this information, which appears to indicate that the companies need to address the issue publicly. Until that happens, Twitter will continue to be a rich source, and perhaps the only source, of this information.

There are many other important takeaways here for security professionals. This analysis shows that there are multiple opportunities to be gleaned, including:

- Social media is a rich source of security-relevant information.
- The general scale and scope of these issues can be ascertained from social media analysis
- Details about specific security and safety incidents can be obtained from social media
- Specific details are given on social media that can serve as evidence
- Policy gaps can be identified from information learned on social media

In addition, specific actionable security information for Stay, Inc., and Ride, Inc., were identified from this analysis, including:

-
- The need to educate users. For example:
    - Users of these services should be encouraged to always be vigilant.
    - Users should understand that there is a benefit for them to independently research the area of a homeshare and rideshare before booking.
    - Users should review the safety

information for homesharing and ridesharing service provider websites before using the services.

- Users should be encouraged to take pictures and video, if it is safe to do so, as these may be used as evidence in fraud-related matters involving these services.
- Homeshares and rideshares can be cancelled unexpectedly by hosts or drivers resulting in guests and passengers being stranded. As such, users should have a backup plan.
- There are extra fees that may be charged for both services, such as cleanup fees. Users should be informed about the fees as well as coached about how to avoid these fees and how to react when charged.
- Users should not expect to reach the service providers in the event of an emergency and should instead have the number of a local law enforcement agency handy.
- Users should be counseled to inspect homeshares and rideshares upon entering. If deemed unsafe, users should leave and report it immediately.

- Incidents requiring further investigation. For example,
  - Hacking, which appears to be a pervasive security issue, requires close attention.
  - Drivers reported to be under the influence of drugs or alcohol.
  - Homeshares and rideshares reported to have guns or drugs.
  - Situations where users report harassment or assault.
  - Reported hidden cameras in bathrooms and bedrooms.
- The need to consider policy changes. For example:
  - To reduce the instance and number of times a user is hacked, a dedicated resource for people to report hacking could benefit the users and the company.

CONCLUSION

Ultimately, this study revealed that social media can provide a wealth of information to security professionals. With the assistance of artificial intelligence, large amounts of social media data can be analyzed efficiently and effectively. Although AI requires technical know-how, which represents costs for AI contractors, this research demonstrates how security issues, including those that had not previously been identified from other sources, can be discovered, to which security professionals can respond.

FINDINGS

There are security risks in every industry. As such, security professionals are charged with ensuring the safety and security of property and assets of the companies they work for. This role includes performing surveillance, pursuing investigations, and reporting findings. In addition, a security professional should be responsible for looking out for new or escalating risks to identify and fix weaknesses that could be exploited by criminally minded individuals.

Security practitioners need actionable information to do their jobs efficiently and effectively. Important decisions about the safety and security of their organizations, the employees, and customers can be challenging when they simply lack information. The true security status of many organizations, including major service providers, can be challenging to ascertain as corporate security data for many companies is private and not available to the general public. As a result, it is often necessary for security personnel to use public data to determine any potential security-related issues.

Yet, when security professionals are scanning for security risks, they have largely overlooked an important source of public data: social media. As defined on Investopedia.com, social media consists of:

> *Computer-based technology that facilitates the sharing of ideas, thoughts, and information through the building of virtual networks and communities. By design, social media is internet-based and gives users quick electronic communication of content. Content includes personal information, documents, videos, and photos. Users engage with social media via computer, tablet or smartphone via web-based software or web application, often utilizing it for messaging" (Dollarhide, 2020).*

The nature of social media as informal communications on daily experiences makes it a good source of information about security issues. For example, an individual may report to their friends on Facebook about a problem they encountered with a specific business or location, or users of a service might complain on Twitter about security and safety-related issues they have experienced. Company-specific pages also allow consumers to contact them directly. This practice is encouraged by Consumer Reports, a nonprofit consumer advocacy group, when consumers need to "get the attention of a company... [especially] when frustrated with its products or services." (Doyle, 2019) Companies often try to collect similar information by creating toll-free hotlines or comment forms on websites; however, many individuals may not want to make a formal complaint. Alternatively, they turn to social media out of frustration. Since most sources of social media are broadcast to the general public, it becomes relatively easy for a security professional to access.

That said, the amount of social media data available to the general public is voluminous and is not provided in a format that can be easily digested. Thankfully, artificial intelligence (AI) is making this process a bit more feasible than it was in the past. With AI, this data can be analyzed in a more efficient manner, allowing for the researcher to quickly glean security-related intelligence from the data.

Thus, this research study asked: Can social media data, analyzed using AI, provide the security practitioner with unique, actionable information that can improve safety and security? To answer this question, the following sub-questions were explored:

- Does social media data contain valuable security-related data?
- Can AI technologies enhance the analytical process?
- Can information gleaned from AI improve security practices?

The study found that social media datasets contain information that can be useful to the security provider; however, it also recognized that analyzing that data can be an onerous task. AI is not an exact science. Hard work is required from a capable security team, and perhaps an AI expert, to get the data analyzed and achieve reliable findings.

The main purpose of this research project is to provide security practitioners guidance on how to mine social media data for security-relevant intelligence. This report informs the security

practitioner about the processes necessary to conduct a rigorous qualitative analysis of publicly available social media data with the help of AI, using two case studies where security data was analyzed as examples. This will be helpful to practitioners who can review the process involved in this study and determine whether it can and should be replicated by their security teams.

Specifically, this research examined tweets—posts on the social media platform Twitter, about two companies in two areas of the growing gig economy: homesharing (here referred to as "Stay, Inc.,") and ridesharing (here referred to as "Ride, Inc.,"). AI methods were used to help analyze vast amounts of social media data. In the end, this exercise demonstrated the value of analyzing social media data for safety and security purposes. It also highlighted some challenges about the methodology that security personnel should be aware of. In the end, this research identified several security risks related to Stay, Inc., and Ride, Inc., that were otherwise relatively unknown to the public.

LITERATURE REVIEW

To date, there is relatively little research on artificial intelligence (AI) in the security field. That said, AI has been recognized for its potential because of its ability to review vast amounts of data quickly and help users draw conclusions about security risks and actions. The following examples demonstrate a few of AI's possible uses that have been explored in the literature.

- Supporting national security:
    - Monitoring for threats can be improved as "neural networks can scrutinize surveillance video and alert soldiers to specific frames that contain objects of interest such as vehicles, weapons, or persons. Facial-recognition software could alert soldiers when an individual of interest is observed in video surveillance or in real time" (Wasilow and Thorpe, 2019, p. 37)
    - AI can be used to identify targets more discreetly to minimize casualties and collateral damage (Gill, 2019).
    - Tactical decision-making can be enhanced by "identifying and assessing tactical courses of action, coordinating distributed warfare resources, and incorporating predictive war-gaming into tactical decisions" (Johnson 2019, p. 64).
- Protecting computing infrastructure: AI can be used to analyze how hackers will try to infiltrate systems (Fugate & Ferguson-Walter, 2019).
- Investigating and preventing crime:
    - AI can analyze metadata from sex services advertisements on the Internet to help identify whether women are voluntarily advertising or being coerced to do so (Radulov, 2019).
    - Credit card fraud and other financial frauds, including money laundering and securities fraud, have been identified through the use of AI (Kirkland et al. 1999; Fawcett et al., 1998; Senator, 1995).
    - AI can be used to identify fraud, waste, and abuse in health care billing by examining large sets of data and identifying anomalies (Liu, 2016).
- Reducing the risk of terrorist attacks: AI can develop plans for randomized patrolling and monitoring, which prevent terrorists from easily analyzing security patterns and weaknesses (Pita et al., 2009)

Acknowledging that security policies, regulations, and ethical frameworks have lagged behind the introduction of AI (Wasilow & Thorpe, 2019), AI provides a great potential to security professionals when it comes to analyzing social media data. Because AI can be trained to analyze and look for specific details on the web or any electronic space, it can be trained to look through huge datasets culled from social media to identify safety and security concerns and new areas of crime. The present study uses Twitter data.

DEFINITIONS AND IMPORTANT CONCEPTS

WHAT ARE SAFETY CONCERNS

In this report, safety concerns are those issues that an individual might worry about when staying in a property advertised by a home share company or accepting a ride offered through a rideshare company that occur unintentionally or by accident. These issues, from a fire to a car

crash, may be caused by human carelessness, inattentiveness, lack of training, or other unintentional events (Fennelly, 2016).

WHAT ARE SECURITY CONCERNS

Although scholars have noted the challenge when it comes to defining security (Brooks, 2010), in this report, security concerns are those issues that an individual might worry about when staying in a property advertised by a home share company or accepting a ride offered through a rideshare company that are a result of intentional, malevolent human actions. These may include theft, vandalism, physical violence, terrorism, or other intentional attacks (Fennelly, 2016).

WHAT IS TWITTER?

Twitter is a popular microblogging social network that allows people to send and read public messages about any topic. Twitter is unique in that it limits the size of your messages to 280 characters, and it allows users to send and read messages from anyone, unlike some social media platforms where connections must be established before communication will occur (Kwak et. al, 2010). Posts on Twitter are known as "tweets."

Worldwide, Twitter's active users amount to 330 million each month and 145 million each day (Oberlo.com). Most of these users are between 35-65, with higher-than-average education levels and incomes (Wojcik & Hughes, 2019; Oberlo.com). Research shows a small number of active users (10 percent) tend to do the most "tweeting" (80 percent) (Wojcik & Hughes, 2019).

Twitter has revolutionized the complaint process for consumers. In the past, it was more difficult to complain about a company's services or products. Consumers often did not know where to report, and if a complaint was submitted, it was usually private, i.e., via letter, which did not create a sense of urgency for the company (Istanbulluoglu, 2017, Einweiller & Steilen, 2014). Now, via Twitter, the "fastest media platform," consumers have an "easy and effortless" opportunity to instantly display their complaint to millions of people. (Pfeffer, et. al., 2014, as cited in Istanbulluoglu, 2017, p. 75; Einweiller & Steilen, 2014, p. 197).

Consumers use Twitter to get the company's attention. For example, in this study, researchers observed Tweets to @StayIncHelp such as "I assume you guys watch social media?" (Twitter user, 5/17/20) and "Not sure how many more social outlets I need to reach out on before I get an answer?" (Twitter user, 9/31/20). These tweets show that consumers expect the company to respond and interact with them regarding their complaint. These examples are also consistent with the research literature, which shows that people often use Twitter to ask rhetorical questions (Paul et. al, 2011).

Twitter is also preferable to consumers as a vehicle to complain because they anticipate that the company will respond quickly. Twitter complainants expect a response within 1-3 hours of their Tweet (Istanbulluoglu, 2017). In this sense, Twitter complaints serve as "early warning signals" to the organization of growing problems, which could be ameliorated by taking immediate action to address the situation (Einweiller & Steilen, 2014, p. 196). Research shows that if organizations do not respond appropriately, they can suffer reputational damage (Einweiller & Steilen, 2014).

Research supports consumers' likely use of Twitter to post safety and security-related concerns. Since people use Twitter "extensively" to lodge complaints and otherwise disseminate negative content with the expectation of a swift response (Istanbulluoglu, 2017), it makes sense that their complaints could include reports of crime or fraud.

WHAT IS ARTIFICIAL INTELLIGENCE?

There is no single definition for what constitutes AI, but generally speaking, it is a set of techniques that seeks to approximate human cognition using machines (Calo, 2018; Nilsson, 1980). In other words, it is a branch of computer science that focuses on creating smart machines that are capable of performing tasks that usually require human intelligence. There are many techniques of AI from voice and handwriting recognition to statistical analysis of raw data.

The present research uses a subfield of AI called machine learning. Machine learning is "an application of AI that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use the data to learn for themselves" (Machine learning, 2020). The goal is for a computer to

solve new problems based on past experiences (Alpaydin, 2020). Very generally, the program will work through a dataset, then based on its understanding of the data, it will create an algorithm that it will use to identify otherwise unseen patterns or to predict what is likely to happen in future datasets. Machine learning should improve its performance over time, ultimately recognizing patterns in datasets as new data are added (Calo, 2018, p. 404-405).

Spam detectors in email programs are a commonplace example of machine learning. An email program identifies and segregates potential spam based on its understanding of the content of a typical spam email. The user can refine the email program so that the machine learning is more accurate as more information is input into the program. Ideally, over time, only true spam will be segregated by the program.

The goal in this research is to demonstrate how machine learning can be used by security professionals. Machine learning is used here to review social media data, in the form of tweets from the social media platform Twitter, that addresses users' experiences with a homesharing company, Stay, Inc., and a ridesharing company, Ride, Inc. After given parameters on security risks that may occur in these two settings, the computer will identify the extent to which such risks occur in over 650,000 Twitter conversations . This information will be actionable for security professionals to pinpoint the need for further research and take appropriate steps to mitigate problems.

METHODOLOGY SUMMARY

For this research, the targets of the analysis are referred to as Stay, Inc., a well-known homesharing company, and Ride, Inc., a popular ridesharing company. Tweets made over specified years that referred to the homesharing and the ridesharing companies who were the target of this analysis were purchased from Twitter. A subset of these tweets were human-categorized, or "coded," by who was tweeting, i.e., host vs. guest; driver vs. passenger, and the topic of the tweets. The coded tweets were provided to the computer via an artificial intelligence process, which allowed the computer to examine the entire set of tweets and identify the tweeter and what they were tweeting about. After ensuring the computer was producing accurate results, the researchers were able to use keywords to drill down into the tweets posted by the consumer, the guest in the case of the homesharing company and the passenger in terms of the ridesharing company, to understand the nature of the complaints and to identify security and safety concerns.

UNIT OF ANALYSIS

Here, the unit of analysis is a Twitter conversation. Each conversation may contain several individual posts of Twitter users (tweets), including the original tweet and responses. For simplicity, throughout the document, "tweet" and "conversation" will be used interchangeably.

LIMITATIONS

The results here are specific to Twitter users and are not necessarily generalizable to the entire set of users of Stay, Inc., and Ride, Inc., services. Nevertheless, due to the high number of Twitter conversations about these companies, specifically 169,023 Twitter conversations about Stay, Inc., and 484,871 Twitter conversations about Ride, Inc., over ten years, Twitter is a good source of social media data for the purpose of this analysis.

RESULTS

This research resulted in unique, actionable security intelligence related to homesharing and ridesharing, which could be helpful to security practitioners.

Of all the tweets coded by humans and machine, a full 79.3 percent of those tweets were made by Stay, Inc., guests, as opposed to hosts or other citizens. A closer examination of the tweets labeled as from guests showed that their primary concern (72.2 percent of tweeted conversations) concerned actions taken by Stay, Inc., itself, including poor customer service, technical problems with the website or app, and accounts being locked by the company.

Table 10: Results of AI Analysis for Stay, Inc.

Problem	Total Tweeted Conversations	Percentage of Total
Customer service, tech issue, or account locked	91,811	72.2
Host cancels stay	24,847	19.5
Unsafe conditions or property not as described	7,456	5.9
Account hacking	5,872	4.6
Discrimination	1,114	0.9

Other important findings from guests' tweets included:

- **Homesharing hosts cancel reservations at a high rate, often leaving tweeters in security-compromised positions.** In this study, 19.5 percent of Twitter conversations analyzed reported their stay had been cancelled by the host. These tweets showed that when homesharing stays are cancelled, the guests have often already arrived at their destination and are suddenly placed in an unsafe position without shelter.
- **Unsafe conditions in homesharing properties are an ongoing concern among tweeters.** In this study, 5.9 percent of Twitter conversations analyzed reported either that their property was unsafe or not as described on the homesharing platform. These issues are a concern for security managers because they provide evidence that many consumers are compromising their safety and/or are being defrauded.
- **Hacking via the homesharing platform appears common.** This study found that 4.6 percent of Twitter conversations analyzed reported suspicions that their accounts had been hacked. Typically, guests reported that rooms had been reserved on their credit cards on file with the homesharing company, followed by passwords being changed, which left the guests without recourse.

A keyword analysis of guests' tweets about their experience with Stay, Inc., revealed that many tweeters, specifically 1.7 percent of all conversations, reported experiencing scams, assaults, attacks, break-ins, extortion, and thefts on Stay, Inc., homesharing properties or in relation to Stay, Inc., hosts. While every mention of these keywords didn't necessarily correlate to a crime, the tweets provide security professionals with important information: descriptions of alleged crimes that have occurred in homeshare properties and identification of potential gaps in security policies.

Additionally, several safety issues emerged from the analysis of tweets, which suggest a close examination is needed by security professionals. Examples include:

- **Fire hazards are a concern for Stay, Inc., guests.** A search for the keyword "fire" resulted in 0.3 percent of tweets reporting either concerns about fire hazards in their Stay, Inc., rental or complaining about the refund policies relating to cancelled reservations due to widespread wildfires.

- **Guests have encountered guns on Stay, Inc., properties.** Keyword searches also demonstrated that .04 percent of the tweets analyzed included the words "gun," "shotgun" and "firearm." Although the number of tweets is small, researchers wanted to see the nature of the tweets that used these words. It was found that the use of these words varied, but generally referred to a guest being confronted by a host, the police, or a criminal in the area who was using a gun. Alternatively, several tweets reported that guns were found on the property.

Regarding Ride, Inc., the primary type of tweeter were riders, as opposed to drivers or other citizens. Riders' tweets amount to 75.5 percent of all tweets coded by humans and machine. A closer examination of those tweets labeled as from riders showed:

- **Riders expressed concerns about the cost of the ridesharing service for multiple reasons.** A quarter (25.7 percent) of riders' tweets complained that they had unexplained charges, experienced a drastic price surge, or were charged an unfair rate. Alternatively, the passenger asserted that their driver inappropriately lengthened a trip.
- **Many tweeters did not enjoy their ride.** Nearly 1 out of 10 (8.7 percent) passengers experienced unsafe or unpleasant conditions in their driver's car or during the ride.

A keyword study of riders' tweets also revealed the following:

- **Many users of ridesharing have been hacked via the platform.** This study found that 1.4 percent of Twitter conversations analyzed reported suspicions their accounts had been hacked. Many of these riders reported being hacked on multiple occasions—sometimes in rapid succession. From the reports, it appeared more users were targeted on multiple occasions as opposed to Stay, Inc.
- **The potential of being charged for fraudulent "cleanup fees" are a concern for ridesharing users.** Many Twitter conversations (0.5 percent) described issues with cleaning fees. Passengers were particularly concerned about the assessment of "cleaning fees" as high as $150 by drivers who said they had to "clean up" the "messes" left by said passengers in their vehicles, which many described as bogus. Users said these claims were often difficult if not impossible to reverse, especially if substantiated with fake "evidence" on the part of a driver.
- **Female riders have expressed security concerns.** When analyzing conversations discussing female riders (0.35 percent) here were many instances where they reported feeling unsafe during their ride, due to the actions or statements of their drivers. Some reported being assaulted, either sexually or verbally. This prompted some tweeters to request that the company add the ability for them to select the sex of their driver, which is not currently an option today.

COMMON THEMES

When analyzing the tweets, it was observed that there were common reasons why Twitter users turned to the social media platform to report on their homesharing or ridesharing experiences. Several common themes about tweeters' motivations emerged, which are important for security practitioners looking for meaning and value in this type of analysis.

Common security-relevant reasons for tweeting included:

- **Crime reporting.** Many tweets reported crimes that occurred to the tweeter or were observed by the tweeter. They included: murder, suicide, drug use, domestic violence, sexual harassment, kidnapping, extortion, stolen items, hidden cameras, speeding, and hit and run accidents.
- **Warning others.** Many tweets included statements telling others to "be careful." These tweets described security incidents such as hacking and physical assault.
- **Seeking refund.** Many users requested their money back after experiencing a negative circumstance or experience. These tweets often contained details about that experi-

ence. For example, in one case, a tweeter said their property was burglarized during their stay, yet they were unable to get a refund from the homesharing company.
- **Lack of other recourse.** Commonly, tweeters said they were unable to leave negative feedback about their experiences for a variety of reasons. Thus, these tweets provide unique data not found elsewhere.

Since the tweets for both companies examined here reflected similar motivations, it is likely that most tweeters have similar motivations when tweeting about other organizations. Therefore, security-relevant information should be also identifiable from tweets about most companies.

CONSIDERATIONS FOR PRACTITIONERS

It is clear that this type of analysis should add value to the intelligence collection processes any security practitioner may regularly use on behalf of the company they work for. Using machine learning to analyze social media data can help a security professional gain increased understanding of the risks facing their own organization, other companies in their industry, or another target, such as a frequently used vendor. The primary issue for security professionals will be weighing potential benefits of new intelligence against the cost and the uncertainty of whether actionable results will be uncovered.

BENEFITS.

The benefits of obtaining security-relevant information that is timely, actionable, and potentially unavailable via other means is invaluable. Based on the results achieved by the analyses described in this report, this type of examination adds value to the intelligence collection processes any security practitioner regularly uses on behalf of the company they work for. Using artificial intelligence to assist in this analysis can help a security professional gain increased understanding of the risks facing their own organization, other companies in their industry, or another target, such as a frequently used vendor. Thus, organizations who employ the analysis described in this report will be better positioned to protect their assets.

In addition, examining social media data can be a proactive approach to prevent reputation loss. Reputation loss is difficult to quantify but is a paramount concern for any organization. If customers face serious issues such as crime, unsafe conditions, and monetary loss because of their experience with an organization and tweet about it, this issue is now a public matter. The organization risks losing not only the customer who had the bad experience, but all the people who read the tweet as well. In the end, responding to these critical problems is crucial, but so is reviewing social media data in a meaningful way to protect the organization's reputation.

COSTS.

Costs for an artificial intelligence-supported review of social media data will vary depending on the way this analysis is performed. The work demonstrated here can be conducted either "in-house" or by consultants. A company will have to determine whether it makes the most sense to outsource this project or to hire people to conduct this work on a full-time basis. Factors to consider in this decision would be the industry at issue, the potential value of the information gleaned, and the frequency with which this type of analysis would occur, i.e., whether it is a one-time project or an ongoing effort towards particular security goals. Organizations may elect to conduct the analysis once with a consultant and then decide about future projects based on whether the data yields actionable results.

In any case, the following expenses should be considered:

- Cost of Artificial Intelligence/Machine Learning Tool and Interface
- Cost of Twitter Data
- Personnel Costs (Project Manager, Consultant, Researchers and Data Coders)

Regardless of how it is performed, analyzing social media data using artificial intelligence is an expense that will have to be negotiated into the budget, but based on the present research, the authors think that it will be worth the effort. During challenging times, such as the current pandemic, it can be difficult to make the case to

increase the security budget for new analyses. However, based on the amount expended for the present research, for many companies, the cost involved would not be considered substantial (and certain costs could be absorbed using in-house personnel). Additionally, should critical security intelligence be gleaned from their efforts, such as in the present analysis, the organization will be poised to target key areas where they can help identify and mitigate potential liability. Consequently, security professionals pursuing this type of analysis will be able to justify the cost.

RISK VS. REWARD.

In addition to costs, security professionals should understand that while there are many benefits to this type of analysis, there is also some risk. There is a chance that an analysis may not uncover actionable results. That said, scholarly research has shown that Twitter users often turn to the platform when they have experienced a problem with a company. Therefore, it is quite likely that any given company embarking on an analysis of Twitter or other social media will be fruitful. The results of this report, based on a review of tweets specifically about Stay, Inc., and Ride, Inc., are consistent with those findings. Here, the vast majority of tweets were complaints about the companies and the people earning money via their platform as opposed to compliments. Further, although some security findings developed from these tweets might be well known, several new, concerning issues were also uncovered.

This analysis may also help reduce individual and institutional risk. There were many cases where tweeters named specific people, places, and actions that compromised their security. By looking at issues in the aggregate, a security professional can take an extra measure to help ensure that these risks have, in fact, been reduced. For instance, the security professional could ask: Were specific employees mentioned as offenders in multiple tweets? Were related incidents investigated, and what were the outcomes? Were specific locations mentioned frequently? Do these locations pose a particularly high level of risk? Do some assets seem more at risk than others? Can any policy gaps be identified and amended, thereby reducing risk?

CONCLUSION

Ultimately, this study revealed that social media can provide a wealth of information to security professionals. With the assistance of artificial intelligence, large amounts of social media data can be analyzed efficiently and effectively. Although AI requires technical know-how, which represents costs for AI contractors, this research demonstrates how security issues, including those that had not previously been identified from other sources, can be discovered, to which security professionals can respond.

Appendix I: Premise

Targets of Analysis: Stay, Inc., and Ride, Inc.

Stay, Inc., is a leading provider in the rapidly growing homesharing industry. homesharing refers to "a situation where people share their home or a section of their home in exchange for compensation" (Araujo, 2020). Today, homesharing is most often facilitated by companies, like Stay, Inc., that provide website platforms where people who want to share their home connect with people looking for a place to stay. The homesharing company earns a fee or percentage of each booking. Currently, Stay, Inc., connects guests and hosts nearly worldwide.

Ride, Inc., is a ridesharing company. Ridesharing is "a service that arranges one-time shared rides on very short notice, usually through a mobile app" (GCF, n.d.). Ride, Inc., provides this service, which is appealing for riders looking for a cheaper alternative to traditional car services, such as a taxi. It also provides a way for people to make money, serving as a Ride, Inc., driver.

Scanning for Risks Related to the Targeted Companies

Potential risks related to Stay, Inc. Although quantifiable data is not currently available to measure crime in home shares, anecdotal evidence demonstrates this setting is vulnerable to significant crimes and safety issues. Guests are subject to risks of bodily harm in locations they are unfamiliar with and which lack the general security measures of a hotel (Binns & Kempf, 2020). Further, new crimes have emerged that are uniquely related to home shares (Binns & Kempf, 2020). These crimes include hosts creating fake listings, hosts illegally sharing homes, and guests using home shares to commit illegal acts on another's property. Safety issues are heightened because hosts' adoption of protective measures vary.

A widely publicized incident occurred in Costa Rica where a Florida woman stayed at an Airbnb home share with a family member. The only night she was alone was the last night of her vacation. A week later, her body was found partially buried less than 200 feet from the host's property. It was discovered that she had been brutally murdered, likely after a sexual assault. Ultimately, the security guard for the property, who lived next door, confessed to the crime, which was corroborated by DNA evidence. (Fieldstadt, 2020; O'Connell, 2020; Specia and Mzezewa, 2019; Madan, 2018).

Another example are two violent shooting incidents that occurred at VRBO home share properties used to host parties during the COVID-19 pandemic. In one case, in Manteca, California, someone entered the party with a semi-automatic weapon and began shooting into the crowd. Seven people were injured, including two children. The shooter was not immediately apprehended (Johnson, 2020, Gross, 2020).

In another case, in Tobyhanna Township, Pennsylvania, police were contacted after multiple criminal incidents occurred at a VRBO rental house, including drug possession, a robbery, a shooting, and an attempted homicide (Harrar, 2020). The shooter, who was not identified or apprehended, acted with an accomplice when firing through the front door of the property from the outside. Police arrested 15 people on drug charges in connection with this incident (Harrar, 2020).

Likewise, homesharing hosts are vulnerable to theft and damage perpetrated by guests (Binns & Kempf, 2020.) In addition, guests have pursued illegal activities in home shares, including illegal brothels (Nyheter, 2016, BBC, 2017; Berghuis, 2018) and drug labs (Seth, 2017), among others. There are many reasons why homesharing properties are targets for risky behavior. Primarily, it is because "private homes provide more privacy [than hotels] for these illegal activities to occur" (Binns & Kempf, 2020).

Potential risks related to Ride, Inc. Like homesharing users, ridesharing drivers and passengers are also susceptible to privacy and security risks. There is also a risk of physical harm, including rape, vandalism, and theft. According to a safety report published by Uber, a leading ridesharing provider, 0.0003 percent of rides resulted in a "critical safety incident" (2017-2018). Since there are approximately four million Uber trips every single day, this means there are 1,200 critical safety incidents per day experienced by riders and/or drivers amounting to 438,000 incidents per year by the company's own estimation.

One such incident involving another rideshare company, Lyft, occurred in Lockland, Ohio. The driver was held-up at gunpoint, kidnapped, and robbed (Johnson, 2020). This crime occurred after the driver picked up a customer who entered the passenger seat, pulled out a gun, held it to his leg, and ordered him to drive to a local ATM and take out cash (Johnson, 2020).

In ridesharing, women can be put into compromising positions due to the inherent vulnerability of being in a stranger's car. For example, in 2020, a woman was sexually assaulted by her Uber driver in California. She summoned an early (4 am) ride to work—a time when it is still dark, and streets are empty. Halfway through the trip, her driver "cancelled" the trip, turning off the company's app and telling her the ride was free (Kenton, 2020). He then started asking her questions about her sex life. She was scared and asked him to stop, but he ignored her. He drove "erratically" and would not let her leave and threatened her if she tried to seek help (Likas, 2020). She tried to escape, running out of the car, but he chased after her and allegedly put her into a chokehold, strangled, and sexually assaulted her (Kenton, 2020; Likas, 2020). The driver was arrested after a nearby security officer heard the woman's screams and ran over to assist.

Other documented security issues with ridesharing have threatened its existence. In 2019, Uber lost its ability to operate in London. There, regulators refused to renew Uber's license due to the widespread use of unauthorized drivers operating Uber vehicles (MacDonald & Schechner, 2020; Olson & Needleman, 2019). Drivers apparently were sharing or renting out their own accounts to other drivers who would otherwise be ineligible to operate an Uber for "nefarious" reasons—typically because they do not have a driver's license, would fail a background check, or cannot afford to own a car (Olson & Needleman, 2019).

To be more transparent about the security risks that arise from ridesharing, Uber issued a report that quantified the number of sexual assaults, murders, and fatal accidents that had occurred in 2018 and related to their ridesharing platform (Conger, 2019). Specifically, Uber reported 3,045 sexual assaults, 92 percent of which were riders and the remainder drivers. In addition, there were nine murders and 58 fatal crashes. Even though these cases are a small number compared to the total number of rides booked through Uber's platform, the company has taken several steps to better screen drivers and educate both drivers and riders about sexual assault and other safety matters.

Social media as a source of information about security and safety. Given these widely known issues, homesharing and ridesharing are good subjects for a security analysis. Further, because homesharing and ridesharing companies operate exclusively through the Internet and their proprietary smartphone apps, the people who connect through Stay, Inc., and Ride, Inc., are computer savvy and more likely to be involved with social media. Thus, social media is particularly a good source of information about these types of companies.

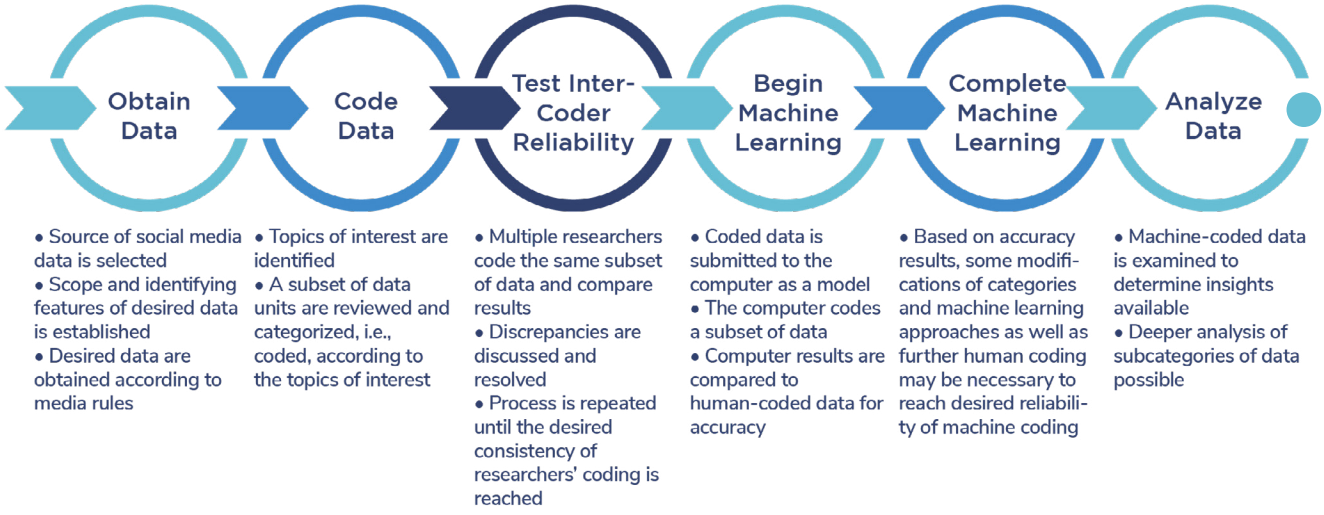
This research sets out to demonstrate how security professionals might use social media and AI to identify security risks by doing the same with Stay, Inc., and Ride, Inc. Tweets about the two companies were reviewed, and with the assistance of machine learning, several security risks that have not received much attention were identified. The methodology used to do so follows.

Appendix II: Methodology

Overview.
This research employed an academic approach to qualitative research, categorizing, or "coding," social media data related to Stay, Inc., and Ride, Inc., according to security themes. Multiple human researchers working in tandem provided assurance about accuracy and reliability of the coding. Then the human-coded data was used as the basis for machine learning, which allowed a computer to analyze the remaining social media data. This final analysis was examined more closely to produce actionable findings. This process is illustrated in Figure 1.

To provide guidance to security professionals, a detailed description of the methodology follows. It should be noted that the same process was used to analyze tweets about Stay, Inc., as was used to analyze the tweets about Ride, Inc.; however, the two analyses were done consecutively. The description follows focuses on the Stay, Inc., analysis, but the reader can assume that the same process was used to analyze Ride, Inc., tweets. Any major differences will be specifically noted in parentheses or separate tables.

Figure 1: Overview of AI Methodology Used to Analyze Social Media Data



Unit of Analysis. As mentioned before, the social media platform Twitter was the source of data for this project. On Twitter, individuals post tweets; however, anyone, even the original poster, can respond to a tweet. That results in a conversation of multiple tweets. The unit of analysis used in this research was the conversation. For simplicity, throughout the document, "tweet" and "conversation" will be used interchangeably.

Obtaining the data. Twitter makes its tweets available for bulk download via a paid subscription to their developer API. The prices vary based on the subscription level but start at $99 for up to 50,000 tweets. The price per tweet gets cheaper as you pay for higher subscription levels. For example, for this research $774 was expended for downloading up to 500,000 tweets related to Stay, Inc. (An additional amount of $2,025 was spent to download conversations related to Ride, Inc.)

The downloaded conversations included one or more of the following terms: "Stay, Inc.," "@StayInc," and "@StayIncHelp." (Also, conversations that mentioned "Ride, Inc.," "@RideInc," and "@RideIncHelp" were obtained.) For ease in interpretation, any tweets not written in English were excluded. The Twitter page analyzed for each entity was designed to receive requests for "help." This was a purposeful sampling method, to target and obtain the most potentially security-relevant information.

For Stay, Inc., we were able to download 365,250 individual tweets which gave us a total of 169,023 conversations dating from 1 January 2013, through to 2 September 2020. (For Ride, Inc., we were able to download more than 1 million individual tweets which gave us a total of 484,871 conversations dating from 14 October 2010 through to 11 October 2020.)

Coding and Inter-Coder Reliability. Coding is a way to systematically categorize data. In this case, Twitter conversations were coded in two stages. First, it was determined whether the initial tweet in a conversation was sent by a Stay, Inc., guest, a host, or a neighbor/citizen. There was also a fourth category, "unknown," which was used when the coder was not able to identify the person tweeting. (Similarly, for the Ride, Inc., analysis, the conversations were initially coded as from a driver, a passenger, a citizen, or unknown.) This categorization was done to be able to identify the unique vulnerabilities of the primary parties tweeting, i.e., hosts vs. guests.

The second step was to identify the nature of the comment. Categories of types of security issues were developed through brainstorming based on a review of prominent media coverage related to problems about Stay, Inc., researchers' personal experiences, and the types of situations that were commonly found when reviewing a preliminary set of 300 random conversations. A sample of the initial terms used to code Stay, Inc., conversations made by Stay, Inc., guests are shown in Table 1. (Initial terms used to code Ride, Inc., passenger conversations are found in Table 2.)

To ensure inter-coder reliability, i.e., that coders were analyzing conversations similarly, two coders analyzed the same set of data separately without knowing how the other coder was categorizing each conversation. This coding was performed on a webapp platform designed by the machine learning company Sciling, the AI contractor for this project. Each coder initially categorized a randomized sample of 300 conversations or about 0.18 percent of the total 169,023 Twitter conversations about Stay, Inc., For Ride, Inc., 300 conversations were coded, making up 0.06 percent of the total conversations. Coders identified the type of tweeter and then assigned up to three categories of types of complaints to each conversation.

Table 1: Initial Sets of Terms Used to Code Stay, Inc. Conversations

Nature of Complaint	Definition
Account hacking	A specific type of scam experienced by a Stay, Inc., guest on the Stay, Inc., platform. Account hacking includes but is not limited to the unauthorized use of a guest's Stay, Inc., account and/or payment methods.
Account locked	Passengers can't get access to their account.
Customer service issue	Any and all customer service complaints raised for Stay, Inc.'s attention, including but not limited to inadequate support from Stay, Inc., in emergency situations, unfulfilled promises from Stay, Inc., or unfair policies. General suggestions or recommendations for improvement of the Stay, Inc., experience are included.
Discrimination	Any form of discrimination or ill-treatment experienced by guests in their interaction with hosts, including racial, LGBTQ, or disability discrimination.
Fake listings/reviews	A specific type of scam experienced by Stay, Inc., guests because of fake listings or untrue reviews. This code includes blackmail attempted for good reviews or reviews modified and/or censored by Stay, Inc.
Host cancels stay/does not show up	Any situation created by the host's action or inaction, which prevents a guest from beginning or from completing the stay confirmed at booking.
Not as described	When the accommodations do not meet the guest's expectations. These include but are not limited to the following: the location of property differs from the map, the property looks better in photos, and advertised amenities are missing.
Scam	Situation experienced or observed by a Stay, Inc., guest that seems to be a part of a fraudulent scheme, such as fake hosts, falsified damages, demands for cash, or unfair reviews of a guest.
Tech issue	Any and all technical problems encountered with Stay, Inc.'s website or app, by a guest. Also includes suggestions made specifically to improve technical aspects of the website or app.
Unsafe conditions	Any situation at the property that presents physical safety or health concerns for guests, including but not limited to bug infestations, unsafe/broken amenities, hostile or intimidating hosts, and hidden cameras in a bedroom or bathroom
Other	A label for any other type of situation that may be complained about, which is not adequately covered by any of the other labels; an "if all else fails" label.

Table 2: Initial Sets of Terms Used to Code Ride, Inc., Conversations

Nature of Complaint	Definition
Account hacking	A specific type of scam experienced by a Ride, Inc., guest on the Ride, Inc., platform. Account hacking includes but is not limited to Account hacking includes but is not limited to the unauthorized use of a guest's Stay, Inc., account and/or payment methods.
Account locked	A passenger can't get access to their account.
Customer service issue	Complaining specifically about Ride, Inc., passengers state customer support was not being helpful or unreachable.
Discrimination	Any form of discrimination or ill-treatment experienced by passengers in their interaction with drivers, including but not limited to racial, LGBTQ, or disability discrimination. Can include being refused rides.
Drastic price surge or unfair rate	When a passenger experiences a drastic surge price increase or what they deem is an unfair charge by Ride, Inc., e.g., prices raised due to demand or increased drastically to drive to an airport.
Driver lengthens trip	Situation where a driver takes wrong turns, drives the wrong way, extends trip length and time.
Driver or vehicle identity issues	Situation where the driver does not match the picture provided in the app, or the vehicle does not match license plate or description. Alternatively, there is failure in the driver's license status or their background check.
Lost or stolen item	The driver is accused of stealing an item from a passenger or the passenger has left an item inside the vehicle.
Lost time	When a driver cancels the ride at the last minute, is not at the pick-up location, or similar situation which leads the passenger to spend excess time in travel.
Positive feedback	When a passenger shares positive information about their experience with Ride, Inc.
Scam	Scams occur when a passenger is wrongly charged for anything caused by the driver. This includes but is not limited to the driver asking for cash, driver asking to bypass the app, the driver fakes an app glitch, the passenger is charged for unrequested upgrades, and the trip doesn't end when you get out of the car.
Tech issues	Any and all technical problems encountered with Ride, Inc.'s website or app, by a passenger, where the app doesn't work properly, including the app miscommunicating with the driver and the location services malfunctioning. This category includes criticism of features or functionality of the app.
Unexplained charges	Any instance where the passenger is wrongly charged and the charge cannot be explained, including but not limited to situations where the passenger is charged, then uncharged, and then charged again.
Unfair reviews	When a driver gives a passenger a review, they believe to be unfair or undeserved.
Unpleasant conditions	Any situation in connection with the rideshare that is not unsafe but is not pleasant, e.g., a bad, nontoxic odor, loud music, a driver talking too much, and unclean but not unsanitary comments.
Unsafe Conditions	Any situation in connection with the rideshare that presents physical safety or health concerns for passengers, including but not limited to a driver not wearing a mask, a driver being physically or verbally aggressive, sexual harassment, or a driver being drunk, speeding, or on their phone, aggressive driving.
Other	A label for any other type of situation that may be complained about, which is not adequately covered by any of the other labels; an "if all else fails" label.

Table 3: Manually Coded Stay, Inc., Conversations by Type of User

Type of User	Number of Conversations	Percentage
Guest	2,404	66.5
Unknown (possible guest or host)	820	22.7
Host	337	9.3
Neighbor or Citizen	55	1.5
Total	3,616	100

Table 4: Manually Coded Ride, Inc., Conversations by Type of User

Type of User	Number of Conversations	Percentage
Passenger	933	42.4
Unknown (possible passenger or driver)	755	34.3
Driver	460	20.9
Citizen	52	0.02
Total	2,200	100

The platform allowed for a comparison of the coders' work to identify the extent to which the coders agreed. Because intercoder reliability is essential for this process, the coders and researchers closely examined the conversations that the coders did not agree on. As a group, researchers and coders came to a consensus as to the accurate coding and amended the coding definitions as needed. With new clarity, the coders coded another set of 300-500 conversations. This process was repeated until inter-coder reliability was at least 75 percent. Finally, the researchers reviewed the remaining conversations that the coders had disagreed on to resolve the conflict and achieve intercoder reliability of 100 percent. In total, 3,616 conversations, or 2.14 percent, had been coded manually. For Ride, Inc.: In total, 2,200 conversations, or 0.45 percent were coded manually. The types of tweeters for these conversations are detailed in Table 3. This information shows that guests are more likely to turn to Twitter to discuss Stay, Inc., than other categories of tweeters. (Likewise, Table 4 shows the most common tweeter who is tweeting about Ride, Inc., is the passenger, making up 75.5 percent of conversations that were human-coded.)

Machine learning. A majority of the human-coded conversations were introduced to the computer with the intent that it would learn patterns from the coded conversations and categorize the remaining conversations. These computer-coded conversations were then compared to the human-coded conversations to calculate a level of accuracy. Sciling reported that their staff used multiple models to train the machine to recognize patterns in the human coding, including addressing the data imbalance between tweets from guests versus the other categories. Table 5 shows the AI results achieved based on the number of human-coded examples of Stay, Inc., conversations, sorted by who tweeted. (Table 6 shows the results achieved for Ride, Inc., conversations.)

To understand Tables 5 and 6, the following definitions are important.

- **Recall** "quantifies the number of positive class predictions made out of all positive examples in the dataset" (Brownlee, 2020). Table 4 illustrates that from all guest conversations, the machine was able to correctly classify 87 percent.
- **Precision** "quantifies the number of positive class predictions that actually belong to the positive class" (Brownlee, 2020). As Table 4 shows, from all the conversations that the machine labeled as "guest," 90 percent were correct.
- **F1 Score (or F-measure)** is "a measure of the test's accuracy and is defined as the weighted harmonic mean of the

Table 5: Accuracy of Machine Classifications of Type of Tweeter for Stay, Inc., Conversations

Type of Tweeter (# of human- coded examples)	Precision	Recall	F1 Score
Guest (4)	0.87	0.90	0.88
Unknown (820)	0.77	0.67	0.72
Host (337)	0.83	0.28	0.41
Neighbor (55)	1.00	0.05	0.10

Table 6: Accuracy of Machine Classifications of Type of Tweeter for Ride, Inc., Conversations

Type of Tweeter (# of human-coded examples)	Precision	Recall	F1 Score
Passenger (933)	0.83	0.83	0.83
Unknown (755)	0.76	0.82	0.79
Driver (460)	0.88	0.69	0.78
Citizen (52)	0.62	0.19	0.29

precision and recall of the test" (Metacademy, n.d.). The F1 score is calculated by dividing the number of observations by the reciprocal of each number in the series, i.e., the precision and recall of the test.

Table 5 shows that the computer was most accurate when coding the guests' tweets, which is likely due to the larger number of tweets to analyze proportionally as compared to the other types of tweeters. Based on the results shown in Table 5, it was clear that there were not enough coded tweets in categories other than "guest" to train the machine. Due to this, along with the fact that guests are the primary users of Twitter to communicate about Stay, Inc., the researchers decided not to do additional coding for "hosts," "citizens," and "other," but rather to focus on the security and safety issues that guests experience in home share settings. (Similarly, Table 6 shows that the machine coding for passengers was most accurate.)

The next step Sciling staff took was to train the machine, in multiple iterations, to recognize both who tweeted (specifically, to identify guests) and the type of complaint they offered. With these initial results, the researchers found the F1 score did not reach 80 percent for most of the categories of complaint type, thereby, making the machine's categorization unreliable, which required additional steps, including merging similar categories and reviewing a sample of machine coding.

The final results of the accuracy of the machine's coding of types of Stay, Inc., guests' complaints are provided in Table 7. (The accuracy of the machine's coding of types of Ride, Inc., complaints are provided in Table 8.) Categories reaching acceptable levels of accuracy are bolded. Table 7 shows that two categories of guest complaints, account hacking and discrimination, met the 80 percent accuracy threshold that the researchers were striving for. Account hacking received an F1 score of 0.86, and discrimination had an F1 score of 0.81, which means that the accuracy of the computer's categorization will be higher than 80 percent.

In addition, three other categories had 72 percent accuracy or higher. These categories included the combined category of customer service, locked account and tech issue (F1 score = 0.72), the combined category of host cancels stay/host does not show up (F1 score =0.78). As increased accuracy was unlikely, the researchers determined these results with more than 70 percent accuracy, would be sufficient for analysis. (For Ride, Inc., Table 8 shows the two categories most directly related to safety and security concerns, account hacked and unsafe or unpleasant conditions, had an F1 score higher than the 80% threshold, 0.86 and 0.82 respectively. Likewise, the category related to scams, including drivers lengthening trips, unexplained charges, drastic price surges, or unfair rates, had an F1 score higher than the 70 percent threshold, at 0.72 percent).

The findings of the computer's coding based on these levels of accuracy are described below in the findings section.

Keyword searches. Using the machine learning results, the same platform created by Sciling was used to do keyword searches within the machine-coded categories, or in the data set as a whole, to identify more granular information about security risks. Specifically, the AI platform enabled the database of 169,023 Twitter conversations about Stay, Inc. (484,871 about Ride, Inc.), to be searched for single words or combinations of words. Then the conversations containing the keywords were analyzed in depth to gain a better understanding of the context in which they were used by the tweeter.

Table 7: Accuracy of Machine Classifications of Stay, Inc., Guests' Types of Complaints

Type of Complaint (# of Human-Coded Examples)	Precision	Recall	F1 Score
Customer Service/ Locked Account/ Tech Issue (2117)	0.61	0.89	0.72
Unsafe Conditions/ Not as Described (941)	0.72	0.79	0.75
Host Cancels Stay/ Host Does Not Show Up (829)	0.74	0.82	0.78
Account Hacking (522)	0.85	0.87	0.86
Other (427)	0.48	0.41	0.45
Discrimination (259)	0.88	0.76	0.81
Scam/ Fake Listing or Review (182)	0.50	0.01	0.01

Table 8: Accuracy of Machine Classifications of Ride, Inc., Passengers' Types of Complaints

Type of Complaint (# of Human-Coded Examples)	Precision	Recall	F1 Score
Customer Service Issues or Tech Issues or Account Locked	0.59	0.68	0.63
Unsafe or Unpleasant Conditions (1014)	0.83	0.82	0.82
Scam or Driver Lengthens Trip or Unexplained Charges or Drastic Price Surge or Unfair Rate (902)	0.79	0.76	0.78
Account Hacked (531)	0.84	0.90	0.87
Discrimination (334)	0.62	0.41	0.50
Other or Positive Feedback (268)	0.64	0.24	0.35
Lost or Stolen Item (79)	0.89	0.30	0.45

The methodology for conducting the keyword research was as follows. An initial list of keywords was curated, designed to retrieve issues impacting safety and security. Some keywords were gleaned from the security literature and others had been observed in the tweets reviewed by coders and researchers. For instance, two keywords searched for were "break-in" and "theft." Searches were conducted using the root words to return all possible results. For instance, the work "hack" to return results including the words "hacked" and "hacking."

If any keyword retrieved a small number of conversations, defined as 275 or less, the conversations were analyzed qualitatively to determine if the conversation identified a safety or security issue. For instance, a conversation including the word "assault" might describe a physical assault that occurred at a home share property, which is an obvious security issue. On the other hand, another conversation might complain about an air freshener being an "assault on the senses," which is not a security problem.

If there were more than 275 conversations, they were further refined using advanced keyword searches with a combination of keywords. The advanced keyword search process illustrated in Table 9 using the words "fake" "price" and "listing" as an example. From the machine learning results and the key word searches, findings were developed.

Limitations: The study population was limited to people who have self-reported issues to Stay, Inc., and Ride, Inc., via Twitter.

Table 9: Example of Advanced Keyword Searches

Keyword Combination	The Type of Conversations that the Keyword Combination Will Retrieve
Fake Price Listing	Conversations that contain words "fake" or "price" or "listing"
"Fake" "Price"	Conversations that include words "fake" and "price" (both must be included)
"Fake Price"	Conversations that include the literal string "fake price" (i.e., both words, with a space between them)
"Fake Price" "Listing"	Conversations that include the string "fake price" and the word "listing"

This fact had several implications to the analysis. First, it excluded any people who are not on Twitter. Thus, findings could not be extrapolated to the entire population of people who use Stay, Inc., and Ride, Inc., Second, this population was skewed towards people having complaints as opposed to providing compliments. Tweets to @StayIncHelp or @RideIncHelp specifically were found to be more likely negative, because the tweeters were looking for assistance with a problem. Finally, Twitter does not validate the content of tweets, and therefore, some of the conversations might be untrue or fake. Nevertheless, these limitations did not negate the value of the findings in terms of identifying actionable safety and security issues for a security professional to address. In fact, some of these limitations may have resulted in a population of data that was more narrowly targeted toward the objective of this study.

Appendix III: Works Cited

Alpaydin, E. (2020). *Introduction to machine learning.* Massachusetts Institute of Technology.

Araujo, M. (2020). How to get affordable Airbnb or homesharing host insurance. Retrieved from https://www.thebalance.com/finding-airbnb-homesharing-insurance-4140684 on September 7, 2020.

Binns, C. A., and Kempf, R. J. (2020). *Safety and security in hotels and home shares.* Springer.

Boyd, D. M., & Ellison, N. B. (2008). Social network sites: Definition, history, and scholarship. Journal of Computer-Mediated Communication, 13, 210-230

Brooks, D. J. (2010). What is security: Definition through knowledge categorization. *Security Journal,* 23(3), 225-239.

Brownlee, J. (2020). How to calculate precision, recall, and F-measure for imbalanced classification. *Machine Learning Mastery,* August 1. Retrieved from https://machinelearningmastery.com/precision-recall-and-f-measure-for-imbalanced-classification/ on August 29, 2020.

Conger, K. (2019). Uber says 3,045 sexual assaults were reported in U.S. rides last year. *New York Times,* December 5. Retrieved from https://www.nytimes.com/2019/12/05/technology/uber-sexual-assaults-murders-deaths-safety.html?auth=login-email&login=email on September 24, 2020.

Dollarhide, M. E. (2020). Social media definition. Investopedia.com. Retrieved from https://www.investopedia.com/terms/s/social-media.asp on September 9, 2020.

Einwiller, S. A., & Steilen, S. (2015). Handling complaints on social network sites–An analysis of complaints and complaint responses on Facebook and Twitter pages of large US companies. Public Relations Review, 41(2), 195-204.

Expert System Team. (2020). What is machine learning? A definition. Expert System. May 29. Retrieved from https://expertsystem.com/machine-learning-definition/ on August 29, 2020.

Fawcett, T., Haimowitz, I., Provost, F., & Stolfo, S. (1998). AI approaches to fraud detection and risk management. AI Magazine, 19(2), 107.

Fennelly, L. (2016). Effective Physical Security, 5th Ed. Butterworth-Heinemann.

Fugate, S., & Ferguson-Walter, K. (2019, Spring). Artificial intelligence and game theory models for defending critical networks with cyber deception. AI Magazine, 40(1), 49.

GCF Global. (n.d.). Sharing economy: What is ridesharing? Retrieved from https://edu.gcfglobal.org/en/sharingeconomy/what-is-ridesharing/1/ on September 7, 2020.

Gill, A.S. (2019). Artificial intelligence and international security: The long view. Roundtable: AI and the Future of Global Affairs, 169-179.

Gross, K. (2020, July 25). "Homeowner speaks out about shooting that left 5 adults, 2 kids injured during house party." Fox40.com. Retrieved from: https://fox40.com/news/local-news/homeowner-speaks-out-about-shooting-that-left-5-adults-2-kids-injured-during-house-party/ on February 4, 2021.

Harrar, H. (2020, Sept. 18). "Police: attempted homicide investigation underway after man fires shots during party at home rental." 69 News. Retrieved from: https://www.wfmz.com/news/area/poconos-coal/police-attempted-homicide-investigation-underway-after-man-fires-shots-during-party-at-home-rental/article_38cf4604-f9e5-11ea-9613-33601398aed9.html on February 4, 2021.

Home. (n.d.). What is twitter and why should you use it? https://esrc.ukri.org/research/impact-toolkit/social-media/twitter/what-is-twitter/ on August 10, 2020.

Istanbulluoglu, D. (2017). Complaint handling on social media: The impact of multiple response times on consumer satisfaction. Computers in Human Behavior, 74, 72-82.

Johnson, B. (2019, Spring). Artificial intelligence--An enabler of naval tactical decision superiority. AI Magazine, 40(1), 63.

Johnson, J. (2020, July 25). ABC10.com. "7 shot at party in Manteca, officials say." Retrieved from: https://www.abc10.com/article/news/crime/party-in-manteca-ends-in-shooting/103-ed41b079-733d-42bc-9f5b-121494ad14d5 on February 4, 2021.

Johnson, K. (2020, Dec. 2). "Lyft driver says he was held-up at gunpoint, kidnapped, robbed." WLWT. Retrieved from: https://www.wlwt.com/article/lyft-driver-says-he-was-held-up-at-gunpoint-kidnapped-robbed/34852355 on February 4, 2021.

Kenton, L. (2020). Uber driver, 30, is arrested for 'choking and sexually assaulting' a 51-year-old passenger he 'chased down after she tried to get out of his car because he kept asking about her sex life during 4am ride to work.' Daily Mail. Retrieved from https://www.dailymail.co.uk/news/article-8718417/Uber-driver-30-arrested-choking-sexually-assaulting-51-year-old-passenger.html on October 7, 2020.

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. Business Horizons, 54(3), 241-251.

Knowledge @Wharton (2014, January 09). The ignored side of social media: Customer service. Forbes. Retrieved from https://www.forbes.com/sites/knowledgewharton/2014/01/09/22014/ on August 28, 2020.

Kwak, H., Lee, C., Park, H., & Moon, S. (2010, April). What is Twitter, a social network or a news media? In Proceedings of the 19th international conference on World wide web (pp. 591-600).

Likas, E. (2020). California Uber driver suspected of kidnapping, sexually assaulting passenger. The Mercury News. Retrieved from https://www.mercurynews.com/2020/09/10/uber-driver-suspected-of-sexually-assaulting-passenger-in-santa-ana-2/ on October 7, 2020.

Lin, Y. (2020, July 14). 10 Twitter statistics every marketer should know in 2020 [Infographic]. Retrieved from https://www.oberlo.com/blog/twitter-statistics on August 10, 2020.

Liu, J., Bier, E., Wilson, A., Guerra-Gomez, J. A., Honda, T., Sricharan, K., Gilpin, L., & Davies, D. (2016, Summer). Graph analysis for detecting fraud, waste, and abuse in health-care data. AI Magazine, 37(2), 33.

MacDonald, A. & Schechner, S. (2020). Uber wins back license to operate in London after years long battle. The Wall Street Journal. Retrieved from https://www.wsj.com/articles/uber-wins-back-license-to-operate-in-london-after-yearslong-battle-11601286874 on October 7, 2020.

Metacademy. (n.d.). F measure. Retrieved from https://metacademy.org/graphs/concepts/f_measure on August 29, 2020.

Morris, M., Teevan, J., and Panovich, K. (2010). A comparison of information seeking using search engines and social networks, Proc. ICWSM, 291294.

Nilsson, N. J. (1980). Principles of artificial intelligence. Morgan Kaufmann Publishers, Inc.

Olson, P. & Needleman, S. (2019). Uber's 'dirty little secret:' shared driver accounts. The Wall Street Journal. Retrieved from https://www.wsj.com/articles/ubers-dirty-little-secret-shared-driver-accounts-11574883278 on October 7, 2020.

Paul, S. A., Hong, L., & Chi, E. H. (2011, July). Is twitter a good place for asking questions? a characterization study. In Fifth International AAAI Conference on Weblogs and Social Media.

Pita, J., Jain, M., Ordonez, F., Portway, C., Tambe, M., Western, C., Paruchuri, P., & Kraus, S. (2009, Spring). Using game theory for Los Angeles airport security. AI Magazine, 30(1), 43.

Radulov, N. (2019). Artificial intelligence and security. Security 4.0. International Scientific Journal "Security & Future", (1), 3-5.

Senator, T. (1995). The financial crimes enforcement network AI system (FAIS): Identifying potential money laundering from reports of large cash transactions. *AI Magazine,* 16 (4), 21.

Suhl, K. (2020, July 14). Top Twitter demographics that matter to social media marketers. Retrieved from https://blog.hootsuite.com/twitter-demographics/ on August 28, 2020.

Wasilow, S., & Thorpe, J. B. (2019, Spring). Artificial intelligence, robotics, ethics, and the military: A Canadian Perspective. *AI Magazine,* 40(1), 37.

West, T. (2019). Uber delivers U.S. safety report. *Uber Newsroom.* Retrieved from https://www.uber.com/newsroom/2019-us-safety-report/ on October 7, 2020.

Uber (2019) "Safety Report, 2017-2018." Retrieved from: https://www.uber-assets.com/image/upload/v1575580686/Documents/Safety/UberUSSafetyReport_201718_FullReport.pdf?uclick_id=1effe8e4-a71f-4de5-a23c-24efa15efbc6 on October 7, 2020.

Zhao, Y., & Flenner, A. (2019). Deep models, machine learning, and artificial intelligence applications in national and international security--Part two. *AI Magazine,* 40(2), 29.

About the ASIS Foundation

The ASIS Foundation, a 501(c)(3) nonprofit affiliate of ASIS International, supports global security professionals worldwide through research and education. The Foundation commissions actionable research to advance the security profession and awards scholarships to help chapters and individuals—including those transitioning to careers in security management—achieve their professional and academic goals. Governed by a Board of Trustees, the Foundation is supported by generous donations from individuals, organizations, and ASIS chapters and councils worldwide.

Support future security research with a gift to the ASIS Foundation. Online at *www.asisfoundation.org*.