Using Facebook as a Pre-Employment Screen: A Case Study of Text Analytics

Bhavik Modi

Appalachian State University

Abstract

Companies can be liable for negligent hiring if due diligence is not exercised. Hiring the wrong person poses a risk to others and the organization and can result in lawsuits against the company. Criminal background checks are commonly used to exercise due diligence during the hiring process and to help avoid claims of negligent hiring. However, these background checks are not an all-encompassing method of screening. This study proposes information posted on Facebook as a complementary pre-employment screening tool and looks to determine if this information can be used to predict an alcohol-related conviction. Additionally, this study looks to determine if there are differences in Facebook posts between people convicted of crimes and a randomly generated sample of people. SAS Text Miner, a text analytics program, was used to answer these research questions. Results showed that people convicted of alcohol-related crimes were found to post different content than participants in the comparison group. A logistic regression model showed that participants in the comparison group were more likely to post slang terms than participants in the conviction group.

Using Facebook as a Pre-Employment Screen: A Case Study of Text Analytics

During the employee selection process, organizations have to be wary of negligent hiring. Attorney Lester Rosen (2017) writes that employers can be sued for negligent hiring if, during the selection process, they did not properly assess whether a new employee could pose a risk to others while on the job, a process called due diligence. In other words, prior to hiring, the employer should be aware if any applicant is unsuitable for the job and could cause harm to others while performing the job. Failing to exercise due diligence can result in organizations being held liable for their employees' actions. Furthermore, employers lose about 75% of all negligent hiring lawsuits, making this an important issue for organizations (Valdez, 2015).

Negligent hiring

While the exact definition of what constitutes negligent hiring varies by state, negligent hiring is generally considered when an employer is held liable for the harm resulting from its employee's negligent acts and failed exercised reasonable care in choosing or retaining an employee for the particular duties to be performed (Kittling, 2010). For instance, if a trucking company were to hire a driver with a DUI on their record, the company could be held liable for negligent hiring if the driver were to get in an accident while driving on company time. Since the employer should have been legally aware of the risks the applicant posed, the court could conclude that the employer did not do their due diligence. In Reagan et al v. Dunaway Timber Company et al (2011), a trucking company was found guilty of negligently hiring a driver with a history of unsafe driving including having his license revoked twice. This driver got into an accident on company time and therefore, the company was held responsible for the driver's actions.

Another case law example of negligent hiring occurred at the Swedish Medical Center (SMC) in Colorado. In Porras, Pecha, and Wolter v. Swedish Medical Center (2016), a class action lawsuit, the SMC was held liable for hiring a surgical technologist, Rocky Allen, who had a prior history of poor job performance and misconduct. At SMC, he exposed patients to an increased risk of contracting bloodborne pathogens, such as HIV and Hepatitis B, by using syringes to inject himself with drugs and then using those same syringes in surgical procedures. Previously, Allen had been court-martialed from the US Navy for abusing drugs and had a history of drug addiction. Despite this, he was hired by SMC and given a job where he was allowed access to narcotics. The court found SMC liable for negligently hiring Allen and placing thousands of patients at risk. These are just two case law examples of where an employer was found to have failed to exercise reasonable care in choosing an employee for the job tasks and responsibilities.

Criminal background checks

A common way that organizations often exercise due diligence during the hiring process is a criminal background check (CBC). A CBC is the review of records containing any information collected and stored in the criminal record repository of criminal history records, involving a pending arrest or conviction by a criminal justice agency ("Criminal history background," n.d.). Employers use CBCs to gather information on an applicant's criminal history and are conducted either by applicant self-report or through an external vendor. Generally, CBCs reveal felony and misdemeanor convictions from county criminal records. However, this limits the scope of the background check to crimes that occurred in that county. For some jobs, more extensive background checks are conducted, such as statewide or nationwide checks. Employers use information from the criminal history to make a decision on whether or not a particular criminal history suggests the applicant would not be a good employee and/or could pose a risk to others in the workplace (Aamodt, 2015).

A survey conducted by the Society for Human Resource Management (SHRM) found that 86% of employers conduct CBCs for some of their employees, 69% of employers conduct CBCs on all their applicants, and only 14% of employers do not conduct CBCs on any of their applicants (SHRM, 2012). In many cases, applicants with a criminal history are allowed to explain the circumstances behind their criminal record, but this process allows employers to claim they did their due diligence when hiring, make more informed decisions regarding which individuals to employ, and determine what job responsibilities these employees should be assigned if hired.

Despite the wide spread use of CBCs, there are some shortcomings. CBCs only identify people who have committed crimes in the past and fail to reveal non-criminal behaviors that still pose or could pose a risk to others or the company. This shortcoming is highlighted by the fact that companies have been held liable for negligent hiring when the employee had no prior criminal history but still exhibited hazardous behaviors. For example, in Bassam Jafar v. Elrac, Inc. (2011), the defendant Eugene Baum, an employee of Elrac, Inc., was driving under the influence, four times over the legal limit and killed Jafar’s two children while driving for Elrac. Though he did not have any criminal history, Elrac, Inc. was sued for negligent hiring because of the defendant’s past behaviors. Elrac, Inc. lost the case due to the fact that Baum had a history of drinking and although he had never been arrested or charged with an alcohol-related offense, Elrac, Inc. was found to not have exercise reasonable care in hiring Baum. To the court, Baum’s past behaviors related to alcoholism were deemed both foreseeable and that they posed a risk to others while on the job, since

Baum's job as a traveling salesperson required him to drive in order to perform his regular duties.

Outside of this case law example, systematic research examining the predictive validity of CBCs reveal that criminal histories are only a minimally valid method of predicting employment outcomes. According to a meta-analysis by Aamodt (2015), criminal history has a validity coefficient of only 0.07 in the prediction of counterproductive workplace behaviors (CWBs). In comparison, certain personality traits have much higher validity coefficients when related to CWBs, such as narcissism (r = 0.35), Machiavellianism (r = 0.20), and conscientiousness (r = -.44; Mount, Ilies, & Johnson, 2006, O'Boyle et al., 2012). Based on these study findings and the results of negligent hiring and retention court cases, CBCs ought not to be used as the sole tool in the pre-employment screening process and other potential methods of due diligence should be explored as a way to avoid negligent hiring.

Social networking sites

One such method that has been suggested as a possible screening tool for learning more about a job candidate during the application process is information posted on social networking sites (SNSs). In particular, Facebook is appealing to companies because it is thought to give employers a unique glimpse into the personal lives of applicants; employers can gauge an applicant's true personality, behaviors, and interests, which can help determine person-organization (P-O) fit (Roulin & Bangerter, 2013). In fact, many employers use Facebook in their selection process to specifically assess characteristics potentially predictive of P-O fit. Employers also use Facebook to view negative behaviors of applicants that may not show up on CBCs (Jobvite, 2015).

In 2015, about 40% of organizations said they used social media as a screening tool, whereas that number was only 33% in 2013 (SHRM, 2016). Results from another study showed that 55% of employers have reconsidered an applicant's position after looking at their behaviors on their SNS profile (Jobvite, 2014). Furthermore, 50% of hiring managers rejected applicants because of Facebook posts about drinking or using drugs, and a third of hiring managers rejected applicants because they saw badmouthing and discriminatory remarks on their profiles (Willey et al., 2012).

Currently, there is little empirical evidence that shows if information posted on SNSs can or should be used as a predictor of job performance, CWBs, on-the-job crimes, or other job-related outcomes. That is, there is little empirical support for the predictive validity the use of SNS information in employee selection. One of the purposes of this study is to determine if information on social media can predict criminality and, thus, could be used as a complement to CBCs. If such a link could be established, then conducting a "social media" background check could be part of the due diligence process and employers would be less likely to be held liable if they checked an applicant's social media profiles.

SNSs also has the opportunity to provide different information than a CBC. As previously mentioned, CBCs only identify those who have committed crimes in the past. Facebook could allow employers to identify non-criminal behaviors that are precursors to criminal activities and are not detected by CBCs. Compared to CBCs, certain behaviors on Facebook could have stronger predictive validity for job-related outcomes such as on-the-job crimes or CWBs. The current study examined if content posted on Facebook could predict a future alcohol-related criminal conviction. If a link is found then by using Facebook, companies could be able to prevent hiring someone that has the potential to commit a crime and avoid negligent hiring lawsuits by including both a CBC and social media background check as part of their due diligence.

While not yet demonstrated in the research literature, past court cases looking at negligent retention suggest that companies could be held liable for hiring someone and not checking their public Facebook posts prior to hiring. For example, Hertz was sued by a customer when an employee posted a derogatory status about the customer on Facebook (Howard v. Hertz, 2016). The court ruled that since the employee had a history of posting offensive Facebook statuses, it was foreseeable to the company and therefore, Hertz should have expected that this problem would arise. This ruling implies that companies could also be held liable for negligently hiring someone based on their social media presence. Thus, it is important for employers to be aware of how their employees are using SNSs. Hiring someone who posts maliciously online could come back to hurt the company and result in a negligent hiring lawsuit.

Workplace behaviors

Vehicular accidents. According to the National Highway Traffic Safety Administration (NHTSA), an average vehicular accident that occurs on the job costs the employer $24,000. If a worker is injured in the crash, the cost increases to $65,000, and if a fatality is involved, those cost skyrockets to above $670,000. Average costs such as these result in annual total costs to employers for on-the-job crashes of approximately $25 billion (Blincoe et al., 2015). The industry most effected by on-the-job car crashes is the transportation industry with a cost of on-the-job vehicle crashes of around $5.5 billion with more than half of these accidents and costs coming from the trucking and warehousing sectors (Network of Employers for Traffic Safety, 2016). As these figures demonstrate, on-the-job vehicular accidents are a large financial liability and cost employers billions of dollars each year. Better screening of drivers could help reduce the on-the-job vehicle crashes and lead to massive savings to employers.

Alcohol-related crashes. Historically, about 40% of all vehicular fatalities occur when a driver has consumed alcohol above the legal limit (Blincoe et al., 2015). Of that total, 1.5% of drivers with alcohol in their system during a fatal crash were on the job. Additionally, the total cost of alcohol-related crashes to employers in 2015 was more than $6 billion, making them one of the most expensive categories of crashes to employers (Network of Employers for Traffic Safety, 2016). Per the US Department of Transportation (DOT), companies hiring for jobs that require driving tasks can be held liable for hiring someone with an alcohol-related conviction or even an alcohol problem that could escalate into criminal behavior (Disqualification of Drivers, 2017).

As mentioned previously, Bassam Jafar v. Elrac, Inc. (2011) showed that companies can be sued for negligently hiring someone with a history of alcohol abuse. Thus, it is important for companies, especially companies in the transportation industry, to screen out applicants not only with an alcohol-related conviction, but also with behaviors that may lead to a future alcohol-related issues. While CBCs cannot provide information to gauge the latter, information posted on Facebook may give more information and insight into drinking behaviors that could lead to a crime, allowing companies to conduct their due diligence to the fullest extent and avoid negligent hiring.

The current study

The study sought to determine if publicly available data from Facebook can be used as a valid pre-employment screening tool to predict alcohol-related convictions. In this study, textual information posted on publically available Facebook profiles of individuals convicted of an alcohol-related crime were compared against publically available textual posts from Facebook profiles from a randomly generated comparable sample. Once the textual post information was captured, text analytics was used to determine the common themes contained in the posts, and the differences between the two groups were used to determine if these themes could validly predict alcohol-related criminality.

Specifically, the study's research questions were:

1. Do the topics posted on Facebook differ between individuals convicted of an alcohol-related crime and a randomly generated sample?
2. Can a predictive model be created that determines whether a person will be convicted of an alcohol-related crime based on the textual information posted on their Facebook profiles?

Method

Participants and Procedure

Participants in this study were 247 convicts from Lafayette, IN and 299 randomly sampled residents of Lafayette, IN. Participants in the convict group were collected by requesting publicly available information from the Tippecanoe County Courthouse. All convictions were made in 2014. A team of research assistants manually gathered data for the convict group such as place of residence, year of birth, other convicted charges, and matching Facebook URLs. In order to collect a representative sample of Lafayette residents for the comparison group, a search on Facebook profiles was conducted based on the frequency of first name letter distributions in the United States, compiled by Joshua Falk (2012) using data from the US Social Security Administration. For instance, in the US, 6% of names begin with the letter “K”. As such, 18 profiles (.06 * 299) from Lafayette with names that begin with K were selected for inclusion in the study. All these Facebook profiles were collected while logged into a blank Facebook account with no friends or information as to limit any biases and inconsistencies that would come with being logged in to a personal Facebook account. Facebook profiles were filtered by location to restrict our control group to Lafayette.

IRB approval was received before collecting data (see Appendix A).

Measures

Facebook textual statuses and comments were manually collected from each Facebook profile and stored in separate Excel files on an encrypted drive. Only posts made by the owner of the account were collected. Comments by others on posts were excluded, along with life updates such as changes in relationship status, jobs, or schools.

Coding Scheme

Facebook profiles were manually scraped and coded by a team of undergraduate and graduate research associates (see Appendix B for instructions used in the data collection). From each profile, all textual statuses and comments written by the participant from one year prior to the conviction date were collected. For the comparison group, research associates scraped text from one year prior to randomly generated dates that matched the date ranges of those being scraped for the conviction group. Research associates were kept blind to whether or not a Facebook profile was of someone who was convicted of a crime or from the comparison group.

Analysis plan

Text Analytics. Text analytics and theme extraction was used to analyze the textual data. Specifically, SAS Text Miner, a part of SAS Enterprise Miner, was used to extract topics and themes from text and assign frequency weights to these themes which were used to analyze and made comparisons between groups based on common themes extracted from their text posts.

The themes were generated based on how frequently certain words and topics appeared in a group's Facebook posts. Each theme was then assigned a frequency weight, which is a value that tells how often that theme appears in the text. A higher frequency weight means that the topic appeared more often in a participant's profile posts. These values are used to compare the comparison group with the convict group to determine any differences in topics between the groups' textual Facebook posts.

Logistic Regression. The themes created through SAS Text Miner were then used as predictors in a logistic regression model and sought to predict group membership (i.e., conviction or comparison group). Both the presence of the theme and theme weight was used in the regression model.

Results

Facebook Themes

There were four common themes found in the sample's textual posts. The themes were not mutually exclusive and participants were able to be assigned to multiple themes. These themes were labeled: Positive, Neutral, Slang, and Spanish. The first theme was termed as "Positive" and described posts that had generally positive affect. While the content of what was contained in the posts varied, posts in this theme displayed a general positive affect. Examples of the terms included in this theme were "great day", "music", "happy", “family”, “love”, and “awesome.” Next, the “Neutral” theme included terms that did not have much emotional value. Posts in this theme were general status updates, such as “Going to church soon”, “Can’t wait for summer”, or “Preparing myself for this doctor visit.” Terms included “god”, “woman”, “future”, “people”, “world”, and “doctor.” The third theme was entitled “Slang.” Profiles categorized into this theme included a number of slang and curse words. For example, some terms were “lol”, “man”, “gonna”, “shit”, “haha”, “cuz”, and “idk.” The final theme was a Spanish theme. This theme was included to account for Facebook profiles with a significant amount of text in Spanish. Hence, all terms in this theme were in Spanish.

Table 1 presents the summary statistics of the themes extracted from and the theme weights of the overall sample. The theme weights are metrics that measure how frequently the themes appear within the sample, based on a combination of overall frequency of use and how many total profiles used words within that theme. Result showed that 21% of profiles contained posts belonging to the Slang theme, 17% had posts belonging to the Positive theme, 10% to the Neutral theme, and only 1% of profiles were categorized into the Spanish theme. The results presented in Table 1 also show that 76% of profiles in the sample were not classified into any themes due to either not having a very limited number or no publically available post information.

Differences in groups

In order to compare the themes between the comparison and conviction group and answer the study’s first research question, a series of independent samples *t*-tests were conducted. The first series of tests examined the extent to which the themes appeared in the two groups by examining the binary variables representing the presence (1) or lack of presence (0) in participants' profiles. As seen in Table 2, three out of the four themes were found to be significantly different between the groups. The Positive theme was found to be higher in the comparison group (M = 0.22, SD = 0.42) than the conviction group (M = 0.12, SD = 0.32); $t(544) = 3.10$, $p = .002$. The Slang theme was found to be higher in the comparison group (M = 0.27, SD = 0.45) than the conviction group (M = 0.13, SD = .34); $t(544) = 4.07$, $p < .001$. The Neutral theme was also found to be higher in the comparison group (M = 0.14, SD = 0.35) than the conviction group (M = 0.05, SD = 0.22); $t(544) = 3.51$, $p < .001$. The Spanish theme was not significantly different from the control group (M = 0.01, SD = 0.12) to the conviction group (M = 0.02, SD = 0.13); $t(544) = -0.27$, $p = .786$. Additionally, the comparison group had significantly fewer profiles with no posts (M = 0.71, SD = 0.45) than the conviction group (M = 0.81, SD = 0.39); $t(544) = -2.65$, $p = .008$.

Next, the difference between the groups in the theme weights were also examined using a series of independent samples *t*-tests. The theme weights represented the frequency or intensity in which a theme appeared in a profile and were scored on a scale from 0 to 0.93 in the current sample, with higher numbers representing higher frequency and/or intensity of a theme in a participant's profile. As shown in Table 3, none of the theme weights were statistically different between the two groups. These results show that while there were differences in whether a theme was present in the profiles of the two groups, the frequencies and intensities of the themes were not different.

Together, these results answer the first research question. When looking solely at the presence of different themes, there were differences between the conviction and comparison group in terms of what they post on Facebook. That is, the two groups differed in whether or not a theme was present in their Facebook profiles. There were more profiles from the comparison group categorized into the Neutral, Positive, and Slang themes. However, the *t*-test for theme weights showed no differences between the two groups. Interestingly, this means that the two groups did not differ in the frequency and intensity of which the themes were present. These results collectively suggest that while people convicted of alcohol-related crimes post differently from regular people, the intensity in which they post about these topics does not differ. These results also indicated that the group convicted of alcohol-related crimes are more likely to have profiles that had either limited number of or no posts on their Facebook profile compared with the comparison group.

Predicting Group Differences

To answer the second research question, two logistic regression analyses were conducted. The first logistic regression used the dichotomous presence of a theme variable as predictors of group membership (i.e., conviction or comparison). As displayed in Table 4, results indicated that only one of the four predictors, Slang, was statistically significant. Specifically, the presence of the Slang theme was predictive of comparison group membership.

The second logistic regression model used the term weights as predictors of group membership. The results, as shown in Table 5, were that none of the term weights were statistically significant predictors of group membership. These results are consistent with the results from the *t*-test in Table 3. That is, the frequency and intensity of the words on profiles do not have predictive value in terms of determining whether someone will be convicted of an alcohol-related crime.

These results provide an answer to our second research question. None of the term weights were predictive of group membership, and only one of the four theme categories were predictive. The results show that Slang may be able to predict group membership. However, the fact that the term weight for Slang was not predictive shows that only the presence of the theme predicts group membership, not the intensity of the themes.

Discussion

During the selection process, companies have to be aware of negligent hiring, and the problems that can be caused by not doing their due diligence when bringing someone into their organization. Negligent hiring can lead to harm towards others and the company and can result in litigation against the organization. One of the ways companies attempt to exercise their due diligence and avoid negligent hiring is through CBCs. Screening with CBCs attempts to weed out people who may pose a risk to the well-being of an organization and its members. However, research has shown that CBCs may not be enough to mitigate the risk of negligent hiring as they are generally limited geographically to a specific county and can fail to report hazardous behavior that is not criminal, but may lead to criminal activity in the future.

Due to the shortcomings of CBCs, information posted on SNSs has been suggested as an additional screening tool. In fact, many organizations already use Facebook in their selection process. SNSs have the potential to be a supplementary screening tool to CBCs, since they provide different information. SNSs can be used to assess P-O fit, judge an applicant's personality and interests, and view any negative behaviors in which they may engage. This study analyzed the viability of Facebook as a pre-employment screening method. Specifically, this study looked at predicting alcohol-related crimes from textual Facebook posts. The main questions asked were whether or not people convicted of an alcohol-related crime posted differently on Facebook compared to normal people and if the textual content posted on Facebook could be used to create a predictive model of alcohol-related crimes.

Results showed that people in the conviction group posted differently from people in the comparison group. Specifically, the comparison group used more Positive terms, Slang terms, and Neutral terms than the comparison group. However, these results were surprising. One of the themes extracted was entitled “Slang.” This theme included terms such as ‘lol’, but also included many swear words. The results showed that the comparison group used these slang terms more frequently than the conviction group, which is counterintuitive. Instead of people convicted of alcohol-related crimes using more slang terms and swear words, the results showed that people in the comparison group used more of those words in phrases in their profile posts.

Logistic regressions were then ran to determine if Facebook posts could predict someone being convicted of an alcohol-related crime. The results showed that the presence of only one of the four themes (Slang) were statistically significant in predicting whether a Facebook profile belongs to the conviction or comparison group. The results from the regression, like the results from the univariate *t*-test, showed that profiles with slang terms were categorized in the comparison group more often than in the conviction group. Essentially, a Facebook profile with more slang had higher odds of being a comparison group profile rather than a conviction group profile. Again, these results were unexpected. One possible explanation for this could be that convicts are more likely to hide their shameful statuses and provide a more presentable profile. They may also have more privacy settings on their profiles. This is seen in the *t*-test results in Table 2; more participants in the convict group had a small number of or no posts.

These findings suggest that Facebook has the potential to be used as a pre-employment screen; however, further research needs to be done. The *t*-tests results showed that while there was a difference between the two groups in the presence of the themes, there were no significant differences in theme weights between the two groups. The predictive model from the logistic regressions showed that only one of the four themes were predictive of alcohol-related convictions. Though not all four themes were predictive, these results suggest that there may be potential in using Facebook to glean information that could predict a criminal conviction. These results potentially open a new pathway for the use of social media in the employee hiring process.

This study has a number of limitations that need to be addressed in future research before social media background screens can be used in pre-employment decisions. Firstly, all textual posts and comments were manually collected by a team of undergraduate and graduate students. Each student was assigned up to 80 profiles to scrape. This process may have caused some human error in the text being scraped, especially since the number of profiles scraped by each student was so high. Automatically web-scraping this data could result in cleaner data, though this would require permission from Facebook. However, while this is a limitation of the study, it could also closely track to what is or would be happening in the field as it is likely that hiring managers would also be examining applicants' Facebook profiles manually.

Another limitation was with the theme extraction results from the text analytics. More specifically, there was a good amount of correlation between the theme topics that were extracted. Thus, it seemed as though the automated features of the text analytics program was unable to pick up variation, or at least more subtle variations, between the Facebook profiles.

Upon closer analysis, Facebook profiles with a large amount of text were often categorized into all three English categories. While this is not inherently an issue, the fact that the theme weights were not able to distinguish between the two groups suggests that further refinement of the themes might be needed and that many of the automated theme extraction processes might be insufficient when making pre-employment decisions. Future research may look at creating theme dictionaries independent from the social media posts and applying them to each group. This would allow for more variation between themes and could provide more significant results. Furthermore, the text analytics process did not generate an alcohol-related theme with terms such as “drinking”, “beer”, or “partying.” An alcohol-related theme could be able to predict alcohol-related convictions.

Facebook itself also provides some limitations. Though the lack of text is a limitation of text analytics, it can also be applied to Facebook. Of our sample, 76% of the profiles either had too little content to code or were simply not made publicly available. This limits the scope of Facebook as a potential screening tool since not everyone will have enough content to provide meaningful information. This lack of publically available information on Facebook profiles also calls into question if social media background checks are feasible. Given that 25 states do not allow potential employers to ask applicants for their social media passwords (“State Social Media Privacy,” 2017), relying on publically available information on SNS such as Facebook might not be a viable route for gather additional information about an applicant.

Social media has plenty of potential for use in organizations. In fact, companies are already using social media in their selection process to judge and filter applicants (e.g. Jobvite, 2014; SHRM, 2016; Willey et al., 2012). Applying social media to the employee selection process can provide many benefits. Organizations may be able to assess person-organization fit through social media (Roulin & Bangerter, 2013). Furthermore, companies may be able to uncover dangerous behaviors that could pose a risk to the company or its employees. Though social media can potentially be a viable selection and screening tool, much more research needs to be done. In today's age where technology is so integrated in the world, social media could change the way of HR and selection.

References

Aamodt, M. G. (2015). Using background checks in the employee selection process. In C. Hanvey & K. Sady (Eds.), *Practitioner's guide to legal issues in organizations*. (pp. 85-110). New York, NY: Springer.

Bassam Jafar v. Elrac, Inc. (N.J. Super. Ct. App. Div., 2011).

Blincoe, L. J., Miller, T. R., Zaloshnja, E., & Lawrence, B. A. (2015). *The economic and societal impact of motor vehicle crashes, 2010 (Revised)* (Report No. DOT HS 812 013). Washington, DC: National Highway Traffic Safety Administration.

Criminal history background information check law and legal definition. (n.d.). Retrieved from https://definitions.uslegal.com/c/criminal-history-background-information-check/

Disqualification of drivers, 49 C.F.R. § 383.51 (2017).

Falk, J. S. (2012). *United States baby name letter statistics*. Retrieved from http://home.uchicago.edu/~jsfalk/misc/baby_names/

Howard v. The Hertz Corporation, Civil No. 13-00645 (D. Haw., 2016).

Jobvite (2014). *Social Recruiting Survey*. Retrieved from https://www.jobvite.com/wp-content/uploads/2014/10/Jobvite_SocialRecruiting_Survey2014.pdf

Jobvite (2015). *The Jobvite Recruiter Nation Survey*. Retrieved from https://www.jobvite.com/wp-content/uploads/2015/09/jobvite_recruiter_nation_2015.pdf

Kittling, N. M. (2010). *Negligent hiring and negligent retention: A state by state analysis.* Paper presented at 4th Annual Section of Labor and Employment Law Conference, Chicago, IL.

Mount, M., Ilies, R., & Johnson, E. (2006). Relationship of personality traits and counterproductive workplace behaviors: The mediating effects of job satisfaction. *Personnel Psychology*, *59*(3), 591-622. doi:10.1111/j.1744-6570.2006.00048.x

Network of Employers for Traffic Safety (2016). *Cost of motor vehicle crashes to employers*. Vienna, VA: Network of Employers for Traffic Safety.

O'Boyle, E. H., Forsyth, D. R., Banks, C. G., & McDaniel, M. A. (2012). A meta-analysis of the dark triad and work behavior: A social exchange perspective. *Journal of Applied Psychology*, *97*(3), 557-579. doi:10.1037/a0025679

Porras, Pecha, and Wolter v. Swedish Medical Center, U.S.D.C. Colorado, (2016).

Reagan v. Dunaway Timber Company, (W.D. Ark. Harr. Div., 2011).

Rosen, L. S. (2017). *The safe hiring manual: The complete guide to employment background checks for employers, recruiters, and job seekers* (3rd ed.). Tempe, AZ: Facts on Demand Press.

Roulin, N. & Bangerter, A. (2013). Social networking websites in personnel selection: A signaling perspective on recruiters' and applicants' perceptions. *Journal of Personnel Psychology*, *12*(3), 143-151. doi:10.1027/1866-5888/a000094

SHRM. (2012). *The use of criminal background checks in hiring decisions*. Alexandria, VA: Society for Human Resource Management.

State social media privacy laws. (2017, January 11). Retrieved from http://www.ncsl.org/research/telecommunications-and-information-technology/state-laws-prohibiting-access-to-social-media-usernames-and-passwords.aspx

Valdez, C. (2015). Hands-tied hiring: How the EEOC's individualized assessment is taking discretion away from employers' use of criminal background checks. *Cleveland State Law Review, 63*(541), 541-566.

Willey L., White B.J., Domagalski T. & Ford J.C. (2012). Candidate-screening, information technology and the law: Social media considerations. *Issues in Information Systems, 13*(1), 300-309.

Appendix A

Appalachian STATE UNIVERSITY

INSTITUTIONAL REVIEW BOARD
Office of Research Protections
ASU Box 32068
Boone, NC 28608
828.262.2692
Web site: http://researchprotections.appstate.edu
Email: irb@appstate.edu
Federalwide Assurance (FWA) #00001076

To: Cameron Brown
Psychology
CAMPUS EMAIL

From: Robin Tyndall, Institutional Review Board
Date: 2/20/2017
RE: Notice of IRB Approval by Relying on External IRB Review of Off-Campus Research
Agrants #:
Grant Title:

STUDY #: 17-0132
STUDY TITLE: Looking for Red Flags: Using Facebook as a Background Screen in Hiring
Submission Type: Initial
Expedited Category:
Approval Date: 2/18/2017
Expiration Date of Approval: 11/10/2017

This submission has been approved by Utah State University for the period indicated in the submitted documentation. The Appalachian State University IRB will rely on the review of your study by the external IRB.

Investigator's Responsibilities:

Federal regulations require that all research be reviewed at least annually. It is the Principal Investigator's responsibility to submit for renewal and obtain approval before the expiration date. You may not continue any research activity beyond the expiration date without IRB approval. Failure to receive approval for continuation before the expiration date will result in automatic termination of the approval for this study on the expiration date.

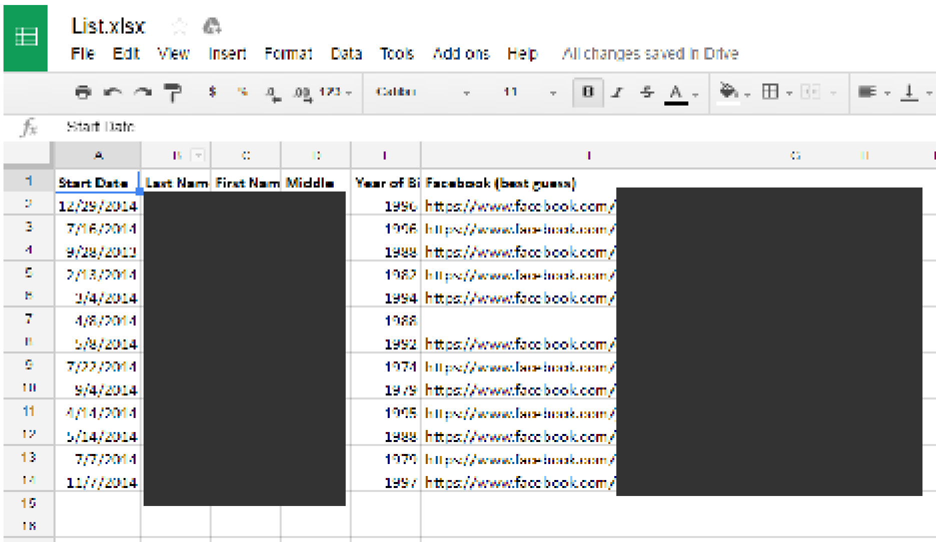
You are required to obtain IRB approval for any changes to any aspect of this study before they can be implemented. Should any adverse event or unanticipated problem involving risks to subjects occur it must be reported immediately to the IRB by logging into the new IRB Information System (IRBIS) using the following link: https://appstate.myresearchonline.org/irb/index_auth.cfm. Note the Dashboard located on the left hand side of the page. Under the title, "Create New Submission," click the link titled, "Unanticipated Problem," and follow the instructions to filling out the application. Please contact the IRB office at either 262-2692 or irb@appstate.edu if you experience any issues or have any questions. We will be happy to assist you.

Best wishes with your research!

CC:
Shawn Bergman, Psychology
Bhavik Modi, Psychology

Appendix B

7.

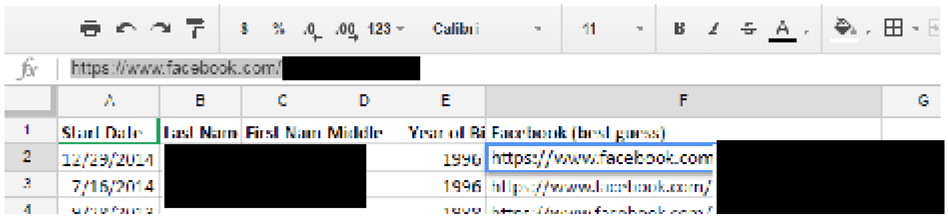


- **Inside of this "List" document is every profile that you have been assigned to**
- **This page will be different and unique for every person in the research team.**
- **After examining the list, we're going to want to take the first person you've been assigned to (EX: Nicholas Swope) and use a separate excel document to record their statuses**

9.

- **After finding a profile to analyze in the List folder, return back to the list of template excel files in your main directory.**
- **We're going to want to re-name each Template file to correspond with the respective name of the person we are coding for (look above).**
- **If I was going to code for [redacted]'s facebook profile, I would need to change the name of my template file to have his name in the file**
- **You will repeat this process for every profile that you analyze**

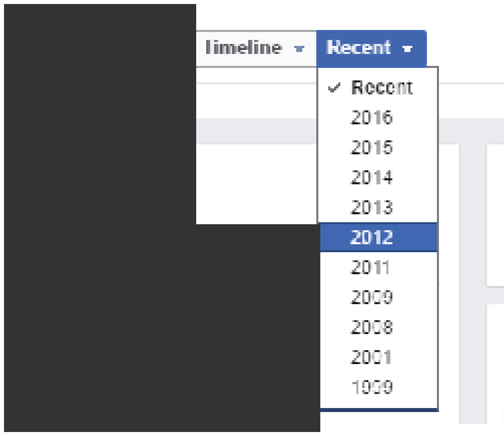
11.



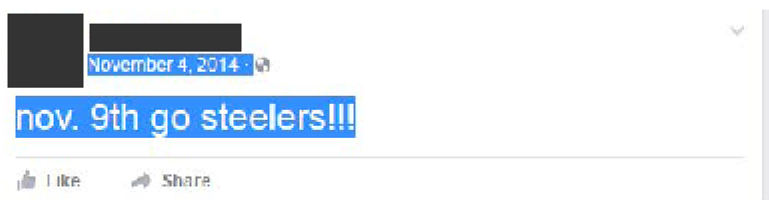
- Return back to the "List" document and copy the URL for the individual's Facebook page. You're ready to start coding!
- IMPORTANT: Only look at statuses before the "Start Date" given in the "List" document
  - o We are only looking at a 1 year period before the Start Date
  - o EX: If the "Start Date" is 12/29/2014, then only look at statuses between 12/29/2013 and 12/29/2014

[redacted]

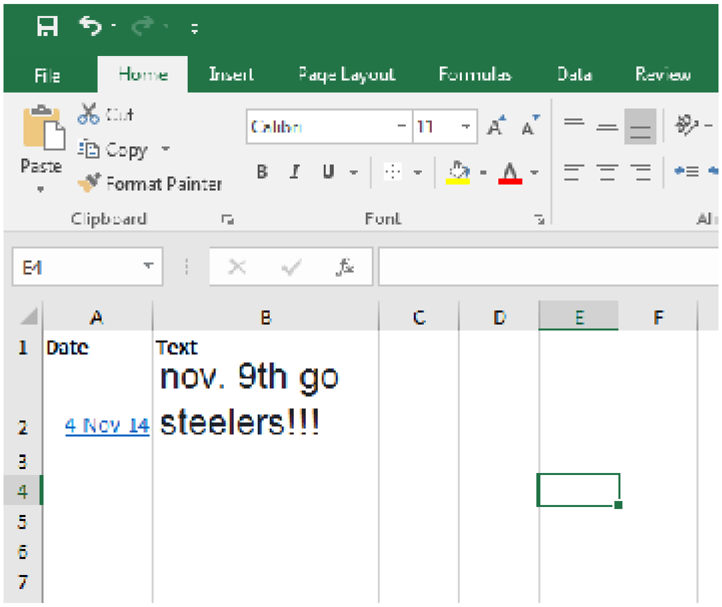
- After you reach the homepage of your respective FB profile, look for the "Recent" button at the top of the screen (NOTE: You may need to scroll down for a few seconds before this task bar pops up).



- Clicking on the "Recent" button allows you to choose which year you would like to examine for the respective profile. REMEMBER: Examine the "Start Date" in the original List document to figure out which year to examine. Any date before the "Start Date" given is perfectly fine, just not anything afterwards!



- So what are we looking for? We are explicitly looking for pure text statuses (as shown above)
- Go through each text status + date and copy and paste it into the Excel document that you created for the individual.
- Copy as many statuses as you can! Don't pay attention to what the status says, just copy every one of them that you see.



- Keep both the Excel document and Facebook page open at the same time to make the process easier and more efficient! (Keep them both in windowed mode)
- Repeat this step with multiple text statuses until you can't find anymore!
- REMEMBER: Different privacy settings will mean that some profiles will have almost zero statuses available publically, yet some will have a plethora. Make sure to record as much information as possible when it is openly available to help with our analyses.

Table 1

Descriptive Statistics

Variable	*n*	*M*	*SD*	Min	Max
Spanish	546	0.01	0.12	0	1.00
Positive	546	0.17	0.38	0	1.00
Slang	546	0.21	0.41	0	1.00
Neutral	546	0.10	0.30	0	1.00
Spanish Weights	546	<0.01	0.02	0	0.30
Positive Weights	546	0.05	0.08	0	0.62
Slang Weights	546	0.08	0.11	0	0.53
Neutral Weights	546	0.03	0.07	0	0.93
None	546	0.76	0.43	0	1.00

Note: Spanish, Positive, Slang, and Neutral represent the presence (1) or absence (0) of the respective theme. Theme weights are values that indicate frequency or intensity in which a theme appears in participants' profile.

The Spanish theme accounts for profiles with a significant amount of content in Spanish. The Positive themes includes words with positive affect. The Neutral theme includes words without much emotional salience, such as mundane status updates. The Slang theme includes slang and curse words.

None is coded as: 0 = categorized in zero themes; 1 = categorized in one or more theme.

Table 2

Independent Samples T-Test: Themes

Variable	Group	*n*	*M*	*SD*	*t*	df
Spanish	Control	299	0.01	0.12	-0.27	544
	Convict	247	0.02	0.13		
Positive	Control	299	0.22**	0.41	3.10	544
	Convict	247	0.12**	0.32		
Slang	Control	299	0.27***	0.45	4.07	544
	Convict	247	0.13***	0.34		
Neutral	Control	299	0.14***	0.35	3.51	544
	Convict	247	0.05***	0.22		
None	Control	299	0.71**	0.45	-2.65	544
	Convict	247	0.81**	0.39		

Note: Spanish, Positive, Slang, and Neutral represent the presence (1) or absence (0) of the respective theme.

The Spanish theme accounts for profiles with a significant amount of content in Spanish. The Positive themes includes words with positive affect. The Neutral theme includes words without much emotional salience, such as mundane status updates. The Slang theme includes slang and curse words.

None is coded as: 0 = categorized in zero themes; 1 = categorized in one or more theme.

*$p < .05$. **$p < .01$. ***$p < .001$.

Table 3

Independent Samples T-Test: Theme Weights

Variable	Group	*n*	*M*	*SD*	*t*	df
Spanish Weights	Control	299	>0.01	.03	-0.65	544
	Convict	247	0.01	.02	-0.68	523
Positive Weights	Control	299	0.05	.09	0.69	544
	Convict	247	0.05	.08	0.70	542
Slang Weights	Control	299	0.08	.13	0.56	544
	Convict	247	0.07	.10	0.57	539
Neutral Weights	Control	299	0.03	.06	0.04	544
	Convict	247	0.03	.08	0.04	416

Note: Theme weights are values that indicate frequency or intensity in which a theme appears in participants' profile.

The Spanish theme accounts for profiles with a significant amount of content in Spanish. The Positive themes includes words with positive affect. The Neutral theme includes words without much emotional salience, such as mundane status updates. The Slang theme includes slang and curse words.

*p < .05. **p < .01. ***p < .001.

Table 4

Logistic Regression Predicting Conviction or Comparison Group using Theme Presence

	β	S.E.	Wald	df	*p*	Exp(β)
Spanish	.21	.73	.08	1	.776	1.23
Positive	.27	.41	.43	1	.513	1.31
Slang	-.79	.34	5.33	1	.021	.45
Neutral	-.75	.44	2.85	1	.091	.48
Constant	-.02	.10	.03	1	.860	.98

Note: Group membership was coded as 0 = control and 1 = conviction group. Spanish, Positive, Slang, and Neutral represent the presence (1) or absence (0) of the respective theme.

The Spanish theme accounts for profiles with a significant amount of content in Spanish. The Positive themes includes words with positive affect. The Neutral theme includes words without much emotional salience, such as mundane status updates. The Slang theme includes slang and curse words.

$^*p < .05$

Table 5

Logistic Regression Predicting Conviction or Comparison Group using Theme Weights

	β	S.E.	Wald	df	*p*	Exp(β)
Spanish Weights	2.43	3.64	.45	1	.504	11.39
Positive Weights	-.99	1.77	.31	1	.578	.37
Slang Weights	-.23	1.30	.03	1	.857	.79
Neutral Weights	.91	1.74	.28	1	.600	2.49
Constant	-.17	.11	2.45	1	.117	.85

Note: Group membership was coded as 0 = control and 1 = conviction group.

Theme weights are values that indicate frequency or intensity in which a theme appears in participants' profile.

The Spanish theme accounts for profiles with a significant amount of content in Spanish. The Positive themes includes words with positive affect. The Neutral theme includes words without much emotional salience, such as mundane status updates. The Slang theme includes slang and curse words.

*$p < .05$.