information

MDPI

Article

Social Media and the Scourge of Visual Privacy

Jasmine DeHart *, Makya Stell and Christan Grant

School of Computer Science, University of Oklahoma, Norman, OK 73019, USA; makyastell@ou.edu (M.S.); cgrant@ou.edu (C.G.)

* Correspondence: dehart.jasmine@ou.edu

Received: 1 January 2020; Accepted: 19 January 2020; Published: 21 January 2020

Abstract: Online privacy has become immensely important with the growth of technology and the expansion of communication. Social Media Networks have risen to the forefront of current communication trends. With the current trends in social media, the question now becomes *how can we actively protect ourselves on these platforms*? Users of social media networks share billions of images a day. Whether intentional or unintentional, users tend to share private information within these images. In this study, we investigate (1) the users' perspective of privacy, (2) pervasiveness of privacy leaks on Twitter, and (3) the threats and dangers on these platforms. In this study, we incorporate techniques such as text analysis, analysis of variance, and crowdsourcing to process the data received from these sources. Based on the results, the participants' definitions of privacy showed overlap regardless of age or gender identity. After looking at the survey results, most female participants displayed a heightened fear of dangers on social media networks because of threats in the following areas: assets and identity. When the participants were asked to rank the threats on social media, they showed a high concern for burglary and kidnapping. We find that participants need more education about the threats of visual content and how these privacy leaks can lead to physical, mental, and emotional danger.

Keywords: privacy; social media networks; emerging technology; social impact of technology

1. Introduction: Social Media Networks, Privacy, and Technology

Increase in technology use has expanded the fronts that privacy advocates must fight. Privacy concerns for credit breaches (e.g., Equifax in 2017), government breaches (e.g., Office of Personnel Management breach in 2015), and personal conversations require personnel to exercise control of the information disseminated; otherwise, the fundamental concept of privacy is violated. Ensuring that your privacy is secure online and safe from outsiders is key to protecting yourself, your family, and your assets. As technology and communication expands, new facets of breaches arise. These developments promote breeding grounds for privacy leakage due to the changes in social culture, development of technology features, and changes in the targeted audience.

Ninety percent of Facebook profiles contain at least one image, 87.8% of users share their birth date, phone numbers are listed on 39.9% of the profiles (including 28.8% that contain cell phone numbers), and 50.8% of users share current residency [1]. Additionally, revealing information such as birthdate, hometown, current residence, and phone number can be used to estimate the user's social security number and exposes them to potential financial and identity threats [1]. This type of information can be found in posts that contain visual content.

Images that contain privacy leaks may expose intimate information that is harmful to your finances, personal life, and reputation [1]. Visual privacy leaks include any instance in which a transfer of a personal identifying image is shared. Leaks that involve images are especially pernicious because they are so unexpected. For example, private images can include baby faces, credit cards, phone numbers, social security cards, house keys and other personal identifiable information [2]. Anything

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/324931

posted to these Social Media Networks (SMNs) can be exposed to someone even after removal of the content [3]. Users re-use the same content across SMNs, this duplicated content can be used to find profiles across platforms [1]. *On SMNs, the users' information and visual content can intentionally or unintentionally be shared even though there may be a privacy risk contained [1].* From visual content, attackers can extract textual information, including credit card numbers, social security numbers, place of residence, phone numbers and other information [1,4]. This content can consequently create an opportunity for "cyberbullying" of other users [3].

Through this investigation, we explore the state of privacy on social media networks with an emphasis on visual privacy leaks and the future of privacy for its' users. This study creates a foundation to understand the users' concerns about the possibility of having secrets in the future and the threats that emerging technologies will expose them to. In summary, the purpose of our work is to understand:

- The privacy perspective of a user can be subjective (Section 3). With this investigation, we were able to uncover the subjectivity of users through age and gender demographics. This section also demonstrates the differences privacy and visual privacy.
- Visual content privacy leaks are common among users on Twitter (Section 4.1). In this study, we found *severe* and *moderate* privacy leaks on Twitter. In comparison to the data set, these numbers were low. We believe that with relevant and trending search terms we will be able to improve future results.
- Several threats and dangers are heightened due to the accessibility of social media (Section 3.4). This work provided an understanding of the most threatening dangers to users as well as a hierarchy of dangers in correlation to the rankings of the participants.

2. Previous Literature

Previous work studying user privacy on SMNs have focused on multiparty privacy conflicts [5,6], images or text content from users [1,7–14], third-party applications that supervise privacy [11,13–15], cultures in university settings/communities [1], the users' SMNs privacy settings [1,3,16,17], studies on users' attitudes, intention, or behaviors [7,17], and children/teenagers' interactions with SMNs [18,19].

Online privacy is important to the growth of technology and the expansion of communication. Since social media has become a popular social forum, our private lives will continually be lived out in a public domain. To many, privacy on social media networks is user-dependent. People tend to share different content, have different privacy settings, and different perspectives of privacy. There are several papers that examine privacy settings of users' accounts in correlation to their privacy leakage [1,16,17]; moreover, there are privacy concerns that go beyond privacy settings [1]. Through the exploration of users' attitudes and intentions on social media platforms, upcoming developments consider the fact that privacy settings of social networks are failing the users [17]. To help the user customize their privacy settings, authors [20] suggests six privacy profiles: privacy maximizers, selective shares, privacy balancers, time savers, self-censors, and privacy minimalists. Investigating privacy settings of users on social media is important, but it is also important to explore the disclosed information from users [1,21]. The intersection of privacy settings and third-party applications on these platforms create opportunity for more risks [16]. Social networks need to take meaningful action to decrease the exposure of personal information. The risks of engaging on social media could outweigh the benefits. The exploration of these risks was investigated by [3]. Studies have provided third-party applications that will help reduce the amount of visual privacy leaks until social media platforms employ further action [11,15,22,23].

The perception of privacy is highly subjective and user-dependent [10], which is shown by literature focusing on users' attitudes towards privacy [17]. As people engage on social media, the images posted can contain potential privacy leaks for users [10]. In several studies it has been found that images on social media can pose danger [1,3,5,8]. With visual content on social media someone can uncover personal identifying information that can be collected from them [8]. Visual content can also become a gateway for multiparty conflicts among users [5]. These conflicts can arise due to

feelings of ownership, privacy boundaries, and privacy perspectives of the individuals in the content. Looking further into self-censoring and reduction of multiparty conflict, users can implement privacy preserving procedures to reduce identity, association, and content disclosure [24].

Researchers have developed mitigation techniques for visual privacy that range between intervention methods and data hiding [25]. To protect visual privacy, these methods can be implemented before posting the content or after identifying private objects. Most privacy technology uses one or more of these five protection techniques: intervention [23,26], blind vision [2,4,12], secure processing [12,15], redaction [27–30], and data hiding [12,15,31].

Beyond protecting ourselves from dangers, there are also minors at risk. Studies exploring teenager attitudes towards privacy note that teenagers tend to be more open about their lives on social media when compared to older users [18,19] . On social media, teenagers and children are exposed to potential dangers like stalking and sexual predators because of this openness. Studies emphasize the importance of stranger danger and insider threat for minors on SMNs because the real threat lies within the users' friends because of interpersonal sharing [32]. The collection of personal information through social engineering and other techniques could affect national security and government officials on these platforms [7]. With the use of surveys, researchers [7,9,17] can understand what information they share and gauge their understanding of privacy in respect to their ethnicity.

From this literature, we can begin to uncover the importance of privacy and the growing need for evolving technologies to combat online threats. Previous works have discussed concerns with visual content focusing on multiparty conflicts, third-party applications, privacy settings, and the danger of this content. The current state of this field shows the importance to continue investigation and development of visual privacy and mitigation techniques to protect SMNs users. The future of this field is in the development of mitigation techniques, understanding the pervasiveness of visual privacy leaks, and helping users understand the correlation of privacy to threats and dangers on these networks. Our research investigates based on the foundation of these works. With this foundation, we explore the future of privacy on SMNs through participant surveys, data collection from Twitter, and analysis of these results. This work details the attitudes and perspectives toward visual privacy, and the data collection results from Twitter.

3. Attitudes and Perspectives towards Visual Private Information

The user's perspective of freedoms, beliefs, ownership, and vulnerabilities aid to guide their decision-making process when engaging on social media. On these platforms, users determine what information to share based on their perceived freedoms and feelings of security. Each user's perspective of privacy will vary based on their subjectivity.

3.1. Survey Overview

We deployed two surveys that asked participants about their knowledge, experiences, and perspectives of social media networks. This encompassed social engagement behaviors and visual privacy leaks. Many of these survey questions were derived from our study, but had influence from others [7,9,17]. The 250 participants completed the surveys online and were not required to answer every question. The first survey (IRB #10299) is used to gauge participants engagement across various platforms. The second survey (IRB #11349) focused on participants use of Twitter and their engagement with visual privacy.

From the surveys, we look at the participants use of social media networks, definitions of privacy, and observations of privacy leaks on social media networks. To avoid bias, the questions were randomized and the survey concluded with demographic questions [33,34]. Table 1 recaps the questions that were asked of the participants in the first survey. Each response that included text entry from the user was analyzed using text analysis methods. The responses were segmented into categories: age and gender identity; then analyzed using Elbow–Knee plots, clustering, and feature weights.

On average, participants engage on SMNs for 11–20 h per week (Table 1B). Majority of the participants have multiple SMN profiles from different platforms (e.g., Facebook, Twitter, Reddit). In this survey, that the leading platforms are YouTube (98 participants), Instagram (80 participants), and Snapchat (65 participants) which are all image and video-based platforms (Table 1A). From this survey, the content posted the most across SMNs is via images and videos. Forty-seven percent of participants post images and 11% of participants post videos; only 42% of participants post textual content (Table 1C). With images alone or the combination of video content, visual content is becoming the prominent method of posting content on SMNs. These responses support literature stating that with the growth of technology and SMNs, the percentages of content posted as images and videos will continually increase.

Table 1. This table displays the outline of the IRB #10299 survey that was completed by participants. The survey was compromised of multiple-choice questions and short answer.

Item	Question
A	Of what Social Media Networks (SMNs) do you consider yourself a frequent user?
B	How many hours per week do you spend on social media networks?
C	What type of content do you usually post on social media?
D	Do you post any of these types of images or videos on your SMNs?
E	How would you define privacy? (in one sentence)
F	Would you define privacy the same for social media networks?
G	Personally identifying information is information that can be used to uniquely identify, contact, or locate a person. Agree or Disagree?
H	Privacy leaks include any instance in which a transfer of personal identifying visual content is shared on Social Media Networks. Private visual content exposes intimate information that can be detrimental to your finances, personal life, and reputation. Agree or Disagree?
I	Would you consider any of these images to have identifying information?
J	As a typical user of Social Media Networks (SMNs), if you were to post these items would you consider these items to be private?
K	Drag and drop the following dangers in order of most threatening (most threatening ranked 1 and so on).
L	Do you believe there are other dangers on Social Media Networks? If so, list them.
M	What type of threat would these items fall under?
N	Do you believe that conflict (e.g., bullying, domestic disputes) can increase the occurrence of privacy leaks?

3.2. Pre-Processing Raw Survey Responses

Once the data collection process was complete, we began pre-processing the raw data using text analysis [35], natural language processing [36], and regular expressions [37]. We conducted this processing in four steps.

Word Tokenization. This method splits each text entry provided by the participants into sentences and then several tokens. Most of the tokens were split on the whitespace in between tokens.

Lemmatization. This task replaces words that have prefixes and suffixes with their root word. Lemmatization allows us to treat the list of tokens that are used slightly differently as the same word. We use the WordNet dictionary for lemmatization [38].

Combining similar words. Some words are completely different with the same meaning, in these cases we created a list to make explicit substitutions (i.e., birthdate, birthday, bday).

Stopword removal. We removed stopwords and articles using a standard English language list.

With these methods employed, it would reduce the error of the inverse document frequency and the term frequency for word analysis and document clustering.

3.3. Surveyed Definitions of Privacy

To begin the analysis of these survey results, we analyzed the common themes from their definitions of privacy. Among the 250 participants that completed the survey, 154 participants responded to Question E in Table 1. To obtain the most meaningful terms through the set of answers, we compute the TF-IDF scores for each term and then take the average score across all documents where the terms appear. That is, for a term t that appears in an answer d among the set of all answers D. The term frequency (tf) is the frequency a term t appears among any term t in the answer, $\text{tf}(t,d) = \frac{\sum_{|t \in d|} 1}{\sum_{|t' \in d|} 1}$. The document frequency is the number of times a term t appears across all answers $(d \in D)$. The document frequency is given by, $\text{df}(t,D) = \sum_{d \in D} \mathbb{1}(t \in D)$. The inverse document frequency is a factor that down weights terms that appear too often across all documents. These words are viewed as less important. The idf formula is given by $\text{idf}(t,d) = 1 + \log \frac{1+n}{\text{df}(t,D)}$. The average importance of each term in the data set is calculated by the product of the inverse document frequency and average term frequency for each document the term appears as shown in Equation 1.

$$\text{avg(TF-IDF)}(t,D) = \text{idf}(t,D) * \text{avg}_{t \in d}(\text{tf}(t,d)). \tag{1}$$

In their responses, the words information, personal, private, and share are the most relevant words used to detail how they envision privacy on social media and in the real world (Table 2). Majority of the participants were adamant that their definition of privacy would not change regarding their physical life or digital one. However, a small subset of participants stated that their definitions of privacy would not be the same. This definition change could be attributed to the participants feelings about the levels of privacy, the unknown factors that exist on a digital network, or fears of exploitation by companies or scammers on these platforms.

Table 2. The weight for each term is computed by averaging the Term Frequency and the Inverse Document Frequency (TF-IDF) scores over all responses.

Term	Avg. (TF-IDF)
`information`	0.1418
personal	0.1254
private	0.0785
share	0.0655

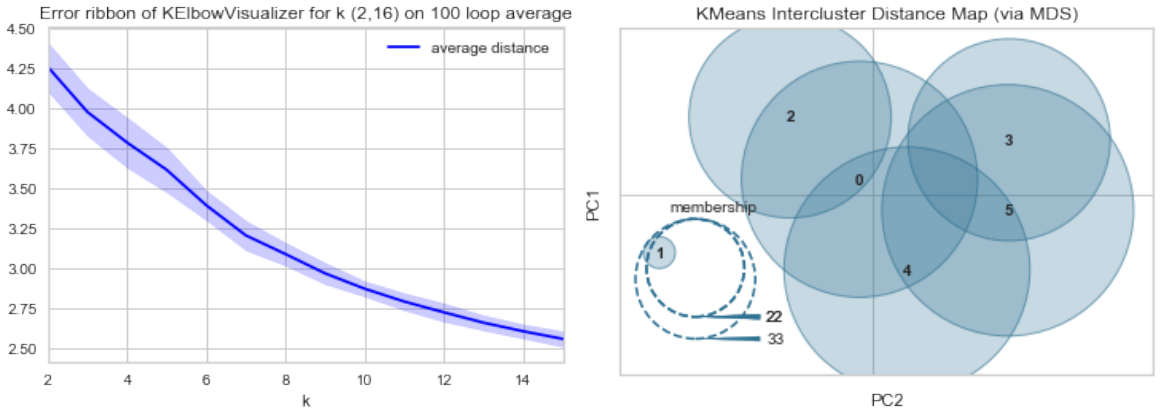
3.3.1. Privacy Definitions by Cluster

The definitions given where further clustered into two groups using Kmeans and yellowbrick clustering via KElbowVisualizer [39]. We deployed Kmeans clustering model and found the elbow of the data using the yellowbrick clustering package. For evaluation we use the *Calinski Harabasz* method to find the optimal cluster size [40]. This method computes the ratio of distribution between clusters and the distribution of points within the clusters. The other scoring metrics provided by yellowbricks include distortion and silhouette. We chose the *Calinski Harabasz* method because it gave us the desired separation focusing on intracluster similarity and intercluster differences—rewarding the best clustering based on the total size and number of clusters. This method uses Equation 2.

$$\frac{SS_B}{SS_W} * \frac{N-k}{k-1} \tag{2}$$

In this equation, SS_B is the overall intercluster distance, SS_W is the overall intracluster distances, N is the total number of data points, and k is the number of clusters. Using this scoring method, we ran the KElbowVisualizer to find cluster values ranging from 2 to 16. In this process we noticed that

the elbow returned at various points for multiple runs. To alleviate this issue, we ran the model 100 times and picked the best mean scores. Based on lowest error and consistency of performance, the best breakeven point was at $k = 6$ (Figure 1a). The intercluster distance shows how strong the correlation is between the clusters and keywords. From Figure 1b, we notice that Cluster 1 is completely different from the remaining 5 clusters. However, cluster 0 and clusters 2–5 have a strong overlap of keywords.



(**a**) (**b**)

Figure 1. Diagrams of the Elbow–Knee scores and errors using Calinski Harabasz method. (**a**) Diagram of the Error Ribbon for the Elbow–Knee Plots using Calinski Harabasz method with cluster size (k) ranging from 2 to 16. (**b**) Diagram of Intercluster Distance Map using YellowBricks Calinski Harabasz method with cluster size (k) = 6.

In those clusters, we see the zeroth cluster included 36 definitions and the words: information, share, and want (Table 3). In Cluster 0, the privacy revolves around authorization, freedoms, and rights of the users on these social media platforms. Cluster 1 included 33 definitions and included the words: social, address, and security. In this cluster, we derived the theme to be protection of personal information regarding physical boundaries, digital security, and personal identifying information. Cluster 2 included 27 definitions and included the words: know, want, and people. In Cluster 2, we hypothesize the definition of privacy emphasized their ability to share information at discretion of the owner without being tied back to that information. Cluster 3 included 30 definitions and included the words: right, ability, and control. The definition for Cluster 3 focuses on accessibility and knowledge of others. The participants in this cluster want to protect themselves their information being shared in a public domain and keep the information disseminated about them in a controlled environment. Cluster 4 included 16 definitions and included the words: passwords, control, and information. The participants clustered in group four defined privacy by access and use of information. It is the user's right to control access of data to keep their information safe. Cluster 5 included 13 definitions and included the words: personal, private, and identify. The last group of participants focus on individual subjectivity about privacy with a focus on personally identifiable information (e.g., social security number, address). Table 3 provides a synopsis of the cluster key words and scores to show their level of importance to each respective cluster.

Table 3. Cluster breakdown of the top terms used to define privacy using average TF-IDF scores.

Cluster (Total Size)		Keyword Score	
0	(36)	information	0.1651
		share	0.1086
		private	0.0870
		want	0.0770
1	(33)	social	0.0984
		address	0.0938
		security	0.0883
		information	0.8171
2	(27)	know	0.1420
		want	0.1315
		people	0.1201
		information	0.1075
3	(30)	information	0.1813
		right	0.1442
		ability	0.0898
		control	0.0874
4	(16)	personal	0.3792
		information	0.3785
		passwords	0.1119
		control	0.09364
5	(13)	personal	0.2380
		private	0.1271
		identify	0.0769
		share	0.0769

3.3.2. Privacy Definitions by Gender Identity

These definitions were further broken down into gender identities. In the section of female participants, the most important words are information, personal, private, share, and social. Of the male participants, the most important words are information, personal, want, control, and private. The remaining participants found the most important words to be birthdate, blood, date, location, and security. It is further noted that the female participants are more concerned with personally identifying information regarding harm and hacking, while the male participants are concerned with financial attacks and exposed information.

In Table 4, the top five keywords are identified for each gender identity. Collectively from every group the keywords are: information, personal, private, share, social, want, control, birthdate, blood, date, location, and security. For further investigation, we began to look at the statistical analysis to find the significance of the words for each gender identity subgroup. In this process we used the analysis of variance (ANOVA) method to analyze the differences between the groups from our participants. Table 5 explore the statistical values produced from this analysis.

To understand the significance of each category, we look at the null hypotheses and p-values associated with the independent and dependent variables. Our null hypothesis states that there are no significant differences in gender identities and the associated keywords. Our p-value threshold is set at 0.05. In the gender comparisons of Female vs. Male, Female vs. Other, and Male vs. Other; we reject the null hypothesis because differences exist in the keywords for gender identities.

Table 4. Top five terms used to define privacy using average TF-IDF scores for each by gender demographic.

Gender Identity (Total Size)		Keyword Score	
Female	(71)	information	0.1361
		personal	0.0985
		private	0.0751
		share	0.0705
		social	0.0400
Male	(82)	information	0.1346
		personal	0.1038
		want	0.0631
		control	0.0545
		private	0.0542
Other	(1)	birthdate	0.3779
		blood	0.3779
		date	0.3779
		location	0.3779
		security	0.3779

Table 5. Calculated ANOVA score for top words used among gender identities to define visual privacy.

Gender Identity Comparison	*f*-Value	*p*-Value
Female vs. Male	5.9749	0.0061
Female vs. Other	6.0464	0.0222
Male vs. Other	6.4194	0.0189

3.3.3. Privacy Definitions by Age Group

These definitions were further clustered into two age groups. In the section of participants ranging between 18–25, the most important words are information, personal, right, control, and private. The definitions of this group revolve around preservation information from hackers and government surveillance. Of the participants over the age of 26, we see the most important words are information, personal, anything, private, and share. The second group's definitions seem to center around a common theme of personal information in relation to external sources while considering alternative factors that could play a role in the dissemination of information.

In Table 6, the top five keywords are identified for each age group. Collectively from each group the keywords are: information, personal, private, share, control, and right. For further investigation, we began to look at the statistical analysis to find the significance of the words for each age subgroup. In this process we used the analysis of variance (ANOVA) method to analyze the differences between the groups from our participants. Table 7 explore the statistical values produced from this analysis.

Table 6. Top words used to define privacy using TF-IDF Scores for each document via age demographic. This table includes the cluster label and the top five words by the highest TF-IDF score.

Age (Total Size)		Keyword Score	
18–25	(111)	information	0.1167
		personal	0.0777
		right	0.0523
		control	0.0512
		private	0.0492
26 & over	(43)	information	0.1395
		personal	0.1212
		anything	0.0869
		private	0.0805
		share	0.0585

Our null hypothesis states that there are no significant differences in age groups in respect to the keywords. In the age comparison of 18–25 v. 26 & up; we retain the null hypothesis because differences do not exist in the keywords for age.

Table 7. Calculated ANOVA score for top words used among age groups to define visual privacy.

Age Group Comparison	*f*-Value	*p*-Value
18–25 v. 26+	0.3275	0.5776

3.3.4. Is Visual Privacy Defined Differently?

From the survey, approximately 6.5% of participants defined visual privacy and general privacy differently. The keywords listed in Table 8 are similar to the word reference by other groups however, these definitions combined the words to form a completely different meaning. They focus on the *information gain* of companies, *lack of control* over your privacy, and the risks on those platforms. This group emphasized how social media and visual privacy creates more risks for the users. The prevalence of visual content and the growth of social media begins to open more doors for attacks and dangers.

Table 8. Top words used to define visual privacy using TF-IDF Scores for each document.

Keyword	Score
private	0.1513
information	0.1280
media	0.1083
social	0.1083
share	0.0952

3.4. Attack Vectors and Existing Dangers

From this survey we investigated what users perceived to be privacy leaks and the dangers of exposed leaks on social media networks. We asked participants if they would consider certain items to be privacy leaks. From this question (Table 1J), we see that 97% of participants identify credit or debits cards, driver's license, social security numbers, and passports as the highest ranked privacy leaks. Following close behind are birth certificate (96%), phone numbers (90%), personal letters (85%), and keys (83%). Participants did not consider images of babies and children to be a privacy leak if posted on social media by their guardians. Sixty-five percent of participants state that they have seen these type of privacy leaks on social media networks. From this, participants to identified keywords or phrases that correlate to those privacy leaks. With this investigation we uncovered hashtags and words such as *#stayoffthesidewalk, #licensedtodrive*, and *#racisttwitter*. Majority of participants stated that they do not recall the phrase that was used in correlation to the images, but did notice privacy leaks on their news feed. The words collected from this survey were used for our data collection process in Section 4.

Next, we ask participants to rank dangers (e.g., burglary, kidnapping, stalking) in reference to what seems to be most threatening (Table 1K). The top threats are kidnapping, burglary, and stalking. Table 9 displays the percentage of votes for the threat in each position. Along with the threats, participants also mentioned cyberbullying, echo chambers, and social isolation (Table 1L).

Table 9. Threats are listed in their respective order based on survey results. In each ranking column shows each threat ranking for dangers and their associated vote percentage for that position; the highest vote for each item is highlighted.

Threat	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6
Kidnapping	**52.38**%	15.48%	10.71%	3.57%	9.52%	8.33%
Burglary	20.24%	**35.71**%	17.86%	10.71%	7.14%	8.33%
Stalking	5.95%	14.29%	**25.00**%	16.67%	23.81%	14.29%
Financial Threat	4.76%	14.29%	23.81%	**30.95**%	17.86%	8.33%
Identity Theft	14.29%	17.86%	14.29%	25.00%	**23.81**%	4.76%
Explicit Websites	2.38%	2.38%	8.33%	13.10%	17.86%	**55.95**%

For further investigation, we began to look at the statistical analysis to find the significance of the dangers for each subgroup. In this process we used the Analysis of variance (ANOVA) to analyze the differences between the groups from our participants. Tables 10 and 11 explore the statistical values produced from this analysis.

Table 10. Statistical analysis of gender related differences of danger assessment results using the Analysis of variance (ANOVA) method. In the table we see the f-value and p-value for Female and Male genders.

ANOVA analysis of danger distribution among gender identity						
Statistic Value	**Burglary**	**Kidnapping**	**Explicit Websites**	**Financial Theft**	**Identity Theft**	**Stalking**
f-Value	5.2662	2.8248	6.0343	1.8150	4.8928	2.9822
p-Value	0.0063	0.0629	0.0031	0.1668	0.0089	0.0541

To understand the significance of each category, we look at the null hypotheses and p-values associated with the independent and dependent variables. For each category, we have a null hypothesis that states if there are no significant differences in gender in the respective category. Our p-value threshold is set at 0.05. In the categories of burglary, explicit websites, and identity theft; we reject the null hypothesis because differences exist in gender. For the categories of kidnapping, financial theft, and stalking; we retain the null hypothesis because no significant differences exist for gender identities. Within the male and female clusters, the female group displayed a higher concern for the threat of being posted on an explicit website unlike their male counter parts.

Table 11. Statistical analysis of gender related differences of danger assessment results using the Analysis of variance (ANOVA) method. In the table we see the f-value and p-value for the age groups: 18–25 & 26 and over.

ANOVA analysis of danger distribution among age groups						
Statistic Value	**Burglary**	**Kidnapping**	**Explicit Websites**	**Financial Theft**	**Identity Theft**	**Stalking**
f-Value	0.3491	0.4125	4.1532	3.7922	3.5000	5.2348
p-Value	0.7059	0.6628	0.0178	0.0250	0.0330	0.0064

For each category, we have a null hypothesis that states if there are no significant differences in age for the respective category. In the categories of explicit websites, financial theft, identity theft, and stalking; we reject the null hypothesis because differences exist in age. For the categories of burglary and kidnapping; we retain the null hypothesis because no significant differences exist for age. With this investigation, we found that the age group above 26 have a higher concern for identity theft. While their younger counter parts tend to have a higher concern for financial theft, explicit websites, and stalking.

The participants allocated privacy leaks into three possible attack vectors (Table 1M). The location attack vector is used to find out where an individual lives and/or current location. The participants classified keys, passports, driver's license, social security cards, and personal letters as an item in location threat. The identity attack vector is used to exploit an individual's identity, even to the very intimate details. The participants classified credit/debit cards, children images, driver's license, social security cards, passwords, and personal letters as an item in identity threat. The asset attack vector is used gain access to an individual's possessions and valuables. The participants classified credit/debit card, keys, passports, driver's license, social security cards, passwords, and personal letters as an item in asset threat.

4. Data Collection via Web Crawling

To understand the pervasiveness of privacy leaks on SMNs, we ingested tweets from the Twitter API using the participant described key words. Each key term was obtained from survey participants who completed our surveys. In our initial survey we ask users to define categories of privacy leaks based on keywords. Next, we examined the keywords that are related to those categories.

We collected tweets and images from Twitter resulting in approximately 1.4 million tweets collected and 18,751 images. We collected data using notable keywords derived from the survey and participant's responses. This data was collected over a two-month time period.

4.1. Tweet Collection

We collected 1.4 million tweets to analyze the hashtags associated with them (Table 12). Twitter was searched with keywords derived from the privacy leak categories and the words given from participants. From the survey, the participants gave key words or phrases that are not currently used on Twitter (i.e., #stayoffthesidewalk). To find a correlation between related images and hashtag, we began a search of the words. The top hashtags from this search were *#racisttwitter, #gameawards, and #wikileaks*. Of the tweets collected, 109,994 tweets contained images.

Table 12. Results of keyword crawling on Twitter.

Keywords or Phrases	# of Images Collected
Credit card debit card	364,825
job offer job acceptance job letter	107,470
key house key car key	174,348
license licensed to drive driver's license	109,520
passport	183,048
password passwords	166,835
racist #racisttwitter	121,638
college acceptance college bound college letter	100,199
#wikileaks	137,208
Total	1,465,091

From the tweets collected, the most relevant results are from the college search which includes the key words college acceptance, college bound, and college letter. In this search we collected trending hashtags in reference to college searches: `#neumannscholarship`, `#nmsubound`, and `#hu24`. These hashtags are associated with college acceptances, scholarship acceptances and college letters.

4.2. Image Collection

Beyond collecting basic tweets, we searched for images associated with keywords and hashtags in the collected tweets. With this search we collected 18,751 images. The images collected were classified into three categories bases on risk: *severe*, *moderate*, and *no risk*. (1) *Severe* risk content contains images that have more than one attack vector (Section 3.4). These images include items that provide actual government issued identification (i.e., social security numbers, driver's license, etc.), items that can be used to identify a person and/or used for facial recognition (driver's license, identification cards), or items that contain insight to a person's location and/or place of residence. (2) *Moderate* risk content refers to images from the asset or identity attack vectors. This content includes images that feature items which can be used to identify a person and/or can be used for facial recognition. However, this content will not provide the user's location, place of residence, nor feature any of their government issued identification. (3) *No risk* content encompasses images that do not include any of the above items. The images were classified by three individuals and placed into categories based on the average agreement. Table 13 shows the total number of images in each category and its respective category after agreement and assignment.

Table 13. Risk Classification from keyword search with Twitter.

Category	# of Images Collected
Severe	160
Moderate	327
No risk	18,264

In each category, we found several privacy leaks. In *Severe*, the most prevalent images are car keys, license plates, and job offers. In *Moderate*, the most prevalent images are work identification, school information, promotion letters images. Table 14 shows the distribution of these images among the categories.

Table 14. Distribution of Content for Risk Categories. This table includes the keyword and the content frequency.

Category		Keyword (Count)	
Severe	(160)	Baby	71
		Driver's License	12
		Financial Document	2
		Hospital	54
		Job	4
		Keys	1
		License Plate	4
		Medication	10
		Medical Records	6
Moderate	(327)	Baby	45
		College Letter	6
		Driver's License	24
		Hospital	123
		Job Promotion	7
		Medical Information	52
		Medication	43
		Work Identification	12
		Workplace	15

The prevalence of images has a higher frequency for the terms `baby`, `hospital`, `medication`, and `medical records`. *Severe risk* includes images containing finances and keys unlike *Moderate risk* which contained more college and work-related images. When asking users to define privacy and identify threats, the participants did not identify hospitals, medical records, or medications as significant concerns. We find that these images trending on Twitter about medical information and hospitals have a higher chance of occurring in comparison to the other keywords.

5. Discussion

This study indicates similarities between users' definitions of privacy. We investigated this topic by creating subgroups using Kmeans and demographic identities to uncover definitions of privacy. It further demonstrates a correlation between dangers of SMNs and demographic subgroups. The analysis confirms that age groups have different levels of concern regarding explicit websites, financial theft, identity theft, and stalking. It also confirms that female and male participants have differences in the level of concern regarding burglary, explicit websites, and identity theft.

In line with our hypotheses, (1) the privacy perspective of a user can be subjective, and (2) several threats and dangers are heightened due to the accessibility of social media. We find that users construct their own meanings of privacy; however, the definitions show significant overlap. The threats on these platforms are heightened because of the accessibility of social media. From this analysis, we find that cyberbullying and explicit content rise to the forefront of concerns. The results do not fit the hypothesis that visual privacy leaks are common on Twitter; however, rare breaches in privacy may still be devastating. The reliability of the data from Twitter is limited by the keyword search terms used. As new trends arise and challenges appear, the keyword associations for the appropriate images change. To obtain accurate results, it would be important to survey recent keyword trends and challenges thoroughly. From the survey, 65% of participants state that they have seen privacy leaks on social media networks; however, we were unable to collect a corresponding amount of visual privacy leaks. In this study, we have collected words regarding trends and challenges in association with visual content. From this data, we see that the most accurate keyword search was regarding `college bound` and `college acceptance`.

These results build on existing evidence of previous work regarding the dangers of social media [1,3,21] and the subjectivity of privacy [3,5]. In this paper, we explore the user thoughts over different aspects of social media privacy and in particular visual privacy leaks. Users have solid personal notions of privacy, but they do not yet understand how privacy leaks can affect them. As new technologies arise, application developers must implement mitigation techniques that allow users to explore the trade-offs between privacy and sharing. This will be increasingly important for non-text and visual sharing methods across SMNs.

Author Contributions: Conceptualization, J.D. and C.G.; methodology, J.D., M.S. and C.G.; validation, J.D., M.S. and C.G.; formal analysis, J.D.; investigation, J.D., M.S. and C.G.; resources, C.G.; data curation, J.D.; writing–original draft preparation, J.D.; writing–review and editing, M.S. and C.G.; visualization, J.D.; supervision, C.G.; project administration, J.D. and C.G.; funding acquisition, J.D. All authors have read and agreed to the published version of the manuscript.

Funding: Jasmine DeHart is supported by the National GEM Consortium and the DoD SMART Scholarship for Service. Makya Stell is supported by the National Science Foundation Oklahoma Louis Stokes Alliance for Minority Participation. Financial support was provided from the University Libraries of the University of Oklahoma.

Acknowledgments: The authors would like to thank Kingsley Pinder Jr. of the University of Arkansas for providing statistical insight of the manuscript; and John Alberse for his initial work on the open source Twitter and Instagram scraper that supported this research: https://github.com/oudalab/viper_scraper.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

SMN Social Media Network
PL Privacy Leak

References

1. Gross, R.; Acquisti, A. Information revelation and privacy in online social networks. In Proceedings of the 2005 ACM workshop on Privacy in the electronic society, Alexandria, VA, USA, 7 November 2005; pp. 71–80.
2. DeHart, J.; Grant, C. Visual content privacy leaks on social media networks. *arXiv* **2018**, arXiv:1806.08471.
3. Rosenblum, D. What anyone can know: The privacy risks of social networking sites. *IEEE Secur. Priv.* **2007**, *5*, 40–49. [CrossRef]
4. Li, Y.; Vishwamitra, N.; Knijnenburg, B.P.; Hu, H.; Caine, K. Effectiveness and users' experience of obfuscation as a privacy-enhancing technology for sharing photos. *Proc. ACM Hum.-Comput. Interact.* **2017**, *1*, 67. [CrossRef]
5. Such, J.M.; Porter, J.; Preibusch, S.; Joinson, A. Photo privacy conflicts in social media: A large-scale empirical study. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017; pp. 3821–3832.
6. Zhong, H.; Squicciarini, A.; Miller, D. Toward automated multiparty privacy conflict detection. In Proceedings of the 27th ACM International Conference on Information and Knowledge Management, Lingotto, Italy, 22–26 October 2018; pp. 1811–1814.
7. Abdulhamid, S.M.; Ahmad, S.; Waziri, V.O.; Jibril, F.N. Privacy and national security issues in social networks: The challenges. *arXiv* **2014**, arXiv:1402.3301.
8. Squicciarini, A.C.; Caragea, C.; Balakavi, R. Analyzing images' privacy for the modern web. In Proceedings of the 25th ACM Conference on Hypertext and Social Media, Santiago, Chile, 1–4 September 2014; pp. 136–147.
9. Srivastava, A.; Geethakumari, G. Measuring privacy leaks in online social networks. In Proceedings of the 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Mysore, India, 22–25 August 2013; pp. 2095–2100.
10. Zerr, S.; Siersdorfer, S.; Hare, J.; Demidova, E. Privacy-aware image classification and search. In Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval, Portland, OR, USA, 12–16 August 2012; pp. 35–44.
11. Buschek, D.; Bader, M.; von Zezschwitz, E.; De Luca, A. Automatic privacy classification of personal photos. In Proceedings of the IFIP Conference on Human-Computer Interaction, Bamberg, Germany, 14–18 September 2015; pp. 428–435.
12. Tierney, M.; Spiro, I.; Bregler, C.; Subramanian, L. Cryptagram: Photo privacy for online social media. In Proceedings of the First ACM Conference on Online Social Networks, Boston, MA, USA, 7–8 October 2013; pp. 75–88.
13. Gurari, D.; Li, Q.; Lin, C.; Zhao, Y.; Guo, A.; Stangl, A.; Bigham, J.P. VizWiz-Priv: A Dataset for Recognizing the Presence and Purpose of Private Visual Information in Images Taken by Blind People. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–21 June 2019.
14. Kuang, Z.; Li, Z.; Lin, D.; Fan, J. Automatic Privacy Prediction to Accelerate Social Image Sharing. In Proceedings of the 2017 IEEE Third International Conference on Multimedia Big Data (BigMM), Laguna Hills, CA, USA, 19–21 April 2017; pp. 197–200.
15. Zerr, S.; Siersdorfer, S.; Hare, J. Picalert!: A system for privacy-aware image classification and retrieval. In Proceedings of the 21st ACM International Conference on Information and Knowledge Management, Maui, HI, USA, 29 October–2 November 2012; pp. 2710–2712.

16. Krishnamurthy, B.; Wills, C.E. Characterizing privacy in online social networks. In Proceedings of the First Workshop on Online Social Networks, Seattle, WA, USA, 17–22 August 2008; pp. 37–42.
17. Madejski, M.; Johnson, M.L.; Bellovin, S.M. *Failure of Online Social Network Privacy Settings;* Technical Report CUCS-010-11; Department of Computer Science, Columbia University: New York, NY, USA, July 2011.
18. Boyd, D. *It's Complicated: The Social Lives of Networked Teens*; Yale University Press: New Haven, CT, USA, 2014.
19. Boyd, D.; Marwick, A.E. Social Privacy in Networked Publics: Teens' Attitudes, Practices, and Strategies. 2011. Available online: https://osf.io/2gec4/(accessed on 20 Jaunary 2020)
20. Knijnenburg, B.P. Privacy? I Can't Even! making a case for user-tailored privacy. *IEEE Secur. Priv.* **2017**, *15*, 62–67. [CrossRef]
21. Veiga, M.H.; Eickhoff, C. Privacy leakage through innocent content sharing in online social networks. *arXiv* **2016**, arXiv:1607.02714.
22. Ilia, P.; Polakis, I.; Athanasopoulos, E.; Maggi, F.; Ioannidis, S. Face/Off: Preventing Privacy Leakage From Photos in Social Networks. In Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security, Denver, CO, USA, 12–16 October 2015; pp. 781–792. doi:10.1145/2810103.2813603. [CrossRef]
23. Zezschwitz, E.; Ebbinghaus, S.; Hussmann, H.; De Luca, A. You Can't Watch This!: Privacy-Respectful Photo Browsing on Smartphones. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, 7–12 May 2016; pp. 4320–4324. doi:10.1145/2858036.2858120. [CrossRef]
24. Loukides, G.; Gkoulalas-Divanis, A. Privacy challenges and solutions in the social web. *XRDS* **2009**, *16*, 14–18. [CrossRef]
25. Padilla-López, J.R.; Chaaraoui, A.A.; Flórez-Revuelta, F. Visual privacy protection methods: A survey. *Expert Syst. Appl.* **2015**, *42*, 4177–4195. [CrossRef]
26. Mazzia, A.; LeFevre, K.; Adar, E. The PViz comprehension tool for social network privacy settings. In Proceedings of the Eighth Symposium on Usable Privacy and Security, Washington, DC, USA, 11–13 July 2012; p. 13.
27. Li, X.; Li, D.; Yang, Z.; Chen, W. A patch-based saliency detection method for assessing the visual privacy levels of objects in photos. *IEEE Access* **2017**, *5*, 24332–24343. [CrossRef]
28. Orekondy, T.; Fritz, M.; Schiele, B. Connecting pixels to privacy and utility: Automatic redaction of private information in images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 8466–8475.
29. Li, Y.; Vishwamitra, N.; Knijnenburg, B.P.; Hu, H.; Caine, K. Blur vs. block: Investigating the effectiveness of privacy-enhancing obfuscation for images. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 21–26 July 2017; pp. 1343–1351.
30. Zhao, Q.A.; Stasko, J.T. *The Awareness-Privacy Tradeoff in Video Supported Informal Awareness: A Study of Image-Filtering based Techniques*; Technical Report; Georgia Institute of Technology: Atlanta, GA, USA, 1998.
31. Boult, T.E. PICO: Privacy through invertible cryptographic obscuration. In Proceedings of the Computer Vision for Interactive and Intelligent Environment (CVIIE'05), Lexington, KY, USA, 17–18 November 2005; pp. 27–38.
32. Johnson, M.; Egelman, S.; Bellovin, S.M. Facebook and privacy: It's complicated. In Proceedings of the Eighth Symposium on Usable Privacy and Security, Washington, DC, USA, 11–13 July 2012; p. 9.
33. Schaeffer, N.C.; Presser, S. The science of asking questions. *Annu. Rev. Sociol.* **2003**, *29*, 65–88. [CrossRef]
34. Kalton, G.; Kasprzyk, D. The treatment of missing survey data. *Surv. Methodol.* **1986**, *12*, 1–16.
35. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
36. Loper, E.; Bird, S. NLTK: The natural language toolkit. *arXiv* **2002**, arXiv:cs/0205028v1.
37. Sarkar, D. *Text Analytics with Python: A Practitioner's Guide to Natural Language Processing*; APress: Berkeley, CA, USA, 2019.
38. Miller, G.A. *WordNet: An Electronic Lexical Database*; MIT Press: Cambridge, MA, USA, 1998.

39. Bengfort, B.; Danielsen, N.; Bilbro, R.; Gray, L.; McIntyre, K.; Richardson, G.; Miller, T.; Mayfield, G.; Schafer, P.; Keung, J. Yellowbrick, 2018. Available online: https://zenodo.org/record/1206264#.XiVI1CMRVPY (accessed on 20 Janurary 2020).
40. Caliński, T.; Harabasz, J. A dendrite method for cluster analysis. *Commun. Stat.-Theory Methods* **1974**, *3*, 1–27. [CrossRef]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).