Awareness about photos on the Web and how privacy-privacy-tradeoffs could help

Benjamin Henne and Matthew Smith

Distributed Computing & Security Group, Leibniz Universität Hannover
{henne,smith}@dcsec.uni-hannover.de

Abstract. Many privacy issues concerning photos on the Web and particularly the social Web have been discussed in the past. However, much of this discussion is based on anecdotal evidence and has focused on media uploaded by users themselves. We present the results of a survey conducted with 414 participants that studies user awareness of privacy issues concerning the sharing of media including media shared by others. We additionally investigate the current perception of metadata privacy, since metadata can amplify threats posed by photos on the Web, for instance by tagging people or linking photos to locations. Furthermore, we present how this metadata can be used to help to protect private information and discuss the concept of a *privacy-privacy-tradeoff* and how this can be used to enable people to discover photos relevant to them and therefore regain control of their media privacy.

Keywords: privacy, awareness, social media, photo sharing, metadata, privacy-privacy-tradeoff

1 Introduction

A multitude of privacy issues with online photos have been discussed in the past years [1, 3, 2, 6, 7, 10, 12], with photography in general being a point of contention for privacy issues for over a century [13]. The thought of being depicted in a photo somewhere on the Web is already a privacy concern for some people: Even a picture of someone at a perfectly harmless location may raise objections. People feel even more threatened by pictures showing them in embarrassing situations, doing socially questionable things, or at a place or with someone they would rather deny having been with. Research has shown that people feel their privacy threatened by photos taken by nearly any other person, no matter if they are from people outside [1] or inside their social circle, including friends and family [3]. Furthermore, media content may not just harm personal privacy, but can also cause immediate effects, since employers, insurance companies and banks use such information to gather knowledge about employees or clients. An increasing number of people have become cautious about sharing personal data in social network services (SNS). Yet, SNS users still create threats to their own privacy by accidentally disclosing compromising pictures of themselves to the public. Access control facilities offered by SNS help people keeping their media private up to a certain degree, though usability or comprehension issues often complicate the effective deployment of privacy settings [6, 10]. Aside from these relatively obvious problems, other threats have not yet received sufficient attention: Shared photos not only affect the uploaders' privacy, but the privacy of all persons visible in the photo. Threats posed by such photos are particular insidious, since the potential victims are not involved in the uploading process and thus cannot take any preemptive measures against being depicted online. While for instance tagging people in photos can be prevented in current SNS, there currently are no countermeasures to the upload itself except legal actions or demanding that the media be taken offline again. Since online sharing of media cannot be simply prohibited, raising awareness about shared media on the Web is the key issue to address privacy concerns arising from the increased use of the social Web.

For privacy threats of shared media to take effect, two requirements have to be fulfilled: To cause harm, media needs to be able to be associated to a person. In addition, the media in question must contain objectionable content for that person. The association and the content can both be either non-technical – i. e. only recognizable by humans – or technical – i. e. content actively linking to a personal profile, or metadata containing a compromising time or location. In this context, the metadata plays an integral role: It stores additional information besides the picture itself and is easily machine-readable. The use of contemporary cameras and especially smartphones amplifies the privacy threat posed by shared media: Current cameras are capable of gathering location information via GPS or Wi-Fi-tracking and automatically embed it into photos. Latest applications additionally integrate facial recognition functions that aim to automatically tag individuals in photos. Modern devices ease the annotation of shared media with information that may give rise to privacy concerns.

In this paper, we focus on threats posed by photos shared by others and analyze their relevance by presenting results from an online survey. We discuss *awareness* of media sharing as a key issue in Sect. 2 and examine the current importance of metadata privacy in Sect. 3. Our analysis is based on the results of an online survey with 414 participants, showing that while most participants are aware of possible privacy threats, they also see a need for a better chance to effectively control which photos depicting them are shared. Using a prototypical system, designed to raise awareness about media sharing, we discuss *privacy-privacy-tradeoffs* that disclose certain private information to a social network privacy service to regain control over more important private information in Sect. 4. Finally, we conclude this paper in Sect. 5

1.1 Survey Design and Participants

The remainder of this paper describes the results of an online survey. We will introduce the individual parts of our survey in combination with the respective results in separate sections.

1,418 members of a university-related mailing list were invited to participate in the survey. The invitation asked for participation in a survey on privacy issues of media sharing. While explicitly mentioning privacy has possibly caused selection bias, we intended to recruit users interested in this topic to investigate a best-case scenario. As an incentive for participation, we offered participants an option to enter a raffle for two $ 60 vouchers from Amazon.

We received 414 complete and valid answers. 53.9 % of our participants were male and 46.1% female. About 25 % of the participants already had at least one university degree. The average age of participants was 23±4 years. 22.2 % indicated a high or very high technical expertise. According to Westin's privacy segmentation index [9], 91.8 % of the participants were classified as privacy pragmatists, 6.0 % as fundamentalists and 2.2 % as unconcerned. Thus most of our participants handle their online privacy pragmatically depending on the situation, indicating that most of them would therefore not simply be uninterested in privacy controls nor demand them regardless of the real threat, but present differentiated opinions on the topic at hand.

Normality testing indicated significant deviations from the normal distribution (Kolmogorov-Smirnov) for most rating variables as expected, which is why we employ non-parametric test measures to discuss our results.

2 Online photo awareness

When reports about employers and banks using social media to gain knowledge about their employees and customers increased, privacy problems of shared media began to catch the public's i.e. the media's attention. However, the extent to which this attention actually translates into user actions or awareness is unknown. Thus, one goal of our survey was to learn about the extent of the awareness users currently have concerning online photos they might be depicted in. Are people really aware of the threat posed by pictures shared by others and their possible impact? Do users realize that pictures they are not tagged in also cause privacy issues?

2.1 Linking media to people

Most of the popular SNS like Facebook or Google+ and media-sharing sites like Flickr allow users to tag objects and people in the media they upload. Media can be commented on, annotated with keywords, or directly linked to a person. The direct link between profiles and photos thereby was initially met with a great outcry of privacy concerns. Such links simplify finding pictures of people beyond the content they consciously share in their profiles. For this reason, current SNS allow their users to either completely forbid others to link them in shared media or to approve links before they become visible to the public. However, such links also have a positive side: When tagged in a photo, users usually receive notifications about the link and consequently about the photo that might raise privacy concerns. Based on this notification, users can check the picture and possibly have unwanted content removed or access restricted [2, 12]. One goal of our survey was to find out to what extent users are aware of the positive effect of such tags. To gather reasons for tagging others in photos, we asked our participants how frequently they tag someone for specific reasons, using a 7-point scale from *not at all* to *very often* (cf. Fig. 1). 30 % of the participants stated that they never tag people in their photos just to notify the tagged user. The remaining 70 % rated this item with a mean rating of 5.34 ($sd = 1.42$), indicating that this is a valid reason for tagging for most users. Likewise, 54.8 % of all participants stated that they never tag someone in a photo to make other people aware of this photo. For the remaining participants, this also appears to be a less important reason with a mean rating of 3.73 ($sd = 1.55$). The numbers indicate that our participants rather tag their friends to notify them about their presence in pictures than to distribute their photos to others.

To assess the perception of being tagged, we asked participants to rate their feelings on the effects of being tagged in photos, on a 7-point scale from (1) *like it very much* to (4) *neutral* and (7) *dislike it very much* (cf. Fig. 2). The results indicate that becoming aware of photos of oneself is not the most important effect of tagging for our participants. This mirrors the Web 2.0 spirit: Most participants state they significantly prefer (Wilcoxon test, $Z = -3.41$, $p = .001$) finding photos of others with a mean of 3.51 ($sd = 1.37$) to finding photos of themselves with a mean value of 3.79 ($sd = 1.8$). However, participants also stated that they rather dislike that others can find their photos because of tags with a mean of 4.77, $sd = 1.55$. These results confirm typical assumptions about social sharing: SNS users like to be able to easily find photos of others while they dislike others being able to easily find pictures of themselves. Feelings about being informed about pictures of oneself tend to be more neutral which indicates that they see only little to no awareness benefits in being tagged. Therefore, people rather tag to follow the Web 2.0 spirit than for privacy reasons.

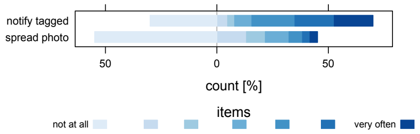


Fig. 1: (q13) How frequently do you tag for these reasons?

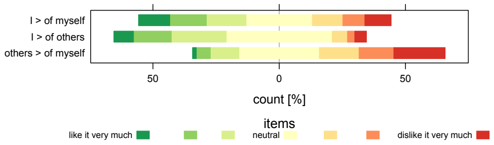


Fig. 2: (q14) Rate the effect of people tags: Who finds photos of whom.

Limits of Tagging The positive side of tagging people with profile links should not be underestimated. Indeed, such links are the only solution available in current SNS to notify people of photos of themselves besides any out-of-band communication between photographers and depicted people. The links offer a certain level of awareness, but one has to keep in mind that they are limited to photos of friends or indirect friends, because outside of these circles, access control, missing social connections and the lack of interest prevent notification.

To judge the seriousness of this deficit, we tried to assess the origin of privacy issues from the users' viewpoint. We asked our participants to rate the extent of a possible privacy violation by photos shared by different groups of people on a 7-point scale from *very low* to *very high* (cf. Fig. 3). Most respondents rated any violation higher than *very low*: Only 1.4 % of the participants rate a possible violation to be *very low* regardless of who shared the photo. The participants rate the violation level of photos shared by friends to be the lowest with a mean of 3.64 ($sd = 1.85$). Photos shared by friends of friends were rated to caused a medium level of violation on average (4.69, $sd = 1.66$) and the media shared by strangers was rated highest with an average rating of 5.23 ($sd = 1.95$). The differences in mean ratings are significant (Friedman test, $\chi^2_2 = 185.41$, $p < .001$)). Additionally, 47 % of participants rated privacy violations by strangers' photos consistently higher than those caused by direct and indirect friends. We conclude that participants perceive threats caused by strangers' photos to be worse than other privacy violations. In contrast to photos posted by direct or indirect friends, photos uploaded by strangers are neither tagged with entailing links, nor do they result in any notification. Therefore, profile links as a privacy feature have serious deficits because they do not cover this scenario.

The results on the extent of a possible privacy violation suggest that participants seem to believe that others do not comply with a "moral obligation", as described in [2], even though most people declare they think about other users' privacy when sharing media: We asked the participants to rate the influence of threats to others and threats to themselves as decision-making criteria for sharing a photo using a 7-point scale from *not at all* to *very much*. Only 2 % of our participants answered that they do not think about threats to others at all when sharing photos on the Web. Within the remaining participants, about 61 % rate threats to others and threats to themselves with the same value. Interestingly, 6.6 % of participants rated threats to others as a sharing criterion higher than threats to themselves.

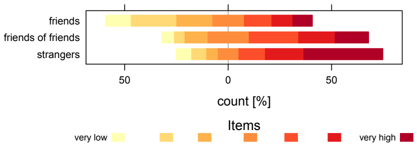


Fig. 3: (q17) Estimate a possible privacy violation of photos shared by ...

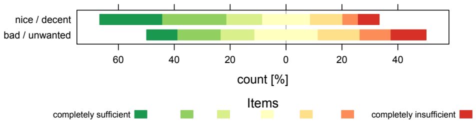


Fig. 4: (q19) How well do you feel informed about all photos of yourself?

2.2 Awareness today

In the context of shared photo awareness, we also need to consider photos that contain identifying information but are not linked to profiles. Compared to photos directly linked to a person's profile and therefore immediately discoverable photos, unlinked photos are more critical: A tag that contains identifying information is attached to a photo, but no link to a person's profile is made. This can technically be implemented in a multitude of ways, ranging from mentioning a name in the headline or a comment in a SNS, to metadata that describes depicted people stored in the image file. While the potential damage of course is smaller, the threat can remain hidden far longer, because no automated mechanism helps to find this image. Currently, the only way to combat this threat is for the concerned person to pro-actively crawl the Web in search of such photos. We asked the participants of our survey to estimate the risk of someone finding a photo of them anytime in the future that this someone should not have seen. They assessed the likelihood of three scenarios of how they could be associated to a picture using a 7-point scale from *very low* to *very high*. While 24 % of the participants rated the risk of someone finding a photo that was previously linked to a SNS profile to be *very low*, only 11 % rated that risk to be *very low* if the photo contained personal references in the metadata or if they are only visible in a photo. This is an obvious result, since the tagged person is notified about photos linked to his or her SNS profile and can therefore be removed if necessary. Users see more future threats in unknown photos with personal references than in those they are only visible in: In the former case, 45 % of the participants rated the risk to be in the worst three elements of the scale, while only 35 % did so in the latter case. This difference is statistically significant (McNemar test, $\chi^2 = 10.32$, $p = .001$). This indicates that participants believed photos with actual personal references in the metadata to be more easily discoverable, for instance using a search engine, than those they are only visible in.

Finally, our study addressed to which extent users are satisfied with currently available options to become aware of photos of themselves. Thus we first queried respondents how they are currently becoming aware of photos of themselves, using a multiple-choice question. 75 % stated that they automatically get notifications by email when tagged in a photo (94 % of these were Facebook users); 52 % of the participants stated that they get to know about photos of themselves by chance; 39 % of them hear about photos of themselves in conversations and 30 % in friends' messages; 18 % actively look for photos; 4.6% get informed by messages from non-friends; and 3.4 % stated that they do not become aware of photos of themselves at all. Automated notifications are only possible in the case of profile-linked tags in current SNS. It is important to note that all the means of becoming aware of photos presented to the participants are not applicable in the case of non-linked tagging or missing tags.

Furthermore, we asked our participants to rate how well they feel informed about several types of photos of themselves on the Web, on a 7-point scale from *completely sufficient* to *completely insufficient* (cf. Fig. 4). Concerning decent photos, their perceived level of available information was a little better than neutral (3.2, $sd = 1.85$) and concerning objectionable photos, their average perception was exactly neutral (4.0, $sd = 1.85$). In detail, 22 % stated that their level of information is *completely sufficient* concerning decent photos of themselves while 25 % chose a level from worse than neutral to *completely insufficient*. In contrast, only 11 % state a level of *completely sufficient* concerning objectionable photos, while 39 % of the participants assert that their level of information about bad photos of themselves was worse than neutral to *completely insufficient*. Again, the differences between these values are statistically significant (McNemar test, $\chi^2 = 50.77$, $p < .001$). We finally asked the survey participants whether they would like to use a service that helps them finding relevant photos, which requires its users to manually screen potential photos. 53.1 % of them answered with a clear yes and 41.8 % were interested in using such a service. Only 3.6 % argued that the effort of screening would overbalance benefits. Others called on the uploaders' moral obligation or denied being depicted online at all.

2.3 Summary

Becoming aware of uploaded photos that a user is visible in is the key issue for combating privacy threats created by online media. Popular services allow their users to tag people in shared media. Mostly, tagging creates a link to the profile of that person. The tagged person is notified and can take action. Respondents did not see very much awareness benefits in such linked tags. Even if these features were fully appreciated, the privacy benefit is limited to photos of direct and indirect friends within their circles of friends. Photos shared by other people and outside of service boundaries cannot benefit from such mechanisms. Yet, users rate exactly those photos to pose the biggest threat for a possible privacy violation. In order to become aware of all relevant photos, photos with non-linked personal references as well as photos without any reference to a person have to be considered. For these types of photos, there currently are no effective possibilities to increase awareness besides manually crawling the web. When asked in which way and how well they are informed about photos they may be depicted in, participants' answers confirm that improvements are needed in the area of online media awareness and privacy. Although prior research has shown that users tend to spend little effort in privacy settings, nearly all of our participants are willing to invest at least some time in screening potential photos, if this offers a chance of being informed about potential privacy violations. A participant even offered to pay a one-time fee for such a service. The challenge is to implement a service that caters for the users' privacy needs and does not create new threats to the users' privacy at the same time.

3 Photo metadata

Metadata is used to add valuable context information to images and helps to order, categorize and even find images in huge media libraries or by search engines. Metadata handling is integrated in nearly every image processing software and digital camera today. Modern devices automatically save several pieces of metadata with each photo, including the current date, time and GPS coordinates and even the camera owner's name. Additionally, an increasing number of applications support semi-automatic tagging of photos with textual location information based on reverse geocoding or tagging people within images. Besides the image itself, metadata of that image can also harm the privacy of a person.

Metadata can link people to images, for instance by storing names of photographers or depicted people. It can also contain information about the time or location of taking a photo that can create or amplify privacy threats.

3.1 Knowledge and Nescience of users

Regarding privacy concerns of metadata, it is important to differentiate between data that is loosely attached to media for instance in the UI of a website and data stored directly in an image file. While the former is typically only accessible within the service and protected by access control, the latter is spread with the image and is generally as persistent as the image itself. Since only a part of all users (61 % of our survey participants) knows the term metadata, we can assume that even less know the difference between these two kinds of metadata storage and their respective implications. In our survey and consequently in this paper, we therefore only use the abstract term *metadata* to refer to additional information of photos, such as time, headline, or tagged people, regardless of how it is stored. During our survey, however, one participant commented: *"No difference was made between embedded metadata and metadata stored externally, that makes a world of differences when spreading a photo"*.

To estimate how users handle metadata, we asked the 253 participants that indicated to know what metadata is to agree or disagree to a set of statements. About 25 % stated that they do not add additional metadata to photos. However, some users might nonetheless do so in SNS, without knowing the term. About 6 % of the 253 participants stated that they remove all metadata from images before they share them on the Web and an additional 35 % stated that they remove parts of the metadata. 2 % said that the online services they use remove metadata on upload. Our participants also admitted to nescience: 58 % answered that they do not know what their SNS or media-sharing sites do with photo metadata. 29 % state that they do not know which additional information is contained in the photos they share. About 27 % of the 253 participants state that they do not think about metadata at all when sharing images on the Web. In contrast, 9 % of the 253 state that metadata is an important part of sharing.

3.2 Private metadata

Most research and online services consider only few pieces of metadata of photos as confidential or related to privacy. We already discussed the practice of tagging people in current SNS in Sect. 2.1. Beyond these kind of tags, the location of a person or the location a photo was taken is most discussed in other papers and one of the few that is also specifically addressed in current services on the Web. The general term *location* mostly refers to GPS-based WGS-84 coordinates. Other location information, such as the name of a city or a point of interest, or an address where a photo has been shot is often not considered. But, since geocoding has become cheap and easy, coordinates and textual location information have to be dealt with equally. However, this is generally not the case: For instance, at the web-based photo sharing feature of Apple's iCloud Photo Stream, WGS-84 coordinates are removed, but any other location information is retained. This shows that we have to extend the current notion of privacy-related information in media metadata. Additional meta-information will raise privacy concerns in the future: The number of cameras that write a camera identifier into photos rises. These ids may not be as unique as a smartphone's IMEI, but still can be used to re-identify a camera owner. Additional concerns may arise from new metadata standards that allow tagging people with names and bounding boxes directly within image files. Up to now, this was only possible and known in the context of online SNS, but applications like Google Picasa or Windows Live Photo Gallery as well as libraries like exiv2 implement these standards today.

With our survey, we aimed to asses the users' view of the privacy implications of different pieces of metadata and how severe they estimate a possible privacy violation caused by the disclosure of such data to be. We asked our participants to rate the possible privacy impact of adding such metadata to media depicting others on a 7-point scale from *very low* to *very high.* Additionally, we asked them to rate the privacy impact of metadata added by others to media depicting themselves using the same scale. Table 1 in the appendix shows details of both.

Comparing the different kinds of metadata, headline, description, and tags are perceived to have the least impact with a mean rating of 3 ($sd = 1.7$) across both questions on the 7-point scale from *very low* to *very high.* The creation date and time of a photo (3.6, $sd = 1.7$), the photographers' name (3.4, $sd = 1.8$), and also broad location information, such as the city or region where a photo was taken, (3.9, $sd = 1.7$) are considered to have slightly less than medium impact. In contrast, the names of depicted people (4.9, $sd = 1.8$) and exact location information, such as GPS-based coordinates or a postal address, (5.2, $sd = 1.7$) are perceived as having a higher impact.

People It is interesting to note the difference between names of depicted people and the photographer's name, since both indicate persons related to a photo. Finding the name of camera owners in photos also implicates their presence at that time and place, as long as the camera or smartphone was not lent to others.

Location Our participants rated location as the kind of metadata with the highest privacy impact. However, recent related work voiced doubts that location still raises much concerns with today's smartphone users, compared to the beginning of the mobile era. For instance, in the "very-upset-ranking" of Porter Felt et al. [11], the participants ranked location-related risks in the bottom half and the actual location was ranked second-lowest out of eleven data types. Fisher et al. [4] show that iOS users seem to pay attention to which apps they allow to use location and do not disable the feature in general. Krumm [8] summarizes different results, showing that people do not seem to care about location privacy. So why does our data differ?

The prior work mainly deals with location in the context of location-based services, the pro-active publication of locations, or the misuse of location permissions by smartphone applications. In all these cases, location and where that information is stored may be less tangible to people. Our survey has been conducted in the context of photo sharing on the Web. In this case, location is at least connected to a picture and eventually to additional meta-information. A photo may be seen to last longer in the public. Photos are indeed not touchable, but much more concrete in the participants' mind than a single location recorded by an abstract service. Caused by the higher familiarity with photos, location data in pictures may raise more privacy concerns than in other contexts. To the best of our knowledge, no previous work compared users' feelings about location data in different contexts. We believe that there are different aspects that may explain the differences of results, which we will investigate in future research.

To examine the influence of the audience when disclosing location data, we asked our participants to rate how they felt if people get to see a photo of them that includes location information, using a 7-point scale from *very unconcerned* to *very concerned* with 4 as *neutral*. When sharing a photo with location data with friends (2.24, $sd = 1.5$) or friends of friends (3.51, $sd = 1.7$), participants state to be more or less unconcerned and more concerned in the case of other people (5.16, $sd = 1.8$). However, when it comes to servers, for instance the service that hosts the photo (5.23, $sd = 1.8$) or a privacy service that searches for depictions (5.28, $sd = 1.9$), people state to be even more concerned, which is contrary to the results of Felt et al [11]. The scenario of a privacy service will be discussed in the next section.

3.3 Summary

In this section, we discussed the role of metadata on the respondents' perception of privacy. While users of SNS know that they can add comments, locations or people tags to images on the Web, the general idea of metadata seems still to be less known to users. Only few people know about metadata that is stored directly in photos. Consequently, few people know about privacy-related data that might already be contained in images before they are uploaded to the Web. Even if they do know about the data, we have to ensure that people are aware of the contained information: For instance, the photographer (and therefore also the likely owner of the camera) was also present when a photo was taken. There also is little difference between GPS-based coordinates and postal addresses due to geocoding. Additionally, we presented results that are in conflict with previous investigations on sharing location data. Further research is needed to examine if and why there is a difference in perception.

In general, the potentially important role of metadata has to be made clear to users who are concerned about their privacy. Additionally, many processes that handle metadata are not forthcoming about which kind of information they handle in which way. For instance, it needs to be clear that if location information is removed, all kinds of location information are removed, including coarse locations or geocoded information. Moreover, there is little awareness of which information is stored in images by software and cameras: A single option in Google Picasa decides if people tags are stored in its database or are written into the files. Most users are not aware of the consequences of this choice. Canon cameras can also write the camera owner's name into the metadata, which also has possible privacy implications. Regarding photos and metadata, transparency and usable privacy mechanisms are needed to lower privacy threats as well as the danger of nescience.

4 A Privacy-Privacy-Tradeoff

Traditional privacy research aims to preserve users' privacy at all cost. We propose that this is not necessary and desirable in real world systems, especially in the social Web that is built around contributing and sharing. Users decide which aspects of their personal data they disclose to others. The Web 2.0 spirit shows that many people are happy about sharing things as long as they benefit from it or appear in a positive light.

Photo metadata can contain various information from technical details about the camera used to context information about the who, when, where and what of a photo. It can be used to preserve the non-visual context of a photo or it can be used to order a huge collection of images. In addition to these traditional use cases, we propose to also use some pieces of metadata for security and privacy purposes. A somewhat related intention can be found in the work of Klemperer et al. [7]: they derive access control rules for images from their keywords. In contrast, we propose to leverage image metadata to protect the privacy of the people affected by an image by allowing people to become aware of it [5].

The following scenario illustrates how metadata can be used to this end: *The service* S *assists users in finding media that might be relevant to them.* S *may be implemented as a value-added service within a SNS. Users of* S *can define private locations on a map or update their current location at the service through "checking in" or similar approaches. Based on co-location checks of users' private areas and the location information of photos uploaded to the SNS via* S*, the service notifies users who may be depicted in a photo based on respective locations. Additionally, SNS profile pictures can be used as training data for face recognition to improve results.* In this example, the service S leverages location metadata and profile pictures of users to make them aware of photos, so that they can protect themselves against unwanted publication.

Most of the necessary metadata is private to the affected people. If we want to use this information for privacy protection, we face some fundamental questions about the privacy of information that potential users have to decide for themselves: Firstly, is all information that at least some people regard as private also private to the user? Secondly, is all information that the user regards as private equally private in the way that the number or groups of people or services, which he allows to get to know the information, are identical? Otherwise, what information would the user share with which people or services? This creates *privacy levels* containing information that is similarly relevant to users' privacy. While most privacy fundamentalists and privacy unconcerned might have exactly one level of privacy, the number of privacy pragmatists in our study was found to be considerably higher. We therefore suggest building privacy mechanisms based on privacy levels.

To confirm the usefulness of this suggestion, we asked the participants of our survey to what extent they agree to the existence of privacy levels as defined above (cf. Fig. 5). On a 7-point scale from (1) *strongly agree* to (7) *strongly disagree* with 4 as *neutral*, the participants provided a mean agreement of 2.63 ($sd = 1.7$). 13.5 % of the participants indicated disagreement (5.6 % strongly disagree), 12.3 % were neutral, and 74.2 % indicated agreement (33.1 % strongly agree). We found no relation between the answers and the Westin segmentation. According to these results, participants generally feel that there are different levels of privacy, while about one third strongly supported this notion.

If privacy levels exist, we can take advantage of them: A privacy service may leverage some information that is less private to a person to secure other information that is more private to that person. We call this a *privacy-privacy-tradeoff*: If privacy levels exist in a system that builds on (public but also private) information – like current SNS and other social sites – users can choose to disclose less private information to secure other, more private information.

In our survey, we validated this idea by asking if our participants agree to this kind of tradeoff. We accompanied this question with a short description of the above scenario where they could choose to "reveal their location to a service to get notified about photos in which they might be depicted". We asked the participants to rate their agreement using a 7-point scale from (1) *strongly agree* to (7) *strongly disagree* with 4 as *neutral*, concerning if they, in general, would disclose some private information to secure other more private information (cf. Fig. 5). Participants gave a mean agreement of 3.49 ($sd = 1.7$). While 22.7 % of the participants indicated disagreement to this privacy-privacy-tradeoff and 24.5 % of them answered neutrally, 52.7 % of the 414 participants voiced their agreement to the tradeoff.

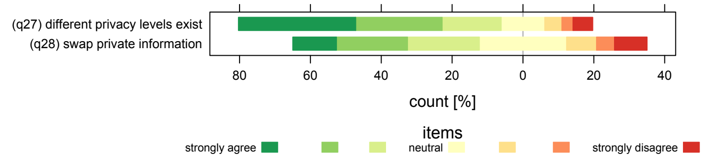


Fig. 5: (q27) Do privacy levels exist? (q28) Would you in general share some private information to secure other more private information?

We also asked participants which information they would trade for being notified about photos they might be depicted in that would otherwise be hard to find or even not accessible. We used the same 7-point scale as above. As shown in Fig. 6, participants mostly agreed using their existing profile pictures (2.97, $sd = 1.8$) to get notifications about photos. These, for instance, could be used to train face recognition. The second information the participants would be willing to disclose to some extent is pre-defined locations (4.0, $sd = 1.9$) that could be used for co-location checks to find photos at static places like home. On average, participants were reluctant to provide additional profile photos that comply with guidelines, like for a passport (4.5, *sd* = 1.9), which would be more suitable to train face recognition. They also indicated slight disagreement on providing their SNS list of friends (4.5, *sd* = 1.8) that could be used to specifically monitor friends' photos. Participants disagreed most to the use of a location-based service to constantly disclose their current location to the photo-service (5.4, *sd* = 1.9). This kind of data would obviously allow for the most effective co-location checks with photos.

Altogether, besides the use of existing profile photos, our participants on average disagree to trade private information when it comes to implementing a real tradeoff. However, if we consider respondents that indicated agreement (including those that would not mind) on trading private information as potential users, we get the following percentages: 67.5 % (82.6 %) might allow the use of existing profile photos, while 35 % (51.7 %) would provide extra photos complying to guidelines. 30.9 % (50.5 %) would allow to use their friends list. 39.6 % (61.8 %) would define private locations on a map and 19.1 % (31.2 %) would use the location-based service to update their current location.

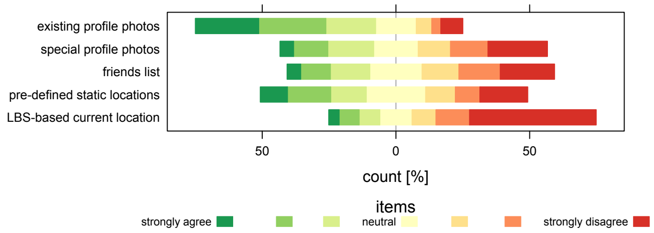


Fig. 6: (q30) What information would you disclose to a photo-sharing service to find photos of yourself that you otherwise would not be able to find or access?

Additionally, we added three questions to our survey that describe specific tradeoff situations to investigate agreement using the same 7-point scale:

q31: *"I am less upset if someone finds out where I have been than if that person gets to see private photos of myself."* — Participants somewhat agreed on average (3.0, *sd* = 1.7); 66.2 % indicated agreement and 15.7 % answered neutrally.

q32: *"I am less upset if my SNS knows where I have been than if my friends and strangers gets to see unwanted photos of myself."*— Again, Participants somewhat agreed on average (3.3, *sd* = 1.8); 60.4 % indicated agreement and 16.2 % answered neutrally.

q33: *"If there is a privacy service that notifies me about unwanted photos in which I am depicted but needs to know where I have been, I would use it. I would tell it where I have been to get to see possible photos of myself."*— Participants provided an average agreement of 3.7 (*sd* = 1.8); 53.2 % indicated agreement and 16.1% answered neutrally.

While all answers differ significantly (Friedman test, $\chi^2_2 = 44.46, p < .001$), answers to the first two questions appear to be weakly correlated (Spearman's $\rho_{1+2} = 0.596, p < .001$) and answers to the third appear to be independent ($\rho_{2+3} = 0.236, \rho_{1+3} = 0.226, p < .001$). Hence, respondents generally indicated agreement to scenarios stating a direct privacy tradeoff, but were more reluctant about disclosing information to a service to get notified of possible photos of them. This may imply that participants do see a privacy tradeoff but are not quite willing to trust another service to keep even less sensitive data private.

4.1 Summary

Our hypothesis that not all private information is equally private to people but is structured into several privacy levels was confirmed by our results; only 5.6 % of participants strongly disagreed. Given that privacy levels exist, we suggested leveraging this circumstance: We proposed to use less private information to secure information that is more private to users. We asked participants to what extent they would agree to a privacy-privacy-tradeoff. In general, 77.2 % agreed or were neutral towards this proposal. However, when participants were asked about a real implementation instead of a general idea, less people agreed to trade private information. While participants agreed to disclose SNS profile pictures for notifications about photos, they were generally more reluctant towards other information, especially location. However, a considerable amount of participants was ready to trade private information and may therefore be considered to be potential users of tradeoff-based privacy mechanisms.

The results of the explicit tradeoff situations confirms this impression: 60.4 % of the participants agreed that they prefer their SNS knowing where they were rather than other people, from inside or outside of their social circle, seeing unwanted pictures. Furthermore, 53.2 % of participants directly agreed to using a service offering this privacy-privacy-tradeoff.

5 Conclusion and Future Work

The results of our survey give a detailed account of the privacy preferences users have concerning the sharing of photos and their perceptions about linking photos to people. The assessment of the users' current degree of awareness shows that improvements are needed in the area of online media awareness and that users are willing to accept additional effort to gain improved awareness.

We investigated the role of metadata and differences in the perceived privacy impact of the unwanted disclosure of specific metadata: Personal references and location data raise most concerns for the users. These findings partly contradict the current views in related work, which state that users are not particularly concerned about location information. We therefore suggest that location privacy needs to be reconsidered in general and especially in the context of shared media, since our survey indicates that there are strong concerns about disclosing this kind of location information.

We also discussed the general idea of a privacy-privacy-tradeoff. Our survey shows that such a tradeoff would be appreciated by a fair number of users. The willingness to use a tradeoff-based service depends on the offered benefits: When participants were asked if they wanted to become more aware of photos of themselves, most agreed. However, the disclosure of meta-information and private data was also considered an issue. Finding the right balance in this tradeoff is an interesting topic of future research. We hope the results presented in this paper can serve as a basis for designing privacy-privacy-tradeoff-based services that take the users' perceptions into account.

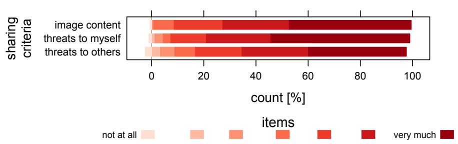
References

1. Ahern, S., Eckles, D., Good, N., King, S., Naaman, M., Nair, R.: Over-exposed? Privacy patterns and considerations in online and mobile photo sharing. In: Proc. SIGCHI Conference on Human Factors in Computing Systems (CHI '07)
2. Besmer, A., Lipford, H.R.: Moving beyond untagging: photo privacy in a tagged world. In: Proc. of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)
3. Besmer, A., Lipford, H.R.: Privacy Perceptions of Photo Sharing in Facebook. In: Proc. of the Fourth Symposium on Usable Privacy and Security (SOUPS '08)
4. Fisher, D., Dorner, L., Wagner, D.: Short paper: location privacy: user behavior in the field. In: Proc. of the 2nd ACM workshop on Security and privacy in smartphones and mobile devices (SPSM '12)
5. Henne, B., Szongott, C., Smith, M.: Snapme if you can: Privacy threats of other peoples' geo-tagged media and what we can do about it. In: Sixth ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec) (April 2013)
6. Johnson, M., Egelman, S., Bellovin, S.M.: Facebook and privacy: it's complicated. In: Proc. of the 8th Symposium on Usable Privacy and Security (SOUPS '12)
7. Klemperer, P., Liang, Y., Mazurek, M., Sleeper, M., Ur, B., Bauer, L., Cranor, L.F., Gupta, N., Reiter, M.: Tag, you can see it!: using tags for access control in photo sharing. In: Proc. of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)
8. Krumm, J.: A survey of computational location privacy. Personal and Ubiquitous Computing 13(6) (Aug 2009)
9. Kumaraguru, P., Cranor, L.F.: Privacy Indexes: A Survey of Westin's Studies. Tech. Rep. CMU-ISRI-05-138, Carnegie Mellon University (2005)
10. Liu, Y., Gummadi, K.P., Krishnamurthy, B., Mislove, A.: Analyzing Facebook privacy settings: User expectations vs. reality. In: Proc. of the 2011 ACM SIGCOMM Internet measurement conference. pp. 61–70. ACM (2011)
11. Porter Felt, A., Egelman, S., Wagner, D.: I've got 99 problems, but vibration ain't one: a survey of smartphone users' concerns. In: Proc. of the 2nd ACM workshop on Security and privacy in smartphones and mobile devices (SPSM '12)
12. Squicciarini, A., Xu, H., Zhang, X.: CoPE: Enabling collaborative privacy management in online social networks. Journal of the American Society for Information Science and Technology 62(3) (Mar 2011)
13. Warren, S.D., Brandeis, L.D.: The right to privacy. Harward Law Review 4(5), 193–220 (December 1890)

A Additional details to participants' answers

A.1 Online photo awareness

Figure 7 shows the answers concerning different decision-making criteria for sharing photos on the Web as discussed in Sect. 2.1. Figure 8 shows answers about the estimated chance that someone anytime in the future finds photos that may raise privacy concerns. The items differentiate the ways how a photo is connected to a person as described in Sect. 2.2.



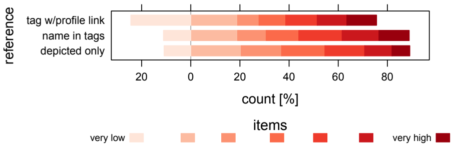


Fig. 7: (q11) Rate the influence of the items as criteria for sharing a photo.

Fig. 8: (q25) Estimate the risk that someone finds an unwanted photo anytime in the future.

A.2 Metadata privacy

As presented in Sect. 3.2, we asked our participants to rate the possible privacy impact of adding metadata to media depicting others on a 7-point scale from *very low* to *very high*. Additionally, we asked them to rate the privacy impact of metadata added by others to media depicting themselves using the same scale. Figure 9 shows the answers to both questions and Table 1 a summary of them.

Table 1: Estimation of the impact of metadata

metadata added by	myself		others		*Wilcoxon*		*Spear-*
with impact to	others		myself		*signed ranks*		*man's ρ*
	mean	*sd*	*mean*	*sd*	*Z*	*p*	($p < .001$)
headline, description, tags	2.94	1.66	3.23	1.75	−4.739	.000	0.73
date & time of creation	3.63	1.70	3.59	1.67	−0.478	.632	0.69
photographer's name	3.49	1.79	3.28	1.83	−3.274	.001	0.65
depicted peoples' names	5.08	1.70	4.76	1.87	−4.204	.000	0.64
broad location (city, region)	3.95	1.62	3.90	1.74	−0.902	.367	0.68
exact location (address, GPS)	5.31	1.68	5.17	1.75	−2.102	.036	0.68

Spearman's ρ indicates that participants' answers to both questions correlate positively (ρ between 0.64 and 0.73, $p < .001$): those who see a higher impact on their own privacy also see a higher impact on other's privacy. We also found a trend that respondents who stated to use location metadata more frequently also saw less privacy impact through that kind of metadata. This may indicate that people who add a particular kind of metadata are more open for the benefits of such information and thus have less concerns about their privacy impact.

The extent of the estimated impact on privacy appears to be independent from the direction of a threat, i. e. regardless of whether a participants's own metadata harms others or foreign metadata harms the participant. For most kinds of metadata, participants perceived that their own metadata has a higher privacy impact on others than others' metadata has on themselves. While some differences were statistically significant, the differences were only slight.

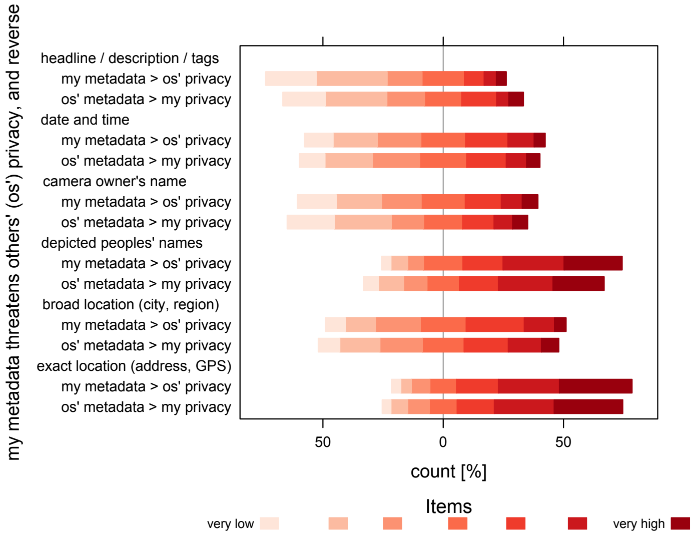


Fig. 9: (q23) Estimate the impact of metadata you add to shared photos on others. (q24) Estimate the impact of metadata others add to shared photos on you.

To compare the privacy levels of different metadata, we asked participants to rate the privacy of different metadata in the context of a privacy-privacy-tradeoff. We used a 7-point scale from *completely public* to *completely private*. The major results as shown in Fig. 10 are congruent with the question about the impact of metadata (cf. Fig. 9). Exact location information is considered to be the most private kind of data, with GPS-based coordinates being more sensitive (91.1% somehow private, 60.4 % completely private, $m = 6.26, sd = 1.2$) than addresses or location names (88.9 % somehow private, 45.7 % completely private, $m = 6.02, sd = 1.4$). Broad locations, like city names, have a mean value of $m = 4.21$ ($sd = 1.5, \text{median} = 4, \text{mode} = 5$). People depicted in the image are the second most private group of metadata, where tags with bounding boxes ($m = 5.46, sd = 1.4, \text{median} = \text{mode} = 6$) in the image are regarded as slightly more private as those without (mean = 5.14, sd = 1.4, median = mode = 5). Again the name of the photographer is regarded as less private (mean = median = mode = $4, sd = 1.7$). The unique id of a camera is also perceived to be more private, with a mean value of 4.72 ($sd = 2, \text{median} = 5, \text{mode} = 7$).

Figure 11 shows feelings about photos with embedded location information that someone might stumble upon as discussed in Sect. 3.2.

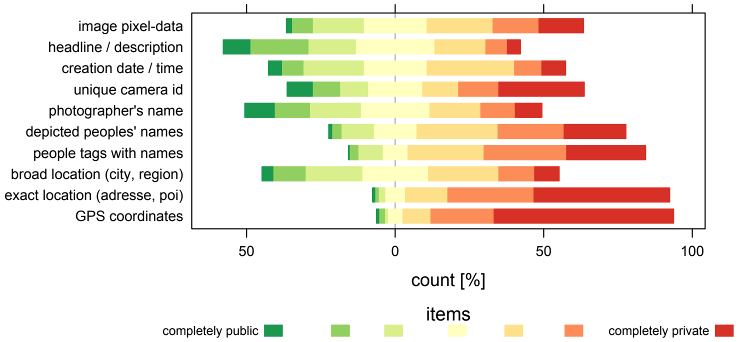


Fig. 10: (q29) How do you feel about the privacy of photo metadata?

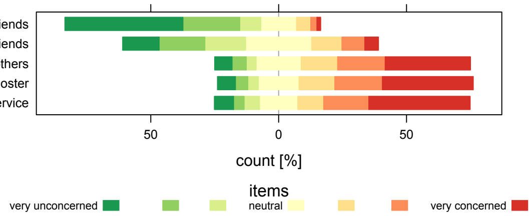


Fig. 11: (q26) How do you feel when these people get to see a photo of yourself that includes location information?

A.3 Privacy-Privacy-Tradeoff

Figure 12 shows answers to the three explicit privacy-privacy-tradeoffs as presented in Sect. 4.

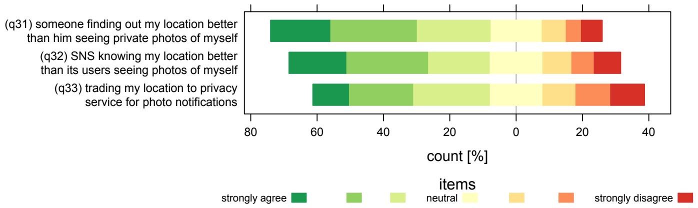


Fig. 12: (q31 - q33) Three explicit privacy-privacy-tradeoffs (cf. Sect. 4)