Tag, You Can See It! Using Tags for Access Control in Photo Sharing

Peter F. Klemperer*, Yuan Liang*, Michelle L. Mazurek*, Manya Sleeper*, Blase Ur*
Lujo Bauer*, Lorrie Faith Cranor*, Nitin Gupta*, Michael K. Reiter†

*Carnegie Mellon University
Pittsburgh, PA
{klemperer, yliang, mmazurek, msleeper}@cmu.edu
{bur, lbauer, lorrie, nitingupta}@cmu.edu

†University of North Carolina
Chapel Hill, NC
reiter@cs.unc.edu

ABSTRACT

Users often have rich and complex photo-sharing preferences, but properly configuring access control can be difficult and time-consuming. In an 18-participant laboratory study, we explore whether the keywords and captions with which users tag their photos can be used to help users more intuitively create and maintain access-control policies. We find that (a) tags created for organizational purposes can be repurposed to create efficient and reasonably accurate access-control rules; (b) users tagging with access control in mind develop coherent strategies that lead to significantly more accurate rules than those associated with organizational tags alone; and (c) participants can understand and actively engage with the concept of tag-based access control.

Author Keywords

Access control; human factors; tagging; privacy

ACM Classification Keywords

H.1.1. User/Machine Systems: Human Factors; D.4.6 Security and Protection: Access controls

INTRODUCTION

Users often have rich and complex sharing preferences for digital content, including online photo sharing [1, 4]. For example, a user may wish to share photos of a work picnic only with co-workers who participated in the event, while blocking those same co-workers from seeing photos taken at a family party. However, non-experts have difficulty using currently available mechanisms to create and maintain access-control policies that capture their sharing preferences [8, 24].

CHI'12, May 5–10, 2012, Austin, Texas, USA.

Tags, including descriptive keywords and captions users add to their photos, could potentially help users create and maintain fine-grained access-control policies more intuitively. Currently, users tag photos for purposes including organization, search, communication, and description (hereafter collectively termed *organizational tags*) [5]. Improvements in automated and assisted tagging — including location tagging via GPS-enabled smartphones and facial-recognition tools found in services like Picasa and Facebook — are making tags increasingly available while decreasing user burden. More broadly, "users have learned to find data by describing what they want ... instead of where it lives" [20], and data management using tag-like organization is emerging in systems such as music management in iTunes, note-taking programs like Evernote or Springpad, and file-tagging mechanisms built into Windows and MacOS.

Systems that use such tags to define access-control policies have been prototyped [3] [6]. However, the usability of tag-based access control has not been investigated using users' own content, tags, and access-control policies. In this paper, we employ an 18-participant laboratory study using participants' own photos to explore the feasibility of tag-based access-control rules for photo sharing. Although tag-based access control could potentially apply to broader categories of digital content, we draw on photo sharing as an initial case study both because users have varied access-control preferences for photos and because systems that allow users to tag photos are already in use.

To explore the efficacy of tag-based access-control rules, we consider the following research questions:

- ***Q1: Can organizational tags be repurposed as-is to create effective access-control rules?***: Users already create tags for purposes including organization, search, description, and communication. To allow tag-based access control to function with minimal overhead, can rules based on these currently available tags capture user preferences?
- ***Q2: Does tagging with access control in mind improve the performance of tag-based access control?***: For tag-based access control to be practical, users must intuit how adding and removing tags affects their access-control policies. When tagging with access control in mind, do users' tags more accurately capture their preferences?

- ***Q3: How do users engage with the concept of tag-based access control?***: Tag-based rule creation should be intuitive and understandable for users. What strategies do users employ when simultaneously tagging photos for both their current, organizational purposes and access control? What do users understand and like about tag-based access control, and what impediments does tag-based access control present? Can users understand and suggest tag-based access-control rules that support their preferences?

We found that organizational tags could be repurposed to create efficient and reasonably accurate access-control rules. When participants tagged photos with access control in mind, they were typically able to develop coherent strategies and create tags that supported significantly more accurate rules than those created from organizational tags alone. We also observed that participants understood the concept of tag-based rules and were able to actively engage in rule suggestion.

We first discuss related work and then detail our methodology. We next present some basic data about our participants' demographics, access-control policies, and tags. We proceed to our main results and analysis and then discuss some study limitations. We conclude by highlighting implications for the design of tag-based access-control systems.

RELATED WORK

This work focuses on photos as a case study to explore how tags could be used for access control. Prior work has examined tag-based access control from both system-building and machine learning standpoints, but only in a limited way concerning usability. The broader literature includes research on both access-control preferences and user-created tags.

Tag-based access control and management

Tag-based access-control systems have been prototyped. Au Yeung et al. implemented a system with access-control policies specified in terms of photos' tags; usability was outside their scope [3]. Hart et al. designed a tag-based access-control system for blogs that performed better than typical privacy tools on a generic blog [6], but they did not examine users' actual preferences for their own content, as we do.

Other work has investigated using tags to predict or recommend access-control policies. Vyas et al. used tags to automatically recommend privacy policies for content, while Squicciarini et al. used a combination of automated image-content analysis and tags to predict privacy policies [25, 23]. In contrast, we allowed users to explicitly define tags for access control.

Prior work also investigated semantic tagging for file and document management [20]. Several distributed file systems have been constructed on this principle [15, 18, 19]. Future systems could include tag-based access control.

Access-control policies and preferences

Studies have demonstrated that users have varied and complex photo-sharing preferences. Miller and Edwards classify digital photographers into two groups: those who share primarily with real-life friends, and those who focus on sharing broadly with online communities [11]. Besmer and Lipford report that concerns about photo privacy are driven by "identity and impression management" rather than by fears about physical safety [4]. Ahern et al. found that sharing decisions were often affected by the location where photos were taken and photo content. They note that sharing decisions are influenced by how easily users can create nuanced policies, demonstrating a need for usable policy management [1].

Beyond photo sharing, research indicates that, while information-sharing preferences can often be categorized into broad classes, exceptions are frequent and important [9, 14]. Other studies report that sharing preferences are frequently dynamic, depending on content lifecycles, the context of sharing requests, and expectations about how shared data will be used [10, 16]. Other work suggests sharing decisions may be governed by the difficulty of setting and updating policies, reinforcing the idea that users must be able to easily set up access-control policies [22, 24].

Tagging

We investigate creating access-control policies from photo-management tags. A number of researchers have examined tagging behavior. Ames and Naaman categorized photo tags as organizational or communicative and intended for oneself or others [2]. Kirk et al. found users often group photos by event [7]. Nov et al. found that users' tagging motivations influence the number of unique tags they create [13].

More broadly, Zubiaga et al. suggest that taggers who categorize rather than describe content provide better inputs to automated classifiers [26]. Gupta et al. provide a survey of tagging motivations, content, and recommendations [5].

METHODOLOGY

We designed an exploratory laboratory study during which participants performed three separate tagging tasks. The first task focused exclusively on *organizational tagging* to help a user organize and search her photos, while the second and third tasks focused on organizational tagging in combination with tagging for access control. These tasks provided insight into participants' tagging behaviors and strategies (Q3). Tags from these tasks were also used to create machine-generated access-control rules that roughly approximated users' policies. Some of these rules were shown to the participants to demonstrate the tag-based access-control concept and stimulate discussion (Q3). We also used the tags and machine-generated rules during post-processing to evaluate the efficacy of organizational tags for access control (Q1) and to compare the performance of organizational-only tags to combined organizational and access-control tags (Q2).

Recruitment

We used advertisements on Craigslist, in the university's newspaper, on a university research participant website, and on posters around Pittsburgh to recruit English-speaking participants who take at least 100 photos per year. Because we were interested in the usefulness of existing organizational tagging

strategies for access control, we required that participants add keyword tags or captions to photos "often" or "always." We eliminated participants who only tagged photos on Flickr or Facebook since Flickr tags tend to be created for sharing [11], and Facebook tags are limited to people's names. We felt that including such users would skew the results, although excluding them may limit generalizability.

Qualified participants were asked to upload 40 photos they had previously tagged. To prompt potential participants to provide photos for which they might have varied access preferences, we provided them with a list of 17 suggested photo categories, including "up to 15 photos that you haven't posted publicly and wouldn't want to share with the general public," "3 photos with trees in them," and "3 photos with your relatives in them." We also asked them whether they would be willing to share the photos with "some," "none," or "all" of several groups of people. Potential participants who answered "none" or "all" to all categories were eliminated. As we discuss in detail in the Limitations section, we believe these measures were successful.

We sent our screening survey to 152 people, 63 of whom completed the survey. Of those 63, we rejected 39 people who only tagged photos on Facebook or Flickr, 5 who did not tag frequently enough, and 1 who lacked English proficiency. The remaining 18 people made up our participants.

Procedures

Qualified participants were invited to our lab for the main part of the study, which lasted between 1.5 and 2.5 hours. During this portion of the study, we observed their organizational tagging behaviors and their strategies for incorporating access control into their tagging schemes. We also presented participants with machine-generated, tag-based access-control rules, both to demonstrate tag-based rules and to gauge their reactions. For this portion of the study, we used the Picasa desktop photo software and custom web interfaces.

Warmup task

We gave each participant a brief tutorial on tagging photos in Picasa. As a warmup task, we asked her to add at least one tag to each of five sample photos unrelated to her own photos.

T1: Organizational tagging

T1 was designed to evaluate how effective organizational tags can be for expressing access-control policies.

Prior to the lab session, we stripped all existing tags from the participant's photos, saving these *original tags* for later reference. In the lab, we asked the participant to re-tag her stripped photos with the objective of finding the photos more easily in the future, adding as many tags as she would like. We asked participants to re-tag their photos so we could observe each participant's tagging behavior using a think-aloud procedure.

Collecting access-control preferences

Next, we collected a set of the participant's access-control preferences for her study photos. These preferences served as ground truth for creating and evaluating access-control rules.

We collected a list of people with whom the participant might want to share photos. We first prompted for three people with whom she had a close relationship, including household members, friends, and significant others. We then prompted for four to seven people with whom she had less close relationships, such as supervisors, friends of friends, neighbors, and colleagues. From here forward, we will refer to this entire set of people as the participant's *friends.*

We then presented the participant with an *access grid* mapping her photos to her friends. We added a "Public" friend column to represent posting a photo publicly, in connection with the participant's real name. For each combination of friend and photo, we asked the participant to select a preference from the following options:

- **Strong allow**: Allow access; would be upset if the friend were not able to view the photo.
- **Weak allow**: Allow access; would not be very upset if the friend were not able to view the photo.
- **Strong deny**: Deny access; would be upset if the friend were able to view the photo.
- **Weak deny**: Deny access; would not be very upset if the friend were able to view the photo.
- **Neutral**: Absolutely no preference between allowing and denying the friend access to the photo.

To confirm understanding, we asked the participant to point out and explain one example for each type of preference (or to explain why that preference would not be needed).

Example rules

At this point, we introduced the participant to the concept of tag-based access-control rules. To aid this introduction, we created a set of machine-generated best-fit access-control rules for each of the participant's friends. Rule generation is described in more detail below. Each friend was assigned zero or more rules of the form "If *tagged / not tagged* with *tag*, then *allow / deny*," combined with *and* and *or* as appropriate. Each friend was also assigned a *default rule* of allow or deny, which applied to any photos not covered by the other rules. Some friends were assigned only default rules — for example, the participant's spouse might be assigned the simple rule of "always allow" access to all photos. The participant's boss, by contrast, might be assigned a more complicated ruleset with three rules: "If tagged with *landscape*, then *allow*"; "if tagged with *work*, then *allow*"; and the default rule "otherwise deny."

We applied the machine-generated rulesets to the participant's photos, resulting in a set of allowed and denied photos for each friend. In the interest of time, we selected two example rulesets to present, including at least one example with at least one non-default rule. If no non-default rulesets were created, we showed the participant a default-only ruleset, as well as a generic non-default example prepared in advance.

Our goal with these examples was to familiarize the participant with the idea of tag-based access-control rules and get preliminary reactions, rather than to evaluate the success of these rules in detail. We designed a simple rule-display interface, shown in Figure 1, intended only to demonstrate the

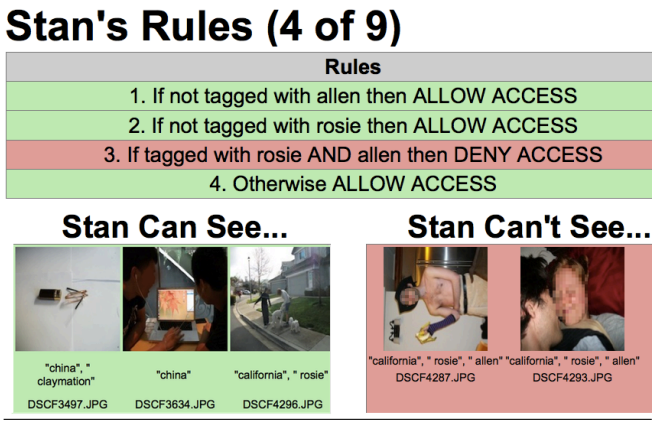

Figure 1. The interface we used to demonstrate machine-generated rules and their effective access policy. (Names changed and faces blurred.)

machine-generated rules and stimulate discussion. The interface displayed the text rules for a given friend at the top of the screen and a thumbnail photo display at the bottom. The photo display distinguished the photos the friend was and was not allowed to see under the rules. Each photo was displayed with its tags to help the participant understand how the rules were applied. This interface was not intended to simulate any real system for managing tag-based access control.

T2: Tagging for access control

After using the sample machine-generated rules to introduce the concept of tag-based access control, we returned to Picasa for T2. We invited the participant to add to and/or delete from the tags she had added in T1, with the joint objectives of finding photos more easily and creating tag-based access-control rules. As before, we observed the participant's tagging behavior and strategies using a think-aloud mechanism.

Detailed review of machine-generated rules

Next, we explored how successful the participant's T2 strategy had been and gathered detailed feedback on the tag-based access-control concept. We used the tags from T2 and the participant's access-control preferences to create a new set of machine-generated access-control rules for each friend. We showed the participant all the resulting rulesets, using the same interface, one friend at a time. For each friend, we asked detailed questions about how accurately the rules reflected the participant's preferences, including for other photos she had taken in the past or might take in the future.

We also used our interface's *"show conflicts" view* to highlight any photos that were misclassified. We asked how upset the participant was about the misclassifications and how they affected her view of the ruleset's overall success. We also asked how she might change the tags or the ruleset to more accurately reflect her preferences.

T3: Refinement and wrap-up

To examine what the participant had learned, we invited her to add to and/or delete tags from T2, this time keeping in mind what she had learned during the detailed rule review. Once again, the goal of the tagging was to make both finding photos and developing access-control rules easier. As before, a think-aloud mechanism helped us observe the participant's behavior and strategy. Although we created rules from these tags for post-processing and analysis, in the interest of time we did not show these rules to the participant. Finally, we asked each participant a series of general questions about her photo-tagging and sharing habits.

Machine-generated rules and analysis

We used machine-generated rules both to demonstrate tag-based access-control rules to our participants and to conduct post-interview analysis. For demonstration, we were mainly interested in creating rules that were somewhat human-readable and would provoke discussion (Q3). In post-interview analysis, we used the machine-generated rules to investigate Q1 and Q2: Could we construct reasonably well-fitting access-control rules from organizational tags, and would tags modified with access control in mind produce better rules? For this purpose, a rough approximation seemed sufficient, so we did not attempt to find an optimal rule-generating algorithm or construct the best possible rules for each participant.

To achieve both goals, we created rules using an open-source implementation of the c4.5 decision-tree algorithm [12].[1] We trained the algorithm for each friend, lumping together weak and strong preferences into *allow* and *deny* categories and using default sensitivity settings. Photos with neutral preferences were ignored during training.

We displayed the results of the training to the participants and used them in our later analyses. We did not use separate training and test sets, because we wanted to establish a baseline scenario for how tags aligned with access-control policies. Future work might separately consider finding the optimal algorithm for generating rules, as well as how well rules generated from one set of photos could predict access-control policies for other photos.

We report the results of generating rules from the participants' original tags, as well as their tags from tasks T1, T2 and T3, in the Results and Analysis section.

DEMOGRAPHICS, POLICIES, AND TAGS

Table 1 lists demographic information for our 18 participants. Half were men, and half were women. The subjects trended young (between 18-32) and technologically focused: 10 of the 18 self-identified as science, technology, engineering and math (STEM) professionals or students. This bias toward youth and STEM professions may limit the generalizability of this study. Other work related to tagging and access control has also focused on similarly young populations [23, 25].

Our 18 participants provided between 40 and 48 photos each and listed 7 to 10 friends, plus Public. Overall, they expressed 6847 access preferences, each for one combination of a photo and a friend. 15.7% of preferences were strong allow, 40.8%

[1] http://www2.cs.uregina.ca/~dbd/cs831/notes/ml/dtrees/c4.5/tutorial.html

Code	*Age*	*Gender*	*Occupation*	*Photos/year*	*Tag software*
P01	23	F	STEM professional	1001-5000	Picasa
P02	20	F	engineering student	101-500	iPhoto
P03	27	F	service industry	501-1000	Picasa, Facebook, Tumblr
P04	32	F	STEM professional	1001-5000	Skydrive
P05	24	F	student	1001-5000	iPhoto
P06	23	M	engineering student	501-1000	Picasa
P07	22	M	engineering student	101-500	Picasa
P08	24	M	student	101-500	Picasa
P09	18	M	engineering student	5001+	Picasa
P10	24	M	STEM professional	5001+	Photoshop Album
P11	26	M	STEM professional	1001-5000	Flickr
P12	28	F	art, writing	5001+	Lightroom
P13	23	M	clothing designer	51-100	Twitter, Yfrog, Photobucket
P14	20	M	engineering student	1001-5000	Picasa
P15	19	F	music student	501-1000	Picasa
P16	29	F	anthropology student	1001-5000	iPhoto
P17	25	M	STEM professional	501-1000	Picasa
P18	18	F	art student	1001-5000	iPhoto

Table 1. Participant demographics.

were weak allow, 11.0% were strong deny, 14.9% were weak deny, and the remaining 17.5% were neutral. The distribution of preferences, however, varied widely across participants. P11 was most permissive, allowing 87.8% of access combinations; P04 was most restrictive, denying 80.0% of access combinations.

In T1, participants used on average 2.6 total tags per photo; P07 used the most, with 5.0, while P08 used the fewest, with 1.0. It is also possible to count the number of unique tags (e.g., counting the tag "family" once whether it was used once or many times). We will refer to this count as "unique" tags. Considering only each participant's unique tags, the average was 1.2 per photo, with a minimum of 4 tags for 48 photos (P13) and a maximum of 130 tags for 40 photos (P03).

RESULTS AND ANALYSIS

Our results show that tags created for organization can often be used to create reasonably effective access-control rules (Q1). Asking users to update their tagging schemes with access control in mind produced even more accurate rules, in many cases with only limited modifications to the tags (Q2). We also observed that most participants quickly understood tagging for access control and were able to develop and apply a modified tagging strategy (Q3).

In the first two subsections, we describe results related to Q1 and Q2, respectively. Taken together, qualitative results from both subsections also address Q3.

Organizational tags can express many access policies

Overall, organizational tags performed well as the basis for access-control policies. Rules generated automatically from the participants' T1 tags were highly accurate, resulting in few false allow or false deny conflicts and indicating that the tags were expressive enough to be useful for such policies. Overall, the rules generated conflicts for only 7.8% of non-neutral photo-friend combinations in T1, with a conflict rate under 5% for one-third of participants. The best case for organizational tags was no conflicts for P11, who had a very simple allow-most policy. The worst observed case was user P08, with 19% conflicts, due in part to his use of long, complex tags that were not repeated across photos. Figure 2 shows the rate of conflicts in each of the three tasks.

As a control, we compared our results to a simple default policy of either allow all or deny all, choosing the more accurate option for each participant-friend combination. In case of ties, we chose deny all to preserve privacy. This default policy, illustrated in Figure 2, produced more than twice as many conflicts as T1 (15.8% to 7.8%, significant, paired Wilcoxon, $\alpha = 0.05$).

Considering T1 conflicts in more detail, we found that for most conflicts (83.5%), the suggested rules disagreed with the less-serious weak preferences, bolstering the case that it is possible to make effective rules from organizational tags. To some extent, this reflects the fact that most non-neutral preferences were weak; however, we find that across participants, the proportion of conflicts with weak preferences was greater than the proportion of all preferences that were expressed as weak preferences (χ^2 per participant, aggregated using Fisher's combined test, $p < 0.05$).

In access control, false allows (erroneously granting access) are often of greater concern than false denies (erroneously denying access). In T1, 57% of conflicts were false allows. We might expect this to mirror the proportion of preferences that were deny preferences (since conflicts with those would be false allows); however, only 31% of preferences were deny preferences, a proportion significantly lower than that of the false allows (χ^2 and Fisher's combined test, as above, $p < 0.05$). More positively, most of those false allows conflicted only with weak preferences: just 13% of conflicts (less than 1% of decisions) were false allows that conflicted with a strong preference. We hypothesize that a classification algorithm tuned against false allows would ameliorate this further.

For most participants, a small number of photos were responsible for most conflicts. In T1, only 27% of all photos were misclassified by the machine-generated rules at least once, for any friend. The worst case was P04, for whom 60% of photos were misclassified at least once; for more than two-thirds of participants, fewer than one-third of all photos were ever misclassified. This suggests that minor improvements in tagging or rule generation might reduce conflict rates considerably.

Simple rules from organizational tags

Another way to evaluate the performance of tag-based access-control policies is to consider their complexity; policies with too many rules or rules with too many clauses could prove incomprehensible to users. We measure rule complexity in two ways: by counting the number of non-default rules generated for each friend, and then by counting the number of unique tags used in those rules.

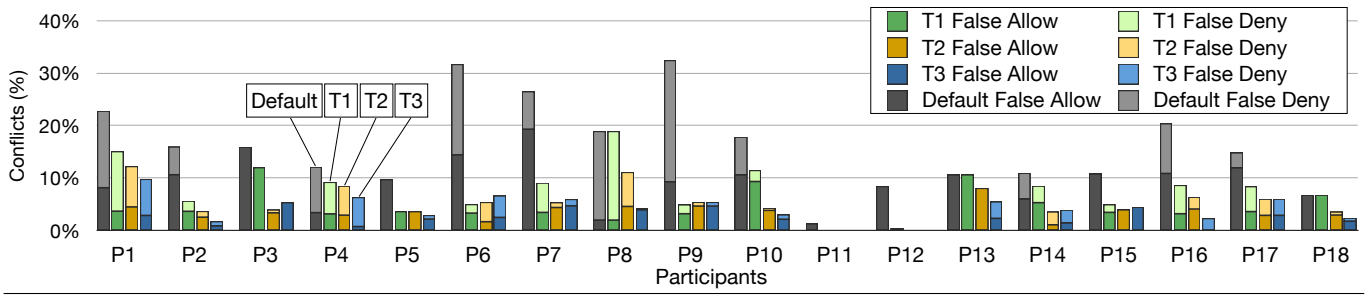

Figure 2. Conflict rate across tasks, for all participants, distinguishing false allows and false denies. For each participant, we show four vertical bars, one each for the default policy, T1, T2, and T3. The default policy is the more accurate of allow all and deny all for each friend. For most participants, the conflict rate was highest for the default and decreased in consecutive tasks. P11 had no conflicts, and P12 had only 0.02% false deny conflicts in T2.

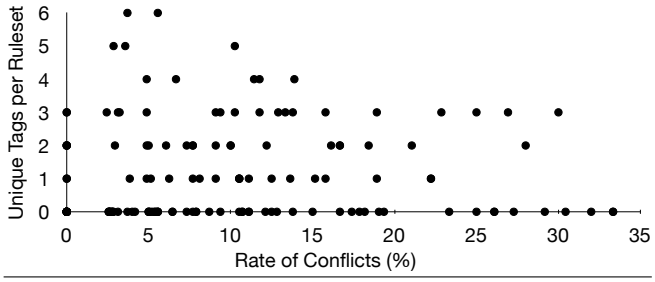

Figure 3. Rule complexity (measured in unique tags per ruleset) versus accuracy (measured in conflict rate), for 168 participant-friend combinations. A value of zero unique tags denotes a default-only ruleset (deny all or allow all). In general, many rules are both simple and accurate, and few rules are both complex and inaccurate.

By both of these measures, the rules we generated were fairly simple. In T1, more than half of the 168 participant-friend combinations resulted in default-only rules. Among the 75 non-default cases, an average of 2.8 non-default rules were made for each friend. Even the worst case was relatively straightforward: the largest four rulesets contained only 4 non-default rules each. Examining the unique tags used in each ruleset yields similar results. On average, non-default rulesets for T1 contained only 2.4 unique tags. The worst case, rulesets containing six unique tags, occurred only twice among the 75 non-default rulesets.

It is also interesting to consider the relationship between accuracy and rule complexity. Prior to the study, we hypothesized that more-accurate rulesets would also be more complex, necessitating a tradeoff between expressiveness and ease of use. To examine this, we plot rule complexity (measured by unique tag count) against accuracy (measured by conflict rate), for all 168 participant-friend combinations, in Figure 3. The results show many instances of rulesets that are both simple and accurate: examples include allowing a spouse to see all photos, forbidding the public to see any photos, or allowing a boss to see only photos tagged with "work." We also see many rulesets in the top-left and bottom-right, indicating at least some tradeoff between accuracy and complexity, with few instances of rules that are both complex and inaccurate.

Reactions to sample rules

After creating machine-generated rules from the T1 tags, we showed each participant two rules and asked for general reactions. Many participants liked these rules and found them intuitive. For example, P05 said the rules for a former teacher, which denied access to photos from a friend's wedding and a graffiti-covered landmark, made sense because the rules "conceptualized what [she] was thinking." P17 said rules created for a friend with a child his son's age matched "the intuitive rule that [he] made" while setting up his preferences: the friend could see only photos containing P17's son.

However, the rules created from the organizational tags were not always fully successful. In some cases, the tags included in the rules were too general or too specific. For example, P09 said a rule denying his roommate access to photos tagged with "gf" was too general, as he wouldn't need to restrict all photos containing his girlfriend. Other participants were generally satisfied with the rules but flagged some exceptions. P14 said a rule for a former teacher "seem[ed] roughly accurate," but he was upset that the teacher could see one embarrassing, slightly lewd picture. In other cases, the machine-generated rules appeared coincidental, fitting the participant's preferences but using tags with little or no relation to the participant's policy decision-making.

Reactions to tag-based access control

After reviewing the sample rules, we asked participants for their overall impressions of tag-based access control, and found that the concept typically made sense. On a five-point Likert scale, 13 participants said the concept made complete sense or some sense (scores of 5 or 4), two were neutral (score of 3), and one said the concept did not make sense (score of 2). Two others said it depended on circumstances (no score).

Among those who said the concept made complete sense, several said making policy using tags would save time or be "more efficient" (P03). Tag-based rules worked particularly well for P11, who had a subtle preference for preventing his family from seeing certain combinations of people: "If they can avoid seeing me and my girlfriend together, I'd probably use it for that."

P06, the only participant to say tag-based rules did not make sense, explained that he would need "a large number of tags

to make it easier to make rules." P09, who chose neutral, expressed related concerns about scalability: "The results are great, but if you added more photos these rules would break down."

Ad-hoc access control with organizational tags

As another indicator that organizational tags may be appropriate for access-control policies, we found that several participants already used photo tagging to help implement their intended policies in various ad-hoc ways. P17 tagged his photos based mainly on the events at which they were taken; he used these tags to help him sort out which photos should be shared with whom on PicasaWeb. In the organization task, P01 tagged photos of herself with her boyfriend to keep track of things she didn't want her family to see.

Similarities between organizational and original tags

As described in Methodology, we also requested participants' original tags — that is, tags they added to the photos prior to the study — to confirm that the tags created in T1 were not highly different from tags the participants normally create. We were able to collect original tags for half our participants.

Our results indicate the two sets of organizational tags were reasonably similar. The overall rate of conflicts for rules generated from the original tags is 10.6%, compared to 8.5% for rules from T1 tags for the same 9 participants (not significant, paired Wilcoxon test, $p > 0.2$). This includes an outlier in P06, whose original photos included captions that were markedly different from the keyword tags he used in T1, resulting in a conflict rate of 25.9% for his original tags compared to 4.9% for his T1 tags. Original-tag conflict rates for the other 8 participants for whom this data was available were within 3 percentage points of the T1 rates. This provides a rough indication that T1 aligns with natural tagging behaviors for the majority of participants.

Tagging for access control provides improvement

Although access-control policies generated from organizational tags performed reasonably well, we found that policies generated from dual organization/access-control tags performed even better. The overall conflict rate improved significantly[2] from 7.8% in T1 to 5.2% in T2, and the worst-case rate improved from 18.9% (P08) to 12.2% (P01). Across participants, the average improvement was 2.7 percentage points.

We observed a smaller but still significant improvement between T2 and T3, as participants fine-tuned their tags after viewing rules. The overall conflict rate improved to 4.2%, and the worst case improved to 9.8% (P01). Individual participants' conflict rates dropped 1.1 percentage points on average.

The ratios of weak to strong conflicts and of false allows to false denies did not change significantly among T1, T2 and T3. Unsurprisingly, participants were more concerned with false allows than false denies. During the detailed rule review, we asked participants to report how upset they were about each conflict on a five-point Likert scale. Comparing median upsetness per participant, we find that false allows cause significantly more upset (paired Wilcoxon test, $p < 0.01$).

[2]Unless otherwise noted, significance tests in this section use a Friedman repeated-measures test to establish differences among the tasks, then paired Wilcoxon tests (chosen a priori) to separately compare T1 to T2 and T2 to T3. $\alpha = 0.05$.

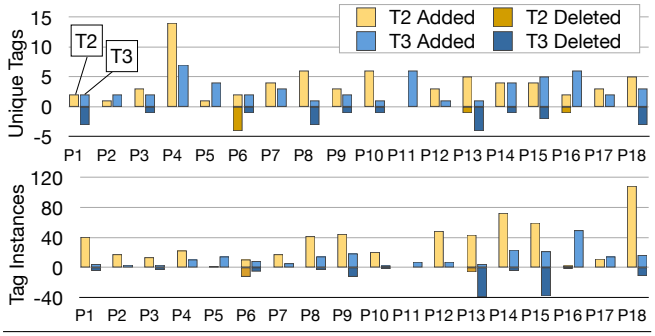

Figure 4. The unique tags and total tag instances added and deleted in T2 and T3, per participant.

Limited modifications for access-control tagging

Overall, participants made few modifications to their T1 tags in T2, averaging 33 modifications each, or less than one modification per photo. On average, participants only added or deleted 4 unique tags. Most of these changes (97%) were additions; only three participants deleted any tags in T2. P18 made the most modifications (108 additions, 5 unique); P11, by contrast, made no modifications.

Participants made even fewer refinements in T3, averaging 19 additions or deletions each (less than 0.5 per photo). P15 made the most (59, 7 unique), and P02 made the fewest (3, 2 unique). The refinement step was more balanced between addition and deletion; 65% of modifications were additions, and 10 participants deleted at least one tag. Figure 4 shows each participant's overall and unique modifications in both tasks.

Rule complexity across tasks

The overall pattern of avoiding complex, inaccurate rules was maintained across T2 and T3. The distribution of the complexity-accuracy tradeoff appeared to shift slightly, with more complex-but-accurate rules and fewer simple-but-inaccurate rules. This is attributable in part to a drop in default-only rulesets, from 93 in T1 to 66 in T2 and 59 in T3, as policies become more precise. However, this change was not signficant.

Strategies for access-control tagging

Participants appeared to follow one of three tagging strategies in T2: leveraging content-based tags for access control, using tags specifically designed to indicate access-control policy, or using a hybrid of content- and policy-based tags. Most participants articulated a clear strategy and applied it consistently, demonstrating that they had quickly grasped the concept of tag-based access control.

Five of 18 participants used the content-based strategy. At the extreme, P11 made no modifications because he felt his organizational tags were sufficient to specify his policy. P05 added "kiddos" to photos containing children, because their parents might not want those photos to be shared. P16 added "Belize" to one photo and changed "outside Sleeping Bear Lake" to just "Sleeping Bear Lake" on another; in both cases her modifications placed the photo in question into a group with other photos for which she had similar policy preferences. Participants using this strategy added, on average, 14 new tags and 4 new unique tags in T2.

Five other participants used an entirely policy-based strategy, creating tags indicating sharing policies like "private," "public," and "for friends." P09, who adopted this strategy, said he "would separate the two ideas [content and policy tags] completely." On average, participants using this strategy in T2 added 51 new tags and 4 new unique tags each.

The remaining eight participants used a hybrid strategy that combined policy- and content-based tags. Some of these participants used tags that conveyed both content and policy information, such as a "family" tag indicating family members could access photos containing family members. P07 used tags like these for family, close friends, and general friends. Others used both content and policy tags to form a combined policy: P10 wanted to restrict photos with the "private" tag from most people, as well as photos with "strange" or "weird" content from some of his less-close friends. Participants using the hybrid strategy added, on average, 31 new tags and 4 new unique tags in T2.

During the detailed rule review after T2, we asked participants to suggest ways to improve the machine-generated rules. Many were able to articulate simple rules that matched their tagging strategies more closely than the automated rules. For example, P02 noted the machine-generated rules did not pick up on her strategy to restrict photos tagged "goofy" from less close friends.

Refinement strategies in T3

As previously noted, participants refining their tags after the detailed rule review made few additional modifications. The modifications they did make generally demonstrated a strong grasp of tag-based access control and a consistent approach to making tags that would facilitate better rules.

Most participants used T3 to adjust the granularity of their access-control-tagging scheme. For example, P03 originally tagged photos she didn't want made public with "drunk"; in T3, she added a "very drunk" tag to distinguish permissions for different levels of drunkenness. P18 changed policy tags on some photos to distinguish between "friends" and "close friends." Others used T3 to make rules more generalizable: P16 added "landscape" tags on top of "Sleeping Bear Lake" to make a broader category, "because if it's restricted to lake pictures," her friends were "not going to be seeing much."

Other participants added tags that were not necessary for creating policies for the photos and friends in the study, but could be useful for broader policies. For example, P17 added a "Pittsburgh" tag to photos taken in Pittsburgh; these photos were already classified correctly for the friends in the study, but he wanted to make a rule to share the photos explicitly with friends from Pittsburgh. T3 was also frequently used to make corrections or fix inconsistencies in tags. P10 consolidated "weird" and "strange" into one "weird" tag after noticing both were used in rules.

Only one participant overhauled her entire tagging scheme in T3: after viewing her rules, P15 switched from a policy-based strategy in which she assigned tags based on who *should* see the photos to an inverse strategy of tagging based on who *should not* see the photos.

LIMITATIONS

There are several important limitations to our study design. First, our results are limited by the participants we recruited and the photos they provided. As discussed in Demographics, our subject pool skewed young and technical; it also included only people willing to upload their photos to our recruitment website. The generalizability of the photos provided was also limited by our request that participants upload previously tagged photos. We requested tagged photos so we could examine the ecological validity of tags created in our lab (as discussed in Results and Analysis). During recruitment, we attempted to encourage participants to upload a range of photos with different access preferences. We believe we succeeded, in part because participants expressed deny preferences for 25% of photo-friend combinations, including photos with sensitive content like alcohol, unprofessional behavior, and skinny dipping. We also asked participants about their current photo sharing preferences: 14 participants said they had published 80% or less of their study photos online. Most used various access-control mechanisms to protect photos they did publish, including setting privacy preferences on websites like Facebook and Picasa Web Albums. Participants distinguished clearly between these protected photos and photos published "publicly" on Tumblr, Twitter, personal websites and blogs. However, we acknowledge participants were unlikely to share their most private photos with us.

A second set of limitations concerns our use of machine-generated access-control rules. The algorithm has no access to the context and meaning of tags and no insight into the policy the participant intended when tagging for access control. As a result, some rules appeared strange or arbitrary to the participants, potentially driving them toward explicit policy-based tags like "private" and "public." We chose to use machine-generated rules to establish a standardized baseline for comparison accross tasks and provoke discussion with the participants. We also did not attempt to optimize our rule-generation mechanism or produce the best possible machine-generated rules. A better algorithm might result in fewer conflicts or strike a better balance among strong and weak conflicts and false allows and false denies.

Other limitations concern scale and generalizability. Our quantitative results measure only how well the rules fit the photos provided by the participants for the study; we talked to participants about how well the rules would generalize to other

photos and friends, but cannot draw firm conclusions. Similarly, we cannot comment on whether the rules would remain tractable when dealing with thousands of photos and hundreds of friends, family members, and acquaintances.

DISCUSSION

Our results indicate that tag-based access-control seems promising. In this section, we discuss potential design implications that arose from our findings and observations.

Automated rule generation

Participants were generally supportive of automated rule generation, and several explicitly said that they would like a system that suggested access-control rules for their photos. As P12 put it, "That's kinda handy." Participants also did not seem to mind suggesting tweaks to automatically generated rules that did not completely capture their preferences. Although asking users to fill in an access-control grid for all their photo and preference combinations, as we did in the study, would be unrealistic in an actual system, such a system could potentially leverage users' willingness to tweak slightly incorrect rules by asking for a small set of ground-truth preferences over time and using these preferences to offer suggestions for baseline rules. Such a feature might be especially useful for people just starting to use a tag-based access-control system: results from our pilot testing indicate that users may better understand a tag-based rule system when it is demonstrated with their own tags and photos, rather than with generic examples.

Additionally, we did not try to optimize the machine-learning algorithm to reduce policy conflicts. Current work in privacy policy prediction promises to reduce conflicts further than observed in this work [23, 25].

Varied approaches to exception handling

Participants showed varying levels of preference flexibility when presented with tag-based rules. We categorize a user as *flexible* if, when presented with conflicts during the detailed rule review, she generally indicated this new access-control setting was also acceptable. In contrast, a *strict* user would attempt to modify her tags and propose new rules to resolve the majority of conflicts.

The 18 participants were evenly split between flexible and strict. P03 was representative of flexible users, saying, "If a couple more things got cut off than I intended, then it wouldn't matter to me so much." P10 felt similarly, explaining, "There's no reason for her to see relative photos, but I don't care too much." In contrast, P08 was representative of the strict users. When he encountered conflicts, he decided to create a number of additional tags and suggest associated rules to resolve those conflicts, saying, "One tag doesn't suffice for these three groups; [even] three tags are not enough."

A system design should account for people with both flexible and strict preferences. Stricter users could potentially be satisfied by providing an option for exception handling; however, providing such an option would need to be balanced with encouraging users to create generalizable rules to promote usability. P18 provides an example of this dynamic. She said that she had a large number of Facebook friends, and, while most of the time she would want to set permissions for groups, she would need some individual exceptions but not so many that the exceptions would became hard to manage.

Interface-supported rule management

Our results indicate that it is possible to repurpose organizational tags to create rules that capture the majority of a participant's preferences. However, a practical system must also help users create and manage rules and understand the impact of tag and rule changes.

Although our rule-display interface was not intended to represent a real system, we asked participants for their impressions of it to gain some insight into potentially useful design features. As expected, most participants used the text rules to understand the future impact of the policy, and the photos to understand the immediate impact. P03 found both the rules and photos useful, saying he was worried about "not remembering how you tagged very specific photos ... I think seeing them is a more visceral response." A few participants mentioned that the photo display might be hard to use for larger sets of photos.

A tag-based access-control system could help users set policy by appropriately displaying relevant photos and rules, as well as demonstrating what would happen if the user changed tags or rules. The Expandable Grid [17], for example, could be used to demonstrate the impact of rule adjustments.

System support for manual tagging

Our users were able to actively and successfully engage in tagging for access control immediately after being exposed to tag-based access control (Q2). A practical system could further encourage tagging for access control in a variety of ways. One possibility is checking tags for consistency: for example, asking whether a user missed one tag in a group of photos she otherwise tagged similarly, noting slight changes in spelling, or highlighting use of close synonyms like P10's "strange" and "weird."

Additionally, a practical system could detect the types of tags frequently used in rules and remind the user to add these tags. We found that users tended to add descriptive, people, or permission tags when tagging for access control in T2. Displaying frequently used tags in such categories might nudge the user to create tags that are more useful for access-control rules. In addition, we investigated only user-provided keyword tags. However, person- and location-based tags were among the most common to appear in the access-control rules. Such tags are well supported by existing automation tools, including Picasa's face-detection feature (turned off during our study). Automated support for tagging is an emerging tool that could reduce user burden and help users add more, and more accurate, tags [21].

CONCLUSION

Overall, we found that tag-based rules are promising for use in an access-control system. Organizational tags can be repurposed to create reasonable access-control policies, and when participants actively create tags for access control, policies based on these tags are yet more accurate. Participants are able to suggest and engage actively with tag-based rules.

These results suggest that it would be possible to create a usable access-control system with tag-based rules and minimal tagging overhead. It may be possible to additionally aid users with appropriate support for automated rule generation, exception handling, intuitive policy management, and automated tag generation and correction.

ACKNOWLEDGMENTS

This research was supported by the NSF via grants CNS-0831407 and DGE-0903659, by CyLab at Carnegie Mellon under grants DAAD19-02-1-0389 and W911NF-09-1-0273 from the Army Research Office, by the ARCS Foundation, and by Cisco Systems, Inc. We thank Greg Ganger, Saranga Komanduri, and Raja Sambasivan for their help and advice.

REFERENCES

1. Ahern, S., Eckles, D., Good, N. S., King, S., Naaman, M., and Nair, R. Over-exposed?: privacy patterns and considerations in online and mobile photo sharing. In *Proc. CHI*, 2007 (2007), 357–366.
2. Ames, M., and Naaman, M. Why we tag: motivations for annotation in mobile and online media. In *Proc. CHI* (2007), 971–980.
3. Au Yeung, C., Kagal, L., Gibbins, N., and Shadbolt, N. Providing access control to online photo albums based on tags and linked data. In *Social Semantic Web* (2009).
4. Besmer, A., and Richter Lipford, H. Moving beyond untagging: photo privacy in a tagged world. In *Proc. CHI* (2010), 1563–1572.
5. Gupta, M., Li, R., Yin, Z., and Han, J. Survey on social tagging techniques. *SIGKDD Explor. Newsl. 12* (November 2010), 58–72.
6. Hart, M., Castille, C., Johnson, R., and Stent, A. Usable privacy controls for blogs. In *Proc. ICCSE* (2009), 401–408.
7. Kirk, D., Sellen, A., Rother, C., and Wood, K. Understanding photowork. In *Proc. CHI* (2006), 761–770.
8. Maxion, R. A., and Reeder, R. W. Improving user-interface dependability through mitigation of human error. *Int. J. Hum.-Comput. Stud. 63* (July 2005), 25–50.
9. Mazurek, M. L., Arsenault, J. P., Bresee, J., Gupta, N., Ion, I., Johns, C., Lee, D., Liang, Y., Olsen, J., Salmon, B., Shay, R., Vaniea, K., Bauer, L., Cranor, L. F., Ganger, G. R., and Reiter, M. K. Access control for home data sharing: Attitudes, needs and practices. In *Proc. CHI*, 2010 (2010), 645–654.
10. Mazurek, M. L., Klemperer, P. F., Shay, R., Takabi, H., Bauer, L., and Cranor, L. F. Exploring reactive access control. In *Proc. CHI* (2011), 2085–2094.
11. Miller, A. D., and Edwards, W. K. Give and take: a study of consumer photo-sharing culture and practice. In *Proc. CHI* (2007), 347–356.
12. Mitchell, T. *Machine Learning*. McGraw-Hill, 1997.
13. Nov, O., Naaman, M., and Ye, C. What drives content tagging: the case of photos on Flickr. In *Proc. CHI* (2008), 1097–1100.
14. Olson, J. S., Grudin, J., and Horvitz, E. A study of preferences for sharing and privacy. CHI EA (2005), 1985–1988.
15. Ramasubramanian, V., Rodeheffer, T. L., Terry, D. B., Walraed-Sullivan, M., Wobber, T., Marshall, C. C., and Vahdat, A. Cimbiosys: a platform for content-based partial replication. In *Proc. NSDI* (2009), 261–276.
16. Razavi, M. N., and Iverson, L. A grounded theory of information sharing behavior in a personal learning space. In *Proc. CSCW* (2006), 459–468.
17. Reeder, R., Bauer, L., Cranor, L., Reiter, M., Bacon, K., How, K., and Strong, H. Expandable grids for visualizing and authoring computer security policies. In *Proc. CHI* (2008), 1473–1482.
18. Riva, O., Yin, Q., Juric, D., Ucan, E., and Roscoe, T. Policy expressivity in the anzere personal cloud. In *Proc. SOCC* (2011), 14:1–14:14.
19. Salmon, B., Schlosser, S. W., Cranor, L. F., and Ganger, G. R. Perspective: semantic data management for the home. In *Proc. FAST* (2009), 167–182.
20. Seltzer, M., and Murphy, N. Hierarchical file systems are dead. In *Proc. HotOS* (2009), 1–1.
21. Sigurbjörnsson, B., and van Zwol, R. Flickr tag recommendation based on collective knowledge. In *Proc. WWW* (2008), 327–336.
22. Smetters, D. K., and Good, N. How users use access control. In *Proc. SOUPS* (2009), 15:1–15:12.
23. Squicciarini, A. C., Sundareswaran, S., Lin, D., and Wede, J. A3p: adaptive policy prediction for shared images over popular content sharing sites. In *Proc. Hypertext and Hypermedia* (2011), 261–270.
24. Voida, S., Edwards, W. K., Newman, M. W., Grinter, R. E., and Ducheneaut, N. Share and share alike: exploring the user interface affordances of file sharing. In *Proc. CHI* (2006), 221–230.
25. Vyas, N., Squicciarini, A. C., Chang, C.-C., and Yao, D. Towards automatic privacy management in Web 2.0 with semantic analysis on annotations. In *Proc. CollaboateCom* (2009), 1–10.
26. Zubiaga, A., Körner, C., and Strohmaier, M. Tags vs shelves: from social tagging to social classification. In *Proc. Hypertext and Hypermedia* (2011), 93–102.