Exploring the Ownership and Persistent Value of Facebook Content

Catherine C. Marshall and Frank M. Shipman
Center for the Study of Digital Libraries and Department of Computer Science
Texas A&M University
{ccmarshall,shipman}@cse.tamu.edu

ABSTRACT
In this paper, we present the results of a study examining 244 participants' attitudes about the value, ownership, and control of social network data. We use Facebook-based scenarios to elicit reactions to hypothetical statements about saving social network content that belongs to others, reusing, repurposing, and monetizing social network data, and removing social network content that is not specifically one's own. Participants also report on their own practices in each of these areas. Findings not only address issues related to ownership, but also explore the use of social networks as documentary records, and the discrepancies between participants' perceptions of how they would like their social network content to be used, and how it is actually used.

Author Keywords
Social networks; social media; saving; reusing; removing; monetizing; digital identity; documentation; records.

INTRODUCTION
Facebook has become a fixture in many people's online lives. Its social networking capabilities have made it a widely accepted and broadly used multi-purpose communication and content-sharing utility [38,43]. It may be a means of keeping a relationship alive [36], a way of seeking information from trusted sources [15], or a place to share emotionally resonant content [44]. Even when it is not foregrounded, many people see Facebook as an implicit constant in their online life and an important means of communication and contact maintenance [11].

As Zhang et al. point out, the online world is still highly volatile [43]. Even well-established services may suffer reversals of fortune. In the face of this volatility, we test participants' attitudes toward the data they have published or shared on Facebook and document users' attitude-shaping practices. We focus on Facebook users and Facebook data, because the service is well-known and fairly stable: participants are likely to have incorporated it into their everyday online lives and to have some investment in content they have shared, social ties they have represented, and profiles they have developed.

In spite of Facebook's ubiquity and entrenchment as a social networking utility, in a recent study, Lindley et al. found that Facebook content was perceived to be of dubious long-term value [19], and, consonant with Zhao et al., that people deliberately curate an ahistorical digital presence in an effort to keep their self-representation current [45]. Follow-on work by Zhao and Lindley revealed that reuse of Facebook content in crafting exercises can change the perception of the material's value [44]. Thus we find ourselves asking, what are the sources of value in social networks, beyond their immediate communicative function? What do people feel they own and control in a co-constructed medium?

In this study, we investigate intertwined notions of value, ownership, and control of social network content. Specifically, we use scenarios and hypotheticals, a technique borrowed from legal education and case-based reasoning [28,32], to identify social norms associated with saving, removing, and reusing social network material in which they have a stake. We test concepts such as the effect of social distance, the difference between individual profiles and shared content, and whether personal data should be used for monetary gain (and by whom). We also ask participants to report on their recent use and curation of social network accounts.

Specifically, we are interested in discovering what people feel they can do with social network content they do not explicitly own, whether social network data can act in the role of archival records (in other words, the reciprocal question of how corporations and public institutions can use individuals' social network content), and the degree to which user-perceived rights deviate from legal reality.

We begin this paper by reviewing several areas of related work. We then describe the method we used to construct the study, recruit participants, and collect data; we also discuss some limitations on our results. Reports of current practice are woven together with quantitative responses to the hypotheticals, organized by the questions they answer. Finally we discuss the answers to our going in questions

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
CSCW '15, March 14–18, 2015, Vancouver, BC, Canada.
Copyright 2015 ACM 978-1-4503-2922-4/15/03…$15.00
http://dx.doi.org/10.1145/2675133.2675203

about ownership and control, as well as exploring the use of social networks as data.

RELATED WORK

Our research is primarily motivated by two confluent strains of related work. One stems from social media archiving, and is directed at answering the question, is there persistent value to the material that is created through and stored in Facebook? To answer this question, it is essential to understand the multiple ways in which Facebook is used, and what contributes to content volatility. The second has to do with ownership, and addresses the question of what people believe they can and should be able to do with co-created social media content. Tied in with these themes are three curation-focused areas: digital identity, privacy, and self-censorship, including questions arising from the deletion or removal of online content.

Social networks and value

Much of our digital identity is formulated and performed online in social media services such as Facebook [45]; in the name of keeping digital media current and digital identity closely curated, personal history is continually renewed and paradoxically destroyed [19]. But Lindley and her colleagues also discovered that users harbored skepticism about the long-term value of the content they have stored in Facebook [19]. To reinvigorate a sense of the content's worth, Zhao and Lindley found that if people crafted something from the content, they became more aware of its value [44]. Gulotta et al. are similarly investigating how value might be explicitly introduced by means of technologies designed to communicate this aspect of their digital possessions [9]. We are exploring value from several standpoints, both archival (value to others and value as historical records [8]) and pragmatic (what people would take with them if they moved to a new service, and whether they are comfortable with the content being monetized by corporate interests or by themselves).

Ownership of co-created social media

The flip side of content value is content ownership. If the content and the social network on services like Facebook has personal or professional value, then we can look at how this content and the social network can be used and reused by others. Previously, we have performed a series of studies looking at the social norms for online content reuse [22, 23, 25, 34, 35]. Generally, we have found a lax attitude toward saving online material, but a much more varied reaction to its reuse, attitudes which may diverge dramatically from law and policy, and which vary by media type and reuse context. Here, we are specifically investigating the reuse of the social graph in other settings.

Curating digital identity

Of course, much of the tension between what is current on services like Facebook, and what is of archival value, stems from a desire to create a robust digital identity. The need for self-representation drives many current identity practices, including a need to manage one's reputation (often seen as frustrating) [42, 40] and to use social media to potentially manipulate search engine results [21]. Farnham and Churchill have shown that people integrate or separate different facets of their online identities to varying degrees [7]; furthermore, sharing and deletion are guided by a sense of audience and social distance [37, 14]. Rainie et al. have established that people that people come and go from Facebook, in part due to curatorial fatigue [31]. These are practices that we used to guide our survey design.

Deletion (and the concomitant act of forgetting) is a valuable part of curation [27]. This may be situation-specific (e.g. after a breakup [33]), or it may be a general part of indicating what's current and what's not of one's online presence [19]. We ask about specific instances of removal to investigate the balance of fixity and fluidity of different portions of one's social network presence, and as a countervailing measure of value.

Finally, privacy is an important influence on many of the behaviors we examine here, especially when it comes to corporate monetization of personal data. Although there are efforts to help users understand their potential exposure [41], some Facebook users rely extensively on hiding in plain sight to achieve privacy through obscurity [13, 38], and attempts to monetize their data may make them feel more exposed (even if distinctions between private and public data follow user privacy settings).

METHOD

This study is implemented as a three-part questionnaire: first, participants responded to 12 demographic questions to characterize themselves. Next, a series of scenarios and 25 associated hypothetical actions were rated along a 7-point Likert scale (from *disagree strongly* to *agree strongly*) to help find the limits of what participants perceive as current social norms. Finally seven multiple choice and open-ended questions elicited participants' reports of their current social networking practices. Two comprehension questions were interspersed among the hypotheticals to help ensure that participants read carefully [15]. A final question asked participants if they would take another survey like this one. The 47 question survey was designed to be completed fairly quickly and to hold participants' interest.

The questionnaire was fielded as an Amazon Mechanical Turk (AMT) HIT. We have refined this technique over the course of four years through experience [24] and the careful application of best practices from the literature [10, 15]. We recruited workers who have at least a 95% acceptance rate for their past work, and pay them as long as they complete the questionnaire regardless of whether we use their data or not. We are conservative about the data we keep. A demerit system means that we discard a participant's data if two or more checks are violated, including: a wrong answer on a comprehension question; no answer or a nonsense answer

to an open-ended question; overly brief work time (less than 50% of the average); or suspicious patterns in the Likert scale data (e.g. all responses are the same). In spite of these stringent criteria, we only had to discard data from 6 participants out of the 250 questionnaires administered, leaving us with data from 244 credible participants.

Participants

Current and recent US-based English-fluent Facebook users were recruited to participate in the study; we want participants to be able to understand the scenarios, and respond by reflecting on what they would do, given the situations posed in the hypotheticals. Triangulation among the initial participation requirements, a question in the demographic section of the survey, and another question in the practice section allowed us to verify that all 244 participants were indeed regular Facebook users (2 had recently deleted their accounts; the other 242 had one or more active accounts). Often participants used other social networking services too, including LinkedIn, Google+, Twitter, Instagram, Tumblr, Myspace, and Reddit.

Participants were active online, reporting a median of between 5 and 6 activities. Almost all used email and shopped online; nearly 70% shared videos. Most respondents (68%) had used the Internet for a decade or more.

In keeping with a trend of increased male participation on AMT, 52% of participants reported to be male. Consonant with prior surveys, almost half (48%) were born in the 1980s (although 13% were born in the 1960s or before and 24% were born in 1990 or later) and the group is predominately college educated (only 8% hadn't been to college). At 60/244 (25%), student participation is lower than usual, close to the lowest we have seen in our past AMT surveys. Thus participants in this study represent a population that is fairly young, well-educated, and Internet-literate, but not necessarily technical, nor invested in any particular technological outcome.

Scenarios and hypotheticals

Hypotheticals are a technique borrowed from legal education [28] and case-based reasoning [32]. We describe a basic situation with concrete details that we hope will be engaging for survey participants and will ground their reasoning in similar assumptions. Then a series of hypotheticals is proposed in which specific aspects of actions are varied to test which responses 'fall off a cliff'. For example, we might describe a scenario in which a character saves her profile to her local disk (say, because she intends to delete her account). Then we might propose that instead of just saving her own profile, she saves *some friends' profiles* locally. From there, we might propose that she saves *her friends' friends' profiles* locally, and finally that she saves a stranger's public profile to her local disk. In these hypotheticals, we vary the details in such a way that they are not predictable, thus circumventing the use of any heuristics or shortcuts in reasoning that participants might employ if they notice that the hypotheticals vary too regularly.

We develop the scenarios using situations we have encountered in previous qualitative studies, in published accounts, or in informally observed everyday practice. From there, we consider the features we want to test (for example, in H1-H4[1] shown in Figure 1a, the feature of interest is the effect of social distance on perceived ownership). We then construct a series of statements, changing aspects of the feature. We iterate over the design and language of the scenarios and hypotheticals, since it is important that they are relatable and understandable. Outside reviewers help us identify possible ambiguities in the language or word usages that might lead to misunderstanding. Monitoring early results (especially via the comprehension questions) is a final fail-safe to ensure participants are reading and understanding the scenarios as we intended.

Naturally, because they seek edges, the order in which the hypotheticals are presented can be important; the idea is to start with a banal premise and vary a feature until the statements become controversial. Even if we misjudge the order (for example, it might be more controversial to save one's friends' friends' private profile than a stranger's public profile), it is apparent in the data where we have violated social norms. Out of order statements are not necessarily harmful; they simply highlight "edges" where we did not expect them. Occasionally, to maintain participant interest, we ordered the statements from most controversial to the least. This has the effect of gradually backing off from controversial ideas until changes render them acceptable.

Analysis

The resulting data requires several types of qualitative and quantitative analysis. The Likert-scale data from parallel hypotheticals is compared using the Wilcoxon signed-rank test (a non-parametric test for comparing related samples) to handle the non-normal distributions. Responses in each category are graphed together to visualize trends. Hypothetical grouping generally reflects themes in the study: saving, removing, and reusing social network content either by moving it to a new service or monetizing it. Responses to open-ended questions about social network use have been open-coded using conventional qualitative methods to reveal primary and secondary themes; where possible, we have triangulated among answers to ensure consistency [38].

[1] In this paper, we use the notations H1 through H16 to refer to specific hypotheticals we tested in the study.

Limitations

Any recruiting technique and study population will bring with them inherent limitations on the reach of the results. We note several such limitations on our results. First, we are recruiting participants who may do surveys regularly and who have an interest in the fate of social media. Thus, they may have strong, well-developed opinions about the topic. Second, although the participants are not necessarily of a single socio-economic type [24], they are younger, better educated, and more technology-conversant than a random sample of the US population would be. Yet we were interested in reaching a diverse group of people who had familiarity with and investment in Facebook; we found the limitations introduced by using the Mechanical Turk platform to offer a reasonable trade-off for access to these users. Similarly, we consciously chose to limit our population to current Facebook users; it might be interesting in a future study to reach non-users, since policy questions necessarily affect them too.

We also recognize limitations with this method; we see it as validating and extending the results of earlier interview studies such as [45] and applying the methods used in [25,35] to extend those results to different content types. Accessing real social network data can further validate some of our claims (e.g. identifying unusual uses or which content types are most frequently removed). Although the scenarios are modeled on real-world behavior and have been carefully debugged, we recognize that some participants may not reach the expected interpretation of the situations that are the basis for the hypotheticals.

RESULTS

In this section, we bring together qualitative and quantitative results to explore four primary questions about ownership and value in social networks that we addressed via open-ended questions and hypotheticals.

The first question is connected with saving content. Past studies have shown that saving most types of content locally is primarily sensitive to limits imposed on the action. In other words, participants mostly react when they are told that someone hypothetically *can't* save something, rather than that they can. However, we suspect that some social network content—profiles, for example—may be different because they are so closely associated with identity. How much effect does social distance have on what a person can save? Does content type matter?

The second question stems from removal. Again, because social network curation is so closely associated with digital identity, will effects identified by past research be more pronounced? For example, do people feel they should be able to remove Facebook posts *about* them? Is permission necessary? Prior research has shown that it is important to allow people to untag Facebook posts about themselves [20]. They are concerned about the content's veracity and timeliness. We explore this idea by asking participants what they have removed from Facebook recently and what they would have liked to remove (but couldn't).

The third question pertains to reuse: many people are aware that they consent to a license agreement when they sign up for services like Facebook. They may even be aware that this license grants the service some ownership rights to their data. How do participants feel that this license—combined with the fact that the service is free—influences the service provider's (and their own) rights to reuse the data in a different service?

Ownership can also be seen through the lens of monetization. If Facebook owns your data, can they sell it to another company? Can users sell their data to another company? Should service providers be able to invade the more private reaches of users' personal data (e.g., private messages to other users) to help profile their interests for advertisers? We did not expose which of these data practices are currently in force; instead we asked participants what *should* be done.

Saving social networks and social network content

Saving is invariably the least controversial user action; many people feel they can download material from Internet services such as Flickr and save it with impunity, especially if they have no plans to share it further. But are all of the components of social networks perceived the same way, even those that are more tightly connected with digital identity? Not only do social networking services like Facebook support media sharing, they also encourage users to create rich profiles of themselves and represent their social networks. In turn, people see their profiles as something they have published just as surely as they have published more conventional content such as blog posts and photos [24]. The hypotheticals tested whether participants felt the same way about saving other people's profile content as they did other people's photos, and whether they felt that their social distance from the person mattered.

Susie has been using Facebook since 2005, when she was a college freshman. She has accumulated hundreds of friends, photos, 'likes', posts, and mail. She would like to delete her account, but keep her connections and the content she has posted. Before she deletes her account:

H1: Susie should be able to save her Facebook profile, including who's connected to whom in her network of friends. **(H1: User can save profile and friend list)**

H2: Susie should be able to save contact information for her friends' friends. **(H2: User can save friends' friends)**

H3: Susie should be able to save the content associated with her account, including the pictures in her galleries and her message box (including email to/from her). **(H3: User can save account content)**

H4: Susie should be able to save anything that she can ordinarily reach on Facebook, including pictures from her friends' and friends' friends' galleries. **(H4: User can save all encounterable)**

H5: Susie should only be able to save photos from her friends' Facebook accounts if the photos are public. **(H5: User can save only public photos)**

Figure 1a. Scenario and hypotheticals to test features of saving

In the initial demographic section of the questionnaire, we asked participants the extent to which they had fleshed out their profiles. 71% felt they had normal profiles and 19% reported profiles with more details than normal. Only 10% had minimal profiles. Furthermore, 70% reported their profiles were either stable or growing. Although 30% reported reduced use of the service over time, by these results and others [31], participants seemed to leave their profiles intact. Thus participants were likely to be able to relate to the user described in the scenarios who has had a Facebook account since she went to college, and has a large social network with plenty of accumulated content. In the first set of hypotheticals, our user is now seeking a grown-up job and plans to delete her Facebook account for fear a prospective employer will scrutinize it closely. What can she save locally from her Facebook account? Figure 1a shows an abbreviated version of the scenario and hypotheticals (the labels used in Figure 1b are shown in bold), and Figure 1b shows the results of testing 1a's five hypotheticals.

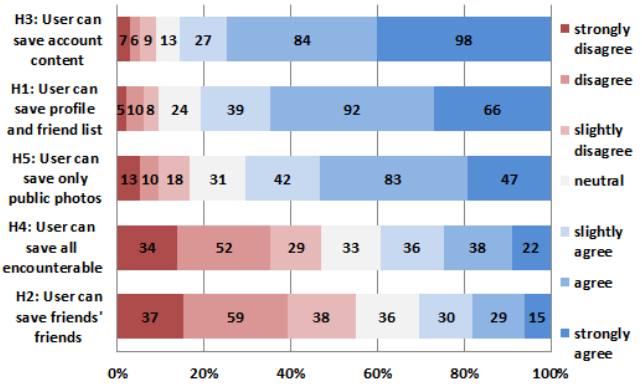

Figure 1b. The results of testing the saving hypotheticals

In this scenario, we are exploring three factors: the effects of social distance, the influence of privacy settings (since saving does not immediately compromise the owner's intent), and whether a profile is similar to (or distinct from) other types of content. Figure 1b shows the results. The strong agreement with statement H3 shows that participants see ownership extending to all of a user's personal content, including their profile. That H1 is less positive (but still positive overall) shows that users distinguish between the ownership of their own profile, and the interconnections among their friends: yes, they can save them, but there is reduced expectation of ownership (Wilcoxon p=0.005).

H2 increases social distance (friends of friends), and demonstrates that social distance strongly influences the acceptability of saving. H5 tests the idea that public content is more loosely owned than private. Indeed, participants generally felt it is okay to save public content, far more so than the comparable private content: being able to *reach* content in compliance with a privacy setting does not mean you can *keep* it (but if it is public anyway, you can). H5 is strongly significantly different from H1 (Wilcoxon p=2.26E-08) and H4 (Wilcoxon p=3.77E-15).

Removing social media content

Social network content is fairly fluid relative to other material published on the Internet [18,12]. Users add and remove content regularly. Removal is a significant manifestation of how ownership plays out through identity curation: it may be just as important to manage one's identity by taking things down as it is by adding new content and self-description. But is all removal equivalent? Is removing a photo from Facebook similar to removing a photo from elsewhere on the web? Or are there other concerns when the content is only visible to friends or acquaintances? Is removal of a photo similar to removal of a profile detail like a birthdate? As Das and Kramer found through their large-scale analysis of self-censorship on Facebook, removal is common: 71% of Facebook users in their sample had removed content [6]. We found self-reports of removal to be similarly common: 191/244 (or 78%) reported removing some portion of an account, including the whole thing. Table 1 shows responses to an open-ended question about what participants have recently removed from Facebook.

Content type removed	#	%
Photos	89	36.5%
N/A or nothing	53	21.7%
Posts/comments	53	21.7%
Profile elements	21	8.6%
Unspecified or 1-off	20	8.2%
Entire account	11	4.5%
Friends	8	3.3%
Links	6	2.5%
Game- or app-related auto posts	6	2.5%

Table 1. Recent episodes of content removal by type

There are four primary reasons participants offered for having removed Facebook content: (1) to curate material and keep accounts tidy (*I hard-deleted my old Facebook profile in August and created a new one in its place. I had too many photos and too many friends, and I wanted a 'fresh' profile/account*. [SN193]); (2) to maintain one's online reputation, either directly (*I removed several alcohol related pictures when I started to look for teaching jobs*. [SN024]) or indirectly, for fear an unpopular item will reflect poorly on the participant (*I removed one of my songs that I posted because I didn't have many likes on it…* [SN169]); (3) to reflect a revised understanding of privacy (*[I] removed [my] birthday as it is often a question on a multifaceted security question and answer* [SN011]); and (4) to respond to evolving knowledge about Facebook itself (*After it came out that my content was being monetized with no reimbursement to me, I deleted my entire account*. [SN061]).

Participants were fairly sophisticated about ways to remove unwanted content: not only did they delete items, they also modified privacy settings, untagged posts and photos, and

sometimes replaced entire accounts to start over. Participants also managed "likes" and comments, especially since they can appear unexpectedly in search results [21].

We also asked participants what they wish they could've removed; this answer may reflect the character of items that they would ask to be withdrawn from an institutional Facebook archive [26]. Responses indicated that participants mostly felt in control of their social network content; 56% (137/244) said they have no need to remove anything beyond what they are capable of removing already. Thus it's not so much that participants always posted the right thing, it's simply that they feel capable of removing Facebook content they don't want to keep.

A few exceptions are worth noting. The most common is photos; almost one quarter (58/244) say they would have liked to have removed one or more photos, usually because a photo was unflattering; a few represented perceived privacy breaches (*I don't like being mentioned in others photos or having people with lower security than me able to post photos of my kids without having to ask my permission.* [SN166]). In most reported incidents with posts (26/244), removing the post wouldn't have done any good: *I was working for GNC and posted some things on twitter and didn't have my account locked. Needless to say, I was fired for the comments*. [SN227] Profile elements are less frequently a concern, because participants are fairly conscientious about their own privacy settings, although they reminded us that privacy settings may be circumvented via social coercion. SN088 reported, *For a background interview I thought an investigator was only going to search for my profile from a public computer but he actually forced me to login on his own computer....*

Hypotheticals tested four aspects of removal: removal of different constituent media types (e.g. photos versus posts), removing solo pictures versus group pictures, removing photos one has taken versus removing photos of oneself, and whether permission mitigates the other hypothetical conditions. Figure 2a summarizes these hypotheticals.

When Susie searches for her own name, she discovers that her Facebook friend Greg has mentioned her name in conjunction with a political protest:

H6: As part of her clean-up effort, Susie should be able to remove (not just untag) her friend Greg's post about her role in the political protest. **(H6: Susie can remove Greg's post)**

H7: Susie should be able to remove Greg's photo of her at the political protest, especially since there's no-one besides her in the photo. **(H7: Susie can remove Greg's photo of her alone)**

H8: Susie should be able to remove Greg's photo of her at the political protest regardless of who else is in the photo. **(H8: Susie can remove Greg's photo with others)**

H9: Susie should only be able to remove Greg's wall post about her role in the political protest if he gives her his permission. **(H9: Susie can remove post with Greg's permission)**

Figure 2a. Scenario and hypotheticals to test removal features

Because text elements (e.g., posts and profiles) seem so consequential—unlike photos, they remain vulnerable to search engines—we wondered if participants would be more sympathetic to our hypothetical job-seeker's desire to remove a post that might impede her job search. Figure 2b shows the outcome. A substantial proportion (65%) saw permission as a way to resolve uncertain ownership: H9 is viewed less disapprovingly than H6 (the same action without permission), Wilcoxon p=.007. The two photo hypotheticals, H7 and H8, test whether the outcome is influenced by reducing potential ownership conflicts—should Susie be able to remove a photo if she's the only one in it (and sees it as harming her reputation)? It does seem to matter whether she's the photo's only subject: the difference between the two hypotheticals is significant Wilcoxon p=3.24E-18.

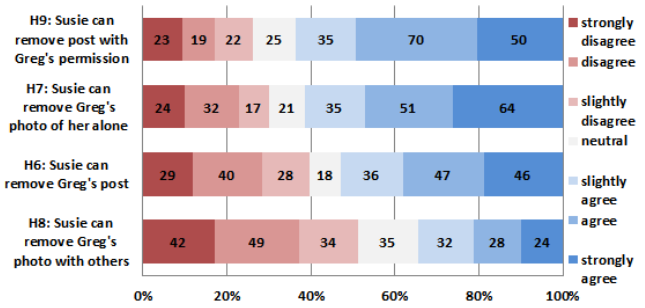

Figure 2b. The results of testing the removal hypotheticals

Reusing social networks

It is fairly easy to come up with realistic scenarios for saving or removing social network data. As our practice results demonstrate, participants readily identify examples of things they have saved or removed. It is more difficult to come up with reuse scenarios that are both specific to the social network setting and align with practice. So we have developed two types of reuse scenarios. The first type involves reuse of an entire account (by moving it to a new service), and the second type couches reuse of complex social network content as data, a form of social network reuse that is apt to be familiar to participants under the guise of 'big data' [1].

Reusing social network content in a different service

Re-entering one's social network data in another Facebook-like service may be seen as needlessly arduous and time-consuming, or it may be an opportunity for a fresh start. We asked an open-ended practice question about this type of reuse: "*If you were to move your social network (say from Facebook to Google+), what would you expect to take with you?*" Participants interpreted this question in several ways: Some thought we were asking what they wanted to take with them (i.e. what was valuable to them); others thought we were asking what they would be allowed to take with them (i.e. who owns the content). Still others speculated on the logistics of taking social network data with them: How would they be able to move this stuff? Would their friends be on the new social networking service too? It is likely that

most participants thought of the question as a mix of all three, since they are difficult to usefully separate: you wouldn't care about moving something you don't value or something that is meaningless in the context of the destination service, such as a friends list if your friends aren't on the service.

The responses were coded to reflect how participants framed their answers. If participants enumerated content types, each was tabulated. Categories that subsumed others (e.g. "everything") did not contribute to counts of individual items; rather we tried to capture the general intent and emphasis of the answers.

The elements that the participants specified seem to depend on Facebook's overall role and value. If the service was used primarily as a communication tool, participants might view the accumulated content as transient and not worth moving (except possibly contacts). On the other hand, if Facebook acted a repository for heterogeneous personal information, participants might focus on moving the valuable portion of the content, e.g., photos and messages.

With those constraints in mind, Table 2 shows that over half of the participants (131/244) specifically wanted to take their photos with them (the count would be 155/244, over 63%, if "I want to move everything" answers are included). Some participants had considerable investment in visual media: *I have over 500 videos and thousands of photos on Facebook if I were to move I would expect to take those with me*. [SN217] They felt ownership of photos they'd taken or were featured in: *[I'd expect to take]… tagged posts and photos containing my name*. [SN191] A few anticipated a fight: *Any content I create and I purposely only upload photos with watermarks on them so even if facebook steals a photo everyone knows it*. [SN061]

Content type	# of responses	%
Photos	131	53.7%
Contacts/friends	85	34.8%
Nothing	50	20.5%
Profile information	39	16.0%
Everything	24	9.8%
Statuses	14	5.7%
profile photos	9	3.7%
Messages	9	3.7%

Table 2. What participants expect to take with them from Facebook, given the ability to select multiple options

Facebook's social network is also important to participants. Eighty-five participants specifically referred to wanting to take their contacts or friends with them (again, folding in the "everything" answers brings the count to 109/244, or 45%). A substantial portion of those who thought of Facebook as a communication tool explained that it wasn't important to move their contacts; they would be able to rebuild what they had on Facebook, especially since only a portion of their friends are current: *…A good bit of my contacts on Facebook I am no longer in contact with are not necessary*. [SN182]

Given the results of other explorations of sources of value in social networks [5,9,19,45], it is unsurprising that participants were less interested in moving content such as wall posts and messages. Profile information seems more valuable, if simply because it's time-consuming to rebuild it (and 'likes' may represent significant curatorial effort); still, only 39/244 (16%) singled out their profile as something they wanted to move, and 9 others (less than 4%) were specifically interested in retaining their profile photos. Similarly, only 9 participants (<4%) specifically said they wanted to move their messages; we might expect more participants to call these out if they were an important part of their Facebook assets.

Finally, we can look at the extremes, participants who said they expected to take everything with them and participants who said they expected to leave it all behind. Although only 24 participants (10%) said they expected to be able to take everything, many of those responses were the most passionate: *Everything, All of my photos and photos that I have been tagged in. All of my mail and messages. All of the comments on my pictures and things that I have shared. All of my links that I have shared on my wall since joining. Anything else on my wall others have shared, pictures. links ect... And all of my game activity and causes information*. [SN188] Several respondents had thought about things like their game activity and curatorial work (again, 'likes'). For some, this aspect of Facebook can represent a significant investment of effort: there is evidence that you are what you like [16]. A few even thought that the weight of the content was enough to keep them from moving: [*I would want to take*] *everything or I would not make the move...* [SN011] Of those who wanted to take everything, several explicitly referred to the content's archival value to them.

Twice as many participants were at the other extreme: 50 thought a move would offer a fresh start (*I would hope to start from a clean slate* [SN068]). Others thought it wouldn't be hard to rebuild their social network and that Facebook owned their content anyway (*…it's not their job to work with competitors* [SN205]). SN203 returned to the communication tool theme, saying that he would expect to take *Nothing if I can't remeber [sic] your name your [sic] not that important to keep*.

After she has a new job, Susie thinks it is safe to have a social network presence again. Following her closest friends, Susie decides to move to a new social networking service, Circles.

H11: Because Circles can automatically rebuild a person's social network from their profile on a different network, Susie can use her old Facebook profile (including links to her friends) to move to Circles. **(H11: User can move profile to new service)**

H12: Because Facebook technically owns the information she stored there, Susie cannot use her old Facebook profile to automatically build a new profile on Circles. **(H12: User cannot move profile to new service)**

Figure 3a. Testing the limits of what a user can remove with and without permission

Given this range of responses to the open-ended question, we might expect this varying perception of value to temper participants' responses to the two hypotheticals shown in Figure 3a, testing different justifications for being able to move (or not being able to move) social network data from one service to another. The first offers a pragmatic justification for reusing Facebook data, and the second reminds participants that the license agreement they signed might forbid this type of reuse.

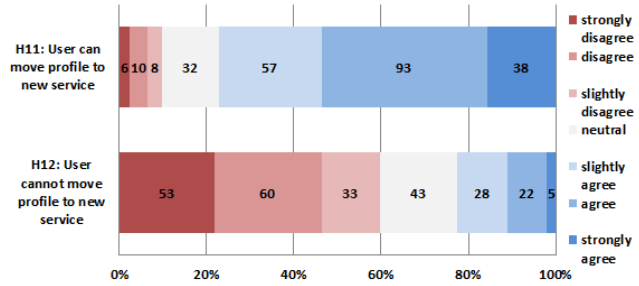

Figure 3b. The results of testing the two social network reuse hypotheticals

Of course, people who agree with the first hypothetical of the set don't necessarily have to disagree with the second. But many of them did. The majority of participants thought they should be able to move their social network data to a new service, and that the license agreement shouldn't prevent this type of reuse: their friends are their friends, not Facebook's (Wilcoxon p= 2.04E-22). This result makes sense when we refer to the practice data: only about 20% were willing to start fresh on a new social network.

Reusing social network content as data.

A second reuse scenario and set of hypotheticals explore the monetization of Facebook data by Facebook and by the user. Because many participants described themselves as having extensive Facebook profiles and user-contributed media collections, it is interesting to pose them as assets that might be valuable to another corporation. Figure 4a shows an abbreviated version of the data reuse scenario and hypotheticals.

Facebook's business model begins to falter. The company needs to develop additional ways to make money. Greg has been using Facebook for many years and has accumulated lots of data.

H13: Facebook should be able to sell the information in Greg's user profile to Amazon so Amazon can create a better profile of Greg's interests. **(H13: FB can sell info to Amazon)**

H14: Facebook should need Greg's permission to sell his profile. **(H14: FB needs permission to sell info)**

H15: Greg should be able to sell the information in his Facebook profile to Amazon to get a cash rebate on his Amazon purchases. **(H15: User can sell FB profile info to Amazon)**

H16: Facebook should be able to analyze the content of Greg's Facebook-internal communication so it can create a better profile of Greg's interests for its own use in selling targeted advertising. **(H16: FB can use internal messaging to improve ads)**

Figure 4a. Testing ownership and monetization

Figure 4b shows participants' reactions to the hypotheticals. Notice that requiring full permission before transferring the data (H14) is the only hypothetical that elicits a strongly positive reaction. But we know from other researchers' practical experience that permission is no panacea: a games archiving project made this assumption (that permission would be granted if requestors offered a polite explanation of their motivation); instead of granting permission, players often denied the request [30]. Given that Facebook data is likely to be personal too, it may be difficult for users to relinquish control of the data without reviewing it in detail.

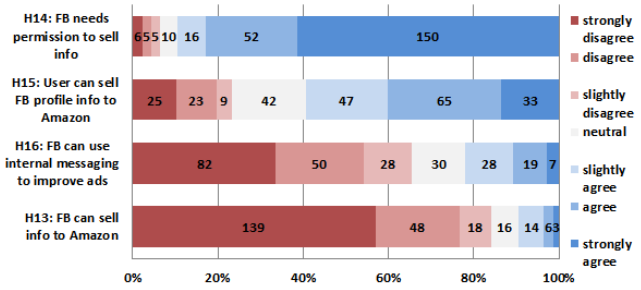

Figure 4b. Results of ownership and monetization hypotheticals

H15 suggests that Greg has sufficient control over his own data that he can sell it to another company (Amazon). This hypothetical is greeted with less enthusiasm than the first (Wilcoxon p= 1.97E-22). But taken together, the responses to both H14 and H15 suggest that users feel they should be able to exert a fair degree of control over what they have put into Facebook regardless of what they agreed to in the service's license agreement.

H16 mimics many modern social media services' personalization strategy: textual message content is analyzed to improve the match of advertising with users' interests. Nearly 75% of participants found this analysis objectionable, in spite of the fact that the participants probably use services that have already adopted this practice (H16 is strongly significantly different than H15, Wilcoxon p= 1.27E-19).

H13, which is based on the premise that Facebook would sell its data assets to Amazon, was the most objectionable hypothetical in our participants' eyes. Over 80% saw this hypothetical as negative, and over 55% thought it was very negative, again strongly significantly different from H16, in which Facebook reused the data internally for commercial purposes (Wilcoxon p= 2.72E-10). In the scenario, Facebook was foundering. Their only asset is the user data they have collected over the last decade. Yet participants were unhappy with the idea that the company might try to monetize this data. This result, taken together with the last set of hypotheticals, suggests that participants might be unhappy if Facebook decided to donate its data to an archive: it does not seem to be solely the monetization of personal data that participants are reacting to. Instead they

may be objecting to the idea that if Facebook can use their data as a corporate asset, they are no longer in control of their personal data, and by implication, their own digital footprints [26].

DISCUSSION

What can we learn from participants' attitudes toward heterogeneous social network content? Certainly activities in services like Facebook shed valuable light on design. For example, Wang et al. have discussed the design implications of Facebook practices for privacy protection features [41] and Sas and Whittaker for removing content in the specific circumstances of a breakup [33]. But we believe that the social norms that are expressed through attitudes and practices should have important bearing on law and policy too.

Although practically speaking, ownership provisions are specified by a service's terms of use, users are seldom cognizant of what they consent to: at the outset, when they create an account, they are not able to foresee how they will use it, whether they are going to share five photos or 5,000 photos. They don't know who will look for their account or how it will be interpreted. Nor can they predict how and whether the service will sustain itself over the years: a company may discover alternate uses for user-contributed content or the service may be acquired. Spammers may find convenient ways to repurpose the data too. Users sign license agreements to gain access to a service; they often neither read the terms of use, nor consider their implications for content they have yet to post.

Thus in this study we have focused on social norms and accepted behavior—what participants think is okay in the face of specific circumstances—rather than on licenses and legal interpretations. Furthermore, we are able to compare attitudes and practices distinctive to Facebook data with attitudes toward other media types and user-contributed content in different services, including multiplayer online games and video-sharing sites. We will focus on three such distinctions: (1) the content users encounter on Facebook may be owned by someone they have an explicit relationship with; (2) Facebook accounts are more closely associated with digital identity and self-presentation, and their use as information often leads to conclusions about people (rather than about, say, the potential earnings of a movie); and (3) the personal nature of Facebook content may lead to an especially keen discrepancy between perceived rights and reality.

Who owns our social networks?

Social networking services represent relationships more explicitly than people were previously accustomed to doing (albeit as many have pointed out, in a flattened, less nuanced way than one might like). Thus when Facebook users reuse the content they encounter there, they know whether it belongs to a friend, a friend-of-a-friend, or someone who is socially more remote. Most people are aware of the degree to which it is private insofar as they are able to encounter it.

So what do people feel they can do with the stuff they don't own? Our past studies have shown that saving public online material is uncontroversial (c.f. [22,23]). In other words, if you encounter content on the Web, people feel you should be able to save it to your computer. This effect only holds as long as the content remains under local control of the person who saved it; even storing the content on the cloud compromises one's ability to save it [25]. Social distance imposes a strong effect on whether a person can save something that is posted on Facebook. Even a friend-of-a-friend's powers are limited. This suggests that posting something in a social networking service is very different than publishing it on the public web, *regardless of whether you can reach it or not*. Ownership effects are stronger inside social networking services than they are outside, on the open web. Once individuals are in an explicitly social setting, normal social forces are once again in effect.

Social media as records

Although other types of online content may represent important portions of an individual's digital footprint, our results suggest that people's Facebook accounts are closely aligned with their digital identity and self-presentation. Furthermore, for better or worse, Facebook profiles have become a primary source of data about individuals and their social networks. Hence we may usefully think of Facebook content in personal accounts as records.

In an archival sense, a record is documentation that might be used as evidence; it may have a regular, cohesive structure; and it may suggest a certain rhythm of fixity and fluidity (for example, records may be updated annually or once in a decade) [18]. In this study, we saw a surprising concern for veracity, offset by a desire to permit self-reinvention or to create false identities (for reasons as diverse as convenience or entertainment). At least a certain portion of social network profiles are seen as an accurate reflection of individuals' characteristics—for example, a Facebook profile might be fairly complete documentation of where someone lives, their marital status, their age, where and how much they've been educated, what they like or dislike, and how to contact them.

It is not unusual for this type of informal material to act as a surrogate for more formal records: often historical research turns to ephemera that have serendipitously survived as a way of confirming hypotheses that cannot be validated by intentional sources of documentation such as censuses, tax records, medical records, or military records. Consider mid-century US city directories: they often contained not only telephone numbers, but also addresses, professions, marital status, spouse's names, neighbor relationships (who lived next-door to whom), household size, and where one worked. Yet city directories were transient, something that would be discarded or replaced annually.

When social media profiles are used this way, they reveal at least three sources of tension: (1) What is their rhythm of fixity and fluidity? (2) Can they be trusted as accurate? Who enforces their accuracy? (3) Finally, does a concern for privacy overwhelm their utility? Our hypotheticals posed preliminary tests to resolve these tensions and shed some light on the emerging norms.

Fixity. Our removal results illustrate the tension between fixity and fluidity: can we trust aspects of a social network to remain intact, or are they intrinsically fluid, changing as frequently as each of us wants to change them? Can a resource like Facebook be archived in such a way as to preserve a social network at a given time? Reported practice tells us that people mostly remove content (as opposed to profile elements), e.g. posts that didn't garner responses or seem ill-advised given subsequent scrutiny. It is less common for people to alter their profiles than it is for them to curate the content. The hypotheticals underscore this interpretation: as long as removal doesn't harm someone else's digital footprint, it is generally accepted and expected. If Facebook is seen as a repository of personal documentation, archival policy must work around this level of fluidity.

Veracity. We looked at veracity through several different lenses. First we must acknowledge that in a service like Facebook, people maintain multiple profiles; the majority of our participants did so, sometimes just to separate a professional and a personal identity, but other times for more mischievous purposes. E.g. SN165 admitted, *Like many users, I use multiple, fake profiles on Facebook for trolling purposes (messing with friends, anonymously commenting on sites, or being annoying in general when bored)*. This approach to identity is tempered by the thought that it's easy to see through this sort of behavior: *I have multiple profiles... But since both profiles can be found by anyone searching, I don't put anything inflammatory on that one either.* [SN194]. But in general, participants seemed to rely on and appreciate the veracity of certain social network data, even if we interpret our results through a filter of social desirability bias [2].

Privacy and control. Although it is unsurprising that participants are concerned with their own privacy, it is important to understand the degree to which they want to remain in control of their content, as demonstrated by their attitudes toward moving and monetizing their own accounts. A vast majority thought that Susie should be able to take her data, and bring it with her to a new service. Not only that, but when they were contemplating their own social media content, if participants thought the content was worth moving, they were less apt to let it go without objection. That almost 60% thought that users could sell their own Facebook content was a far stronger assertion of ownership than we expected (in spite of the precedent of supermarket affinity cards).

And this brings us to the way that perceived rights can—and do—fly in the face of reality.

Perceived rights and reality

The relatively personal nature of Facebook content may lead to an especially keen discrepancy between perceived rights and the real terms offered users when they sign up for the service. Because there is so much emphasis on data privacy and there is so much communication-centric use, we might expect to see a more profound sense of data ownership, and less thought given to its role as a primary corporate asset.

Although there is a certain level of awareness of governmental use and abuse of social network data, as well as countervailing initiatives to keep governments out of individuals' private data, there is a certain willful ignorance about the equivalent practices in the private sector. Not only is this data acknowledged as a valuable asset, it may also be the basis for an ongoing commercial revenue stream. Government agencies may also surreptitiously use the data (in ways other than the NSA's notorious surveillance program). A 2012 *New York Times* article revealed that, "*LexisNexis has a product called Accurint for Law Enforcement, which gives government agents information about what people do on social networks. The Internal Revenue Service searches Facebook and MySpace for evidence of tax evaders' income and whereabouts, and United States Citizenship and Immigration Services has been known to scrutinize photos and posts to confirm family relationships or weed out sham marriages*." [1] Researchers are aware of the intrinsic risks of bringing together 'big data' resources tied to identity, but there is considerable evidence that users are not [4].

CONCLUSION

In this study, we have explored comingled notions of value and ownership in social networks. Importantly, ownership within a social network is perceived differently than it is on the open web, where many people take strong cues from content type and reuse context; instead, even friends-of-friends are thought to have limited rights to the material that they can reach inside the service. Social strictures seem to play a greater role. Following from this principle, the right to remove content is mediated by the counterbalanced concerns of privacy and damage (would removal for reasons of privacy impinge someone else's desire to curate content?).

Because social network accounts are more closely tied to digital identity than other types of online content, we were interested in testing the limits of reuse, particularly reuse of social networks as data (or records). While participants were comfortable with reusing their own social network content in similar circumstances (e.g. in a similar service), controversy entered the picture when monetization or alternate uses were introduced. A significant number of participants felt they should be able to monetize their own

social network data, but the corporate service provider should not be able to traffic in this sort of personal data.

In spite of the unpredictability of social networking services—their use and the nature of their content shifts over time—we feel that it is important to understand the content's larger social role: will it be the basis of a large scale archive? How can corporations and governments use it? What effects will it have on other aspects of life (such as personal legacies)? Although these questions are being resolved in *de facto* ways (for example, by economic forces), we feel that understanding social norms may be crucial to informing more carefully designed policies.

ACKNOWLEDGMENTS

This work was supported in part by National Science Foundation grant DUE–0938074. Cathy Marshall was at Microsoft Research Silicon Valley when this research was performed.

REFERENCES

1. Andrews, Lori. Facebook Is Using You. *New York Times*, February 4, 2012.
2. Antin J. & Shaw, A. Social desirability bias and self-reports of motivation: a study of Amazon Mechanical Turk in the US and India. *Proc. CHI'12*, 2925-2934.
3. Backstrom, L., Boldi, P., Rosa, M., Ugander, J., & Vigna, S. Four Degrees of Separation. (Version 5 Jan 2012) http://arxiv.org/abs/1111.4570.
4. boyd, d. & Crawford, K. Six Provocations for Big Data. Oxford Internet Institute, September 21, 2011.
5. Burke, M., Kraut, R., Marlow, C. Social capital on facebook: differentiating uses and users. *Proc CHI'11*, 671-680.
6. Das, S. & Kramer, A. 2013. Self-censorship on Facebook. *Proc. ICWSM 2013*, 120-127.
7. Farnham S. & Churchill, E. Faceted identity, faceted lives: social and technical issues with being yourself online. *Proc CSCW '11*, 359-368.
8. Good, K. D. From scrapbook to Facebook: A history of personal media assemblage and archives. *New Media & Society 15*, 4 (2013), 557-573.
9. Gulotta, R., Odom, W., Forlizzi, J., Faste, H. Digital artifacts as legacy: exploring the lifespan and value of digital data. *Proc. CHI'13*, 1813-1822.
10. Jakobsson, M. 2009. Experimenting on Mechanical Turk: 5 How Tos. *ITWorld*, September 3, 2009.
11. Joinson, A.N. Looking at, looking up or keeping up with people? Motives and use of facebook. *Proc. CHI'08*.
12. Harper, R., Whitworth, E. & Page, R. Fixity: Identity, time and durée on Facebook. In *Proc. IR 13.0* (2012).
13. Herzog, W. & Stutzman, F. The Case for Online Obscurity, *California Law Review, 101*.
14. Kairam, S., Brzozowski, M., Huffaker, D. & Chi, E. 2012. Talking in circles: Selective sharing in Google+. *Proc. CHI 2012*, ACM Press, 1065-1074.
15. Kittur, A., Chi, E., Suh, B. Crowdsourcing user studies with Mechanical Turk. *Proc. CHI'08*, 453-456.
16. Kosinski, M., Stillwell D.J., Graepel T. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences 2013*.
17. Lampe, C., Vitak, J., Gray, R. & Ellison, N.B. Perceptions of Facebook's value as an information source. *Proc. CHI 2012*, 3195-3204.
18. Levy, D.M. *Scrolling Forward: Making Sense of Documents in the Digital Age*. Arcade, 2001.
19. Lindley, S., Marshall, C.C., Banks, R., Sellen, A. & Regan, T. Rethinking the web as a personal archive. In *Proc. WWW 2013*, 749-760.
20. Litt, E., Spottswood, E., Birnholtz, J., Hancock, J., Smith, M.E. & Reynolds, L. Awkward encounters of an "other" kind: collective self-presentation and face threat on facebook. *Proc. CSCW '14*, 449-460.
21. Marshall, C.C. & Lindley, S. Motivations and Strategies for Self-Search. *Proc. CHI'14*, 3675-3684.
22. Marshall, C.C. & Shipman, F.M. Social media ownership: Using Twitter as a window onto current attitudes and beliefs, *Proc. CHI'11*, 1081-1090.
23. Marshall, C.C. & Shipman, F.M. The ownership and reuse of visual media, *Proc. JCDL'11*, 157-166
24. Marshall, C.C. & Shipman, F.M. Experiences Surveying the Crowd: Reflections on Methods, Participation, and Reliability, *Proc. WebSci'13*, 234-243.
25. Marshall, C.C. & Shipman, F.M. Saving, reusing, and remixing web video: using attitudes and practices to reveal social norms, *Proc. WWW'13*, 885-896.
26. Marshall, C.C. & Shipman, F.M. An argument for archiving Facebook as a heterogeneous personal store, *Proc. DL'14*.
27. Mayer-Schonberger V. *Delete: The Virtue of Forgetting in the Digital Age*. Princeton University Press, 2009.
28. MacCormick & Summers, (eds.) *Interpreting Precedents*, Ashgate/Dartmouth, 1997, 528-529.
29. Manjoo, F. & Yoffe, E. Resume Robbed: What to do when a colleague steals your job title and description for her LinkedIn profile. *Slate Manners for the Digital Age*, January 10, 2012.
30. McDonough, J., Olendorf, R., Kirschenbaum, M., Kraus, K., Reside, D., Donahue, R., Phelps, A., Egert, C., Lowood, H., and Rojo, S., 2010. Preserving virtual worlds final report, http://hdl.handle.net/2142/17097

31. Rainie, L., Smith, A. and Duggan, M. Coming and going on Facebook. *Pew Research Center's Internet and American Life Project*, 2013.
32. Rissland, E. & Ashley, K. 1986. Hypotheticals as heuristic device. *Proc. Strategic Computing Natural Language*, 165-178.
33. Sas, C. & Whittaker, S. Design for forgetting: disposing of digital possessions after a breakup. *Proc CHI'13*.
34. Shipman, F.M. & Marshall, C.C. Are user-contributed reviews community property? Exploring the beliefs and practices of reviewers. *Proc. WebSci 2013*, 386-395.
35. Shipman, F.M. & Marshall, C.C. Creating and sharing records of multiplayer online game play: practices and attitudes. *Proc. ICWSM'14,* 456-465.
36. Skeels, M. & Grudin, J. When Social Networks Cross Boundaries: A Case Study of Workplace Use of Facebook and LinkedIn. *Proc. Group'09*, 95-103.
37. Sleeper, M., Balebako, R., Das, S., McConahy, A.L., Wiese, J. & Cranor, L.F. 2013. The post that wasn't: exploring self-censorship on Facebook. *Proc. CSCW'13*, 793-802.
38. Strauss, A. and Corbin, J. *Basics of Qualitative Research*, Sage Publications, 1998.
39. Stutzman F. & Kramer-Duffield, J. Friends only: examining a privacy-enhancing behavior in facebook. *Proc. CHI '10*, 1553-1562.
40. Vitak, J. & Kim, J. You can't block people offline": examining how facebook's affordances shape the disclosure process. *Proc. CSCW '14*, 461-474.
41. Wang, Y., Leon, P. G., Acquisti, A., Cranor, L.F., Forget, A., & Sadeh, N. A field trial of privacy nudges for facebook. *Proc. CHI '14*, 2367-2376.
42. Woodruff, A. Necessary, Unpleasant, and Disempowering: Reputation Management in the Internet Age. *Proc. CHI '14*, 149-158.
43. Zhang, H., De Choudhury, M., and Grudin, J. Creepy but Inevitable? The Evolution of Social Networking, *Proc. CSCW 2014*, 368-378.
44. Zhao, X. & Lindley, S. Curation through use: understanding the personal value of social media. In *Proc. CHI '14*, 2431-2440.
45. Zhao, X., Salehi, N., Naranjit, S., Alwaalan, S., Voida, S. & Cosley, D. 2013. The many faces of Facebook: Experiencing social media as performance, exhibition, and personal archive. *Proc. CHI'13*, 1-10.