Personalization via Friendsourcing

MICHAEL S. BERNSTEIN
Massachusetts Institute of Technology
and
DESNEY TAN, GREG SMITH, MARY CZERWINSKI, and ERIC HORVITZ
Microsoft Research

When information is known only to friends in a social network, traditional crowdsourcing mechanisms struggle to motivate a large enough user population and to ensure accuracy of the collected information. We thus introduce *friendsourcing,* a form of crowdsourcing aimed at collecting accurate information available only to a small, socially-connected group of individuals. Our approach to friendsourcing is to design socially enjoyable interactions that produce the desired information as a side effect.

We focus our analysis around Collabio, a novel social tagging game that we developed to encourage friends to tag one another within an online social network. Collabio encourages friends, family, and colleagues to generate useful information about each other. We describe the design space of incentives in social tagging games and evaluate our choices by a combination of usage log analysis and survey data. Data acquired via Collabio is typically accurate and augments tags that could have been found on Facebook or the Web. To complete the arc from data collection to application, we produce a trio of prototype applications to demonstrate how Collabio tags could be utilized: an aggregate tag cloud visualization, a personalized RSS feed, and a question and answer system. The social data powering these applications enables them to address needs previously difficult to support, such as question answering for topics comprehensible only to a few of a user's friends.

Categories and Subject Descriptors: H 5.3. [**Information Interfaces and Presentation**]: Group and Organization Interfaces—*Computer supported cooperative work*

General Terms: Design, Human Factors

Additional Key Words and Phrases: Social computing, social tagging, human computation, friendsourcing

The authors are grateful for funding from Microsoft Research, the National Science Foundation, and Xerox.

Authors' addresses: M. S. Bernstein, Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 32 Vassar St., Cambridge, MA 02139; email: mbernst@csail.mit.edu; D. Tan, G. Smith, M. Czerwinski, and E. Horvitz, Microsoft Research, 1 Microsoft Way, Redmond, WA 98052; email: {desney, gregsmi, marycz, horvitz}@microsoft.com.

DOI 10.1145/1746259.1746260 http://doi.acm.org/10.1145/1746259.1746260

ACM Reference Format:
Bernstein, M. S., Tan, D., Smith, G., Czerwinski, M., and Horvitz, E. 2010. Personalization via friendsourcing. ACM Trans. Comput.-Hum. Interact. 17, 2, Article 6 (May 2010), 28 pages.
DOI = 10.1145/1746259.1746260 http://doi.acm.org/10.1145/1746259.1746260

1. INTRODUCTION

Crowdsourcing refers to the online recruitment of individuals to complete tasks too large or too difficult for a single person to undertake. In typical crowdsourcing applications, large numbers of people add to the global value of content, measurements, or solutions by individually making small contributions. Such a system can succeed if it attracts a large enough group of motivated participants, such as Wikipedia [Bryant et al. 2005; Kittur et al. 2007], delicious [Golder and Huberman 2005], Flickr [Marlow et al. 2006], and YouTube [Cha et al. 2007]. These successful crowdsourcing systems have a broad enough appeal to attract a large community of contributors.

However, it can be difficult to bootstrap crowdsourced systems when only a small network of people is qualified to provide information. This small-network problem can rear its head in numerous real-world settings. For example, leaders of a small start-up company may wish to organize its large intranet, or a small creative team might want to encode and upload years of old work to the web. Two key challenges arise in such small-network situations: motivating enough members of a small pool of people to participate, and ensuring the accuracy of the generated information. If a crowdsourced system like Wikipedia gets only 1.6% of its viewers to edit [Hoffmann et al. 2009], that statistic still results in tens or hundreds of editors for a given page—but when the viewership pool is restricted to the scale of tens or hundreds of individuals as in a social network, a 1.6% hit rate suggests that only a small handful of people might participate. Such information is likely to be both incomplete and unverified.

In our work, we investigate the small-network challenge of collecting accurate information about the interests, hobbies, and preferences of people from members of a socially connected group of individuals. This information can be used to personalize users' computing experiences, for example to aid question answering for topics comprehensible only to a few of a user's friends. Such valuable information is typically held by members of tightly knit groups and its collection poses challenges such as motivating a relatively small pool of people to contribute knowledge.

In such problems where the desired knowledge is held by an individual's social network, we bring social application design to bear via an approach we call *friendsourcing*. Friendsourcing gathers social information in a social context: it is the use of motivations and incentives over a user's social network to collect information or produce a desired outcome. We shall specifically take on the challenge of friendsourcing for personalization: gathering descriptive information about an individual for use in enhancing computing services for that person. In this paper we build a social tagging system by drawing on previous work in social tagging of bookmarks [Golder and Huberman 2005; Dugan et al. 2007], images [Marlow et al. 2006] and people [Muller 2007;

Fig. 1. The user has guessed several tags for Greg Smith, including *band, poker* and *stanford*. Tags guessed by Greg's other friends are hidden by dots until the user guesses them.

Farrell et al. 2007]. We adapt elements of Games with A Purpose [von Ahn 2006] and extend their design principles using social controls. While previous work has explored people tagging [Farrell et al. 2007], broadening this beyond the enterprise creates challenges and opportunities around motivation, tone, and accountability issues—issues we address in our work.

We shall explore key concepts with friendsourcing applications in the context of a system named *Collabio*. Collabio, short for *Collab*orative *Bio*graphy,[1] is a game we developed to elicit descriptive tags for individuals within the Facebook social network [Bernstein et al. 2009]. The game (see Figure 1) collects information that friends know about one another, such as people's personalities, expertise, artistic and musical tastes, topics of importance, and even quirky habits. The goal is to leverage properties of the social network such as competition and social accountability to solve the tag motivation and accuracy problems within a social framework.

We will lay out the design space of social tagging games and their incentive system, describe Collabio, and evaluate the success of our design choices via usage logs and a survey of active users. We reflect on the accuracy and nature of Collabio's tags, the power of reciprocity in games like Collabio, and the impact of social controls on behavior. We recruit raters to explore how Collabio tags

[1]http://apps.facebook.com/collabio

augment ones that might have been generated through traditional online Web crawling techniques, and find that the knowledge is often unavailable through these techniques. Finally, we introduce an illustrative trio of applications that we believe would benefit from tags collected by Collabio.

2. RELATED WORK

Friendsourced applications draw on years of CSCW research supporting workgroup practice. These systems facilitate workgroup awareness [Gutwin and Greenberg 2002] and help employees find coworkers who can answer their questions [Ackerman and Malone 1990]. However, these systems typically rely on each user investing effort into updating their own status or expertise in the system, or in installing logging software to do so semiautomatically. Friendsourced applications typically reverse the dynamic and recruit a user's motivated friends rather than require the user to do work on their own behalf.

We first highlight the need for friendsourcing, as well as its potential, via research exploring social networking profiles. Facebook users are confident that their profiles portray them in accurate and positive ways [Lampe et al. 2007], and outside observers who view these profiles or personal web sites do form clear and accurate impressions of the author's personality even with extremely small subsets of the information [Vazire and Gosling 2004; Stecher and Counts 2008; Gosling et al. 2007]. However, combining personality ratings of outside observers with self-evaluations produces a more accurate picture than either the raters or the individual alone [Vazire and Gosling 2004], suggesting that the integration of friends' impressions into profiles may lead to more accurate portrayals. Furthermore, these studies do not consider the large number of Facebook users who do not complete their own profiles—Lampe et al. [2007] found that 41% of profile fields on Facebook are missing. We hope that our work on friendsourcing can result in users having more complete profiles.

Collabio builds on previous work in social tagging systems. Fringe Tagging is a social people-tagging application developed for internal use at IBM [Farrell et al. 2007], and Collabio extends Fringe's exploration of people tagging. Beyond the differences associated with being an enterprise application, Fringe takes a largely pragmatic perspective on motivating people to participate: it "enables people to organize their contacts into groups, annotate them with terms supporting future recall, and search for people by topic area" [Farrell et al. 2007]. Collabio, by contrast, is oriented primarily toward encouraging social connectedness. Both systems collect many affiliation and expertise tags, though we believe that Collabio collects a broader set of tags due to its nonwork focus.

Social tagging systems for organizing photos, bookmarks, and videos are widespread on the Web [Golder and Huberman 2005; Cha et al. 2007; Marlow et al. 2006]. Systems such as Mr. Taggy [Kammerer et al. 2009], Spar.tag.us [Hong and Chi 2009], and Dogear [Millen et al. 2006] explore the social tagging of web content. Collabio adds to the knowledge of these systems because its motivation is derived from "tagging for you" rather than "tagging for me" (Dogear, Fringe) or "tagging for us" (Spar.tag.us). Unlike other tools, Collabio taggers do not receive direct benefit by being active; instead, they hope to incentivize their friends to reciprocate and tag them. The in-place tagging and relevance feedback techniques explored in these tools could be applied to Collabio as well.

Collabio draws inspiration from other people-tagging applications on Facebook. These applications typically maximize entertainment rather than quality of tags. iDescribe[2] and Compare People[3] allow users to place predefined tags on their friends. Using only predefined tags assumes a small set of static descriptors and does not leverage the richness of knowledge in the network. In contrast, Describe Me,[4] Define Me,[5] and Impressions[6] encourage users to create new tags. However, they also allow authors to see and reapply existing tags, hence potentially skewing perception and reducing the actual richness of tags. They also keep author identities anonymous, which we believe leads to undesirable behavior since there is no real motivation to "play nice."

Our approach is inspired by prior work on human computation [von Ahn and Dabbish 2004], which aims to obtain useful information for computers by enticing users to provide it. Games with a Purpose [von Ahn 2006] recasts difficult computational problems as games for humans to play. As an example, to date, computer vision systems have been poor at the general problem of identifying items in images drawn from a large class of objects. The ESP Game [von Ahn and Dabbish 2004] asks two players who cannot otherwise communicate to try and guess matching words to describe the image. When the players agree, both players gain points, and the game has learned a label for the image.

Friendsourcing extends the design principles of Games with a Purpose to address the challenges of collecting data from a small network. Games with a Purpose typically uses points as motivators, randomly pairs players to prevent cheating, and collects information that all players know but that computers do not know. Though friendsourced applications such as Collabio do utilize game motivations such as point scores and leader boards, they lean just as heavily on social motivators such as social reciprocity, the practice of returning positive or negative actions in kind [Gouldner 1960]. Rather than anonymously pairing random players to prevent cheating, we target users within established social groups to contribute data, relying on social accountability and profile management to discourage poor behavior [DiMicco and Millen 2007]. Finally, rather than gather information common to all web-enabled humans, we directly target information that is known and verifiable only by a small social group: information about a friend [Toomim et al. 2008].

IBM's Dogear social bookmarking game shares several of these characteristics, though it is focused around Web bookmarks [Dugan et al. 2007]. In the Dogear Game, the game proposes a Web page, and players try to guess which of their contacts bookmarked the page. If the player guesses correctly, he or she gets points; if the player guesses incorrectly, the system learns that the incorrectly guessed individual might be interested in seeing the Web page. Like the

[2]http://apps.facebook.com/idescribe
[3]http://apps.facebook.com/comparepeople
[4]http://apps.facebook.com/describeme
[5]http://apps.facebook.com/defineme
[6]http://apps.facebook.com/impression

Dogear Game, Collabio gains its most valuable knowledge when players miss and guess less popular tags. The Dogear Game attempts to collect information about many individuals at a time and the information is not seen by other players; Collabio focuses the information on a single individual and makes the collected information visible in an anonymized form to the rest of the social network.

Studies of contribution in online communities motivate several design decisions in Collabio. One danger is *social loafing*: users will exhibit little effort on a collective task when they believe that others will also contribute [Karau and Williams 1993; Latané et al. 1979]. Related to social loafing is *diffusion of responsibility*: when many individuals share the responsibility for an action that one person must perform, each feels less cognitive dissonance when he or she does not act individually [Latané and Darley 1970]. Social loafing and the diffusion of responsibility together lead to the phenomenon known as the *bystander effect*. The bystander effect exists in computer-mediated communication, for example, in chatrooms where a newcomer asks a full chatroom for technical help but nobody steps forward to answer the question [Markey 2000]. These effects play out in Collabio as players try to send out mass invitations to participate, and when there is a group responsibility but little individual motivation to be the first to tag a new user.

Online contribution may also be viewed as a social dilemma. Tags in Collabio are a public good of sorts: all players benefit if others have already tagged a user. However, tagging takes time and effort, and being the first to play means that the game is less fun for you: these incentives result in a *knowledge-sharing dilemma* where everyone would benefit from the shared knowledge but nobody is sufficiently motivated to take time to share that knowledge [Cabrera and Cabrera 2002]. Knowledge-sharing dilemmas can be mediated via societal expectations of participation, realistic threats of sanction in the face of non-action, the ability to monitor behavior, and monetary incentives [Cress et al. 2006; Kerr 1999]. Collabio's approach is to offer point incentives and positive social reinforcement for contributing knowledge.

Previous work has found that individuals are likely to contribute to an online community when they are reminded of the uniqueness of their contributions, given specific, challenging goals, and helping groups similar to themselves [Rashid et al. 2006; Beenen et al. 2004]. Thus, in Collabio, we challenge individuals' (potentially obscure) knowledge of members of their own social group. Both active and loafing users can be motivated by comparing their activity to the median participation of the community [Harper et al. 2007], as in the kind of competition that Collabio has designed into its leaderboards. Loafing can also be overcome via opportunities for reciprocity toward other friends [Resnick and Zeckhauser 2002], motivating our Facebook notifications upon tagging a friend.

Finally, we note that social person tagging is in effect a new venue for computer-mediated communication. Collabio can be used to send short affectionate notes (tag a friend as *cute*), or it can be used to flame and embarrass (tag a friend as *dropout*). Social interaction is known to be more uninhibited when computer-mediated [Kiesler et al. 1984], so we must be careful to design for positive interactions.

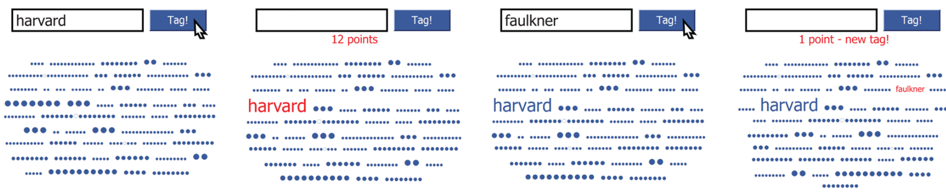

Fig. 2. The tag cloud begins completely obscured. The player guesses *harvard*, receives 12 points for agreeing with eleven other players and reveals Harvard as a large tag. *Faulkner* is next; it does not match existing tags and is added to the cloud.

3. COLLABIO

Collabio is a social tagging game embedded in the Facebook social network. We reported briefly on the Collabio system in previous work [Bernstein et al. 2009]. Here, we provide a fuller treatment of the application, and analyze the design space of such applications. To follow, we describe Collabio's three top level interface tabs: the tab in which users can *Tag!* their friends, the one in which they can manage *My Tags*, and the one in which they can see the *Leaderboard*. We then discuss propagation through the social network, the incentive design space, and issues of cheating and abuse.

3.1 Tag Friends

The main activity of Collabio is guessing tags that others have used to describe friends, so the focus of the user's experience is the tagging page (Figure 1). The user sees the tag cloud that others have created by tagging the selected friend. When presenting this cloud, Collabio only displays tags that the user has already explicitly guessed (Figure 2). Tags not yet guessed are obscured by replacing each constituent letter with a solid circle; for example, the tag *TOCHI* appears as ●●●●●. Whitespace in obscured tags is represented by clear circles such as ○. Thus, the length and makeup of the obscured tag provide hints as to the hidden text. As an additional hint, terms in the tag cloud are alphabetically ordered. The tags in the cloud are scaled so that the popular tags are larger.

As the user tags a friend, one of two things happens (Figure 3). If the tag is unique and has not previously been placed on their friend, the tag is inserted into the cloud. If the tag exists, then it is revealed within the cloud. For each guess, users receive points equal to the total number of people who have applied a tag, including themselves. If they are the only person to have guessed that tag, then they get 1 point; if there are 11 others, they get 12 points. These points continue to accumulate as more people apply the tag, so earlier taggers' scores rise as well. A user can retract a tag by clicking on a small × by the tag. To expose one's score to others, and to stimulate competition, each tagged friend has a "People who know [this friend] best" pane which lists friends who have earned the largest number of points from tagging that friend (Figure 1).

In the current system, if the user is the first to tag a friend, Collabio seeds the tag cloud with terms from the friend's public profile (such as network names, affiliations, or interests), thus ensuring that the tag cloud is never completely empty. These tags are attributed to the "Collabio Bot." We observed early on

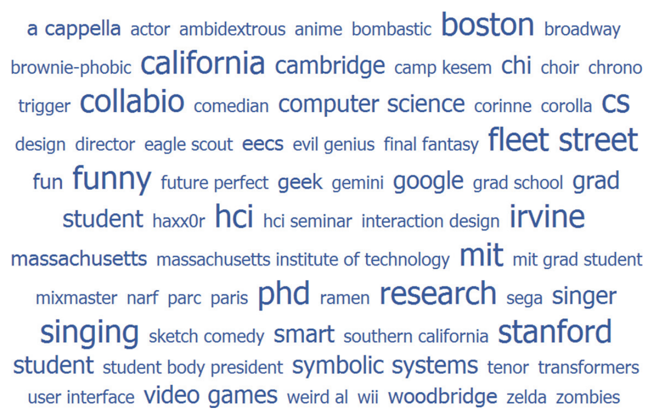

Fig. 3. The *My Tags* page allows the user to view their own tag cloud completely uncovered. Not shown: the top 10 scorers list and a complete listing of each tag's authors.

that users were typically unwilling to tag others who had not already added the application, so this tag seeding is helpful in overcoming reluctance to be the first to tag an individual.

3.2 Managing My Tags

The *My Tags* interface allows users to inspect and manage tags their friends have placed on them. The My Tags page contains three sections: a fully uncovered tag cloud (Figure 3), an expanded top scorers list, and a table explaining which friends used which tags. In order to allow people to maintain control of tags placed on them, Collabio allows them to easily delete tags from their tag cloud by clicking on a small × by the tag.

3.3 The Leaderboard

The third Collabio tab is the Leaderboard. While the individual leaderboards on the *Tag!* tab encourage users to keep tagging a friend until they are listed as one of the Top Friends for that person, the global leaderboards encourage users to continue tagging activity within the application. We present two lists here, one of the friends that have the most unique tags placed on them, and the other of the individuals in the user's social network who have tagged the most other friends (Figure 4).

3.4 Designing for Social Spread

Collabio relies on social mechanisms to spread to new users and retain existing ones. For example, the individual leaderboards are labeled "friends who know [this friend] best" to conflate closeness of friendship with score in the game, and notifications purposely do not share all the new tags to entice the user to visit the application.

As with typical Facebook applications, users can explicitly invite others to play. More subtly, when a user tags a friend, the application sends a Facebook

Most Tags from Friends	
You	1678
Meredith Ringel Morris	484
Desney Tan	372
Kurt Luther	260
Emilie Kim	245
Szymon Chachulski	235
Mary Czerwinski	229
Erin Treacy Solovey	219
Grace Woo	214
Greg Smith	194

Most Friends Tagged	
Sandra Emerson	213
You	204
Miguel Nacenta	61
Lilly Irani	55
Erin Treacy Solovey	55
Meredith Ringel Morris	55
Jeffrey Nichols	45
Biliana Kaneva	40
Desney Tan	38
Szymon Chachulski	36

Fig. 4. Collabio leaderboards feature the friends with the most tags (left) and the friends who have tagged the most others (right).

notification to the friend, whether or not that friend has previously played Collabio. The notification includes the user's name, the number of new tags, and a glimpse of the tags' contents:

> Michael Bernstein has tagged you with **cyclist** and 7 other tags using Collabio. Tag Michael back, or see what you've been tagged with. 2:41pm

A similar version appears on the tagger's wall feed and on Facebook's homepage news feed. Users can also place the occluded version of the tag cloud onto their Facebook profile page. The profile tag cloud demonstrates to visitors the number of tags the individual has acquired and serves as a hook for new users to install and play.

3.5 Incentive Design

One of Collabio's distinguishing characteristics is its incentive system for collecting tags. We designed Collabio's incentive system in a highly iterative manner, controlling and testing dimensions with Wizard of Oz prototypes played as a text-based game over Google Chat. Figure 5 summarizes the space of incentive and game options we considered. Here we discuss the design space and our eventual choices. We expect that many of these decisions will be relevant to other friendsourced systems as well.

Tag Visibility. Tags may be made immediately viewable by all users, as in Fringe [Farrell et al. 2007], or they may feature varying degrees of concealment. Fully public tags can be seen before guessing and enable a lighter-weight interaction (voting up existing tags) which is a low-risk hook in the same way that fixing spelling errors draws new users into Wikipedia [Bryant et al. 2005].

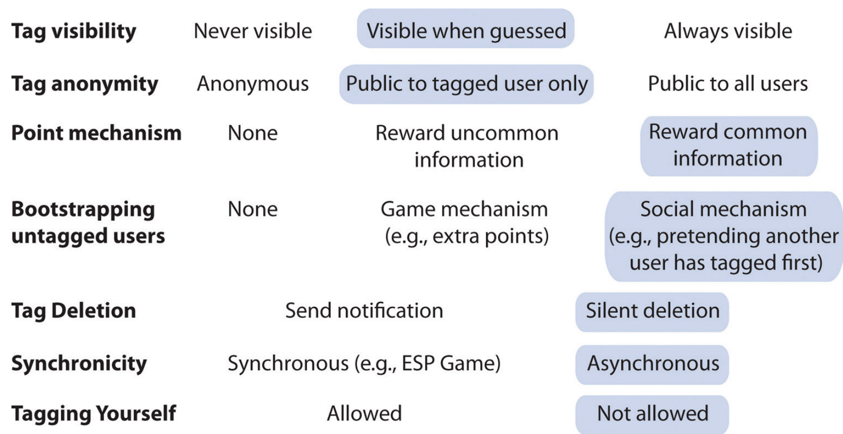

Fig. 5. The design space of social tagging applications. Collabio's choices are highlighted in blue.

However, publicly viewable tags do not provide for very challenging or interesting game mechanics and may prime users to provide tags similar to existing ones. Hiding the tags, by contrast, leads to independent verification and helps guarantee correctness [von Ahn 2006; Surowiecki 2005]. We based our design on games such as the ESP Game and Family Feud, which entailed hiding the tags. Other Facebook tagging games such as iDescribe, Compare People, and DescribeMe make them visible.

Tag Anonymity. Tags may be publicly attributed to the tagger, fully anonymous, or anonymous to players but attributed when the tagged user looks. We hypothesized that making all tag activity anonymous would encourage taggers to use mean and hurtful tags, but lead to more tagging since more terms are in bounds. Conversely, making tag authors public both to the tagger and the tagged user would reduce the number of negative tags, but might cause players to contribute fewer tags because mutual friends will judge your contribution. The anonymity decision impacts the nature of the application's notification system, as well. For example: "A friend has tagged you with cyclist and 7 other tags using Collabio." (How does the user know who to tag back?) Or: "Michael Bernstein has tagged you with 8 tags using Collabio." (Which tags did Michael use? Is this just Facebook spam?) We chose to make Collabio's tags public to the tagged user because we wanted to provide a clear call to action in the notifications, but private to all other taggers because we wanted to minimize perceived peer pressure. Other Facebook tagging games choose other points in this design space; for example, Compare People sends notifications to users that they have been judged more attractive than another friend, but don't share who did the judging, while Describe Me makes tagging authors public, even to new taggers.

Point Mechanism. We prototyped two versions of Collabio: one encouraging common information, and the other encouraging unique information. To encourage unique information, our prototype awarded more points to tags that were guessed by fewer people, provided that at least one other person had already supplied the tag. This version did not lead to very much tag agreement, so in practice players received few points. Without the agreement constituting independent verification, we could also not guarantee tag accuracy. In contrast, we found that the common information prototype (which grants more points when more people agree with a tag) collected quite a bit of unique information. The unique information came about because some guesses were inevitably misses and had not been guessed by other individuals. Guessing easy, obvious tags also provided a hook to bring new users into the game. This scheme maintains the properties of independent verification and information variety while making the game more fun by providing more points, so we chose it.

Bootstrapping Untagged Users. Social applications must recruit new users to grow. In Collabio this need is particularly acute: without other users, there are no tag clouds to uncover. As described, there is no game incentive to be the first to tag a friend. It can also be socially awkward to send Facebook notifications to a friend who has not demonstrated interest in the application. As a result, Collabio must incentivize players to bootstrap a friend's tag cloud. One approach is to reward players for being first with points—an early version of Collabio rewarded 5 extra points to the first player to tag each user. However, we found that the point incentive was not enough to "subsidize" the awkwardness of sending an unsolicited Facebook invitation. So, we turned to another approach—adjusting the social dynamics. Collabio now initializes all players' tag clouds with a set of words gathered automatically from their profiles, so that a player might see (system-fabricated) social proof that others have participated. Collabio attributes these machine-generated tags to the "Collabio Bot" so as not to be overly disingenuous. Once the first human player has finished playing against the auto-generated tag cloud, Collabio automatically removes the Collabio Bot tags and leaves only the human-generated tags. This approach has seen some success. However, as described in Section 7.2, bootstrapping remains a difficult problem. Applications such as Compare People deal with the bootstrapping issue by forcing randomization of who to tag so that players tag new users.

Tag Deletion. Should the tagger be notified when the tagged user deletes a tag from their tag cloud? If so, it may discourage a tagger from contributing, parallel to a Wikipedia author's first edit being reverted. Facebook culture suggests another option: silent deletion, similar to the silent decline of a friendship request. We hypothesized that most taggers would not notice missing tags the next time they visited. Collabio thus uses this silent delete mechanism. Facebook tagging applications such as Compare People, iDescribe, and Describe Me do not allow you to delete tags after they are applied; however, we found deletion to be an important part of Collabio's gameplay.

Synchronicity. We considered whether to make the game synchronous and paired like the ESP Game [von Ahn and Dabbish 2004] or asynchronous and unpaired like the Dogear Game [Dugan et al. 2007]. Our estimates suggested that it would be unlikely for two simultaneous Collabio players to have many friends in common, and that we would be relying mostly on replaying recorded partners as in the ESP Game. For this reason, we decided to make the game asynchronous.

Tagging Yourself. Many Collabio players have expressed interest in being able to tag themselves or seed their own tag cloud. Indeed, the most accurate picture of an individual comes by combining self-perception with others' perceptions [Vazire and Gosling 2004]. Though we did not implement this particular design, and neither do the other Facebook tagging applications, we regard this direction as interesting future work.

3.6 Dealing with Cheating and Abuse

Many games suffer from cheating, collusion, or other malicious actions. Because Collabio activity can only occur between people with a mutually established social connection, we rely on social pressures to prevent this behavior. Specifically, cheating in Collabio would involve annoying your friend by dirtying their tag cloud or sending many notifications, which are undesirable; the tagged individual can also manually retract points or unfriend the tagger.

There are several ways a user could conspire to increase their score. For example, they could ask the person who they are tagging or their friends for the answers. They could also reverse engineer tags using a search strategy on the alphabetized cloud. This behavior does not do active harm to the tag cloud, as it simply reinforces already existing tags. However, it does erode our premise that popular tags were generated by multiple independent sources. Fortunately, this is more work than just guessing at tags, and it is a poor method for drastically increasing one's score relative to everyone else's since mimicking friends' guesses simultaneously increases their scores as well. Another way to artificially increase one's score might be to tag a friend with a large number of nonsensical tags for 1 point each: for example, *a*, *aa*, *aaa*. However, this strategy quickly deteriorates because it does not take advantage of the work others are doing to earn you points and one point becomes worth less and less as more users tag.

Users could also decide to tag an individual with an undesirable tag as a joke or punishment. Since a tag is not automatically revealed to other users until they guess it, the payoff for such a strategy is rather low and non-public, and we did not see much of this in practice. Furthermore, the tagged individual is likely to disapprove of and delete inappropriate tags, thereby eliminating ill-gotten points or reward. We have seen people apply social pressures to friends to discourage such behavior. As regards misspelled or otherwise inaccurate tags, we rely on users' self-interest in maintaining a well-manicured public profile [DiMicco and Millen 2007].

3.7 Implementation

The Collabio application interface is built as an AJAX-enabled ASP.NET web application, which calls a Microsoft SQL Server-backed Windows Communication Foundation web service for data storage and querying. The application is served as a Facebook Markup Language (FBML) page using the Facebook API.

4. FIELD DEPLOYMENT AND EVALUATION

We analyzed tag statistics collected between July 2008 and March 2009 (about an 8 month period). In that time, Collabio gathered 7,780 unique tags on 3,831 individuals in 29,307 tagging events. These tags were generated by 825 different users out of 1,350 who installed the application according to Facebook. The median user who tagged at least one friend received 11 unique tags in return, indicating that even minimal usage of Collabio resulted in a user being relatively well-tagged by friends.

We supplemented this data with a survey methodology aimed at active users, who we defined as users who had tagged at least three friends, were tagged by at least three friends, and had at least nine distinct tags. Using Facebook's notification service, we invited Collabio's 112 most active users to fill out a survey about their experience. Forty-nine users (24 female) responded to the survey. The median age was 27 ($\sigma = 4.1$). The respondents were skewed toward students and researchers with an interest in user interfaces. We offered a small gratuity for responding.

4.1 Tie Strength

Users tended to tag friends close to them: their strong ties [Granovetter 1973; Gilbert and Karahalios 2009]. Survey results suggest that users would usually tag closer friends, but not exclusively so. This preference for stronger ties came about because it was easier to invent tags for them and because it could be awkward to send a Facebook notification to a friend who you had not spoken to in months. Our logs show that the average user tagged 5.8 other friends ($\sigma = 13.6$) with 6.1 tags each ($\sigma = 7.3$).

4.2 Reciprocity

Social reciprocity through Facebook notifications played a critical role in keeping users engaged. When asked about the reasons for tagging, 82% of survey respondents cited that the friend had tagged them first. In practice, 82% of Collabio users who joined after being tagged reciprocated by tagging at least one of the friends who had tagged them.

A reciprocating play cycle would begin when a user decided to tag a friend. That friend would then click on a link in the Facebook notification to see his or her new tags and then tag the friend back, and the original tagger would reciprocate again. Often these back-and-forths would occur between one user and several of their friends simultaneously. This process was enough to trigger many repeated visits to Collabio: the average pair of reciprocating taggers engaged in a combined 4.14 tagging sessions.

Very little activity was required to hook interested users to reciprocate. We expected that several friends would need to invite a user before they granted the Facebook application enough legitimacy to join [Granovetter 1978]. This was not the case. Of the users who began tagging after being tagged by others, 73% began tagging after having been tagged by one other person and 17% did so after receiving tags from two or more friends.

Reciprocity is effective once friends become engaged, but many Facebook users perceive invitations from any application (including Collabio) as spam. Across all Facebook applications, 63% of invitations are ignored;[7] in Collabio, 87% of tagged individuals never tagged others and 39% of installed users never tagged a friend. (The Facebook and Collabio numbers are not directly comparable because one measures invitation click streams and the other tagging activity). The 39% of users who installed the application but never tagged a friend were content to be lurkers or were only interested in viewing their own tags. It is an interesting dichotomy that the notifications are extremely effective at retaining active users, but poorer at recruiting new ones. It seems that reciprocity needs to be augmented by other incentive mechanisms to reach new users.

4.3 Tag Characteristics

In Collabio, single words are often enough to convey the essence of a concept in tag form. Collabio's mean tag length is 8.3 characters ($\sigma = 5.2$). 5,867 tags (~75%) are a single word, and 1,913 tags (~25%) contain multiple words.

Globally, the tags applied to the most individuals in Collabio are descriptors like *kind* and *smart* as well as affiliations such as *Stanford*. These generically positive descriptors point to the general good-natured bent of most Collabio tags, and suggest that we may have succeeded in preventing large-scale trolling.

To learn more about tag content, we asked each survey respondent to rate nine tags in their tag cloud. These tags were drawn from three buckets: *Popular Tags*, the three tags applied by the most friends; *Middling Tags*, tags drawn randomly from the set of tags that occurred at least twice but less often than the Popular Tags; and *Unique Tags*, tags drawn randomly from the ones applied by only a single friend. For users who did not have enough tags to fill the Middling Tags category, we instead presented a randomly generated string and removed the resulting data from later analysis.

For each tag presented, the user provided a rating on a 7-point Likert scale (1 for disagreement and 7 for agreement) for each of two questions: "This is a good tag for me," and "This tag is something I would expect lots of people to know about me." In addition, participants classified each tag into the following categories: school, workplace or group affiliation; professional or academic interest, expertise or title; recreational hobby, interest, or expertise; location; personality trait; physical description; name or nickname; another person in the participant's life; inside joke; don't know; or other.

We found that a large percentage of Collabio's tags are affiliations, interests, expertise and hobbies; the long tail of tags contributes a wide variety of unusual information. Table I reports that Popular Tags were reported to be mainly affiliations; Middling Tags and Uncommon Tags were more commonly reported to capture interests, expertise and hobbies. The Uncommon Tags were commonly categorized as Miscellaneous, including clothing choices, special abilities, and the name of a friend's dog.

[7]As reported by Facebook Developer application statistics, 5/30/09.

Table I. A Breakdown of Information Type by Tag Bucket. Affiliation and Interest Categories were the Most Popular Among the Three Categories

Tag Bucket	Definition	Most Popular Information
Popular Tags N = 147	Three most popular tags for the user	School, workplace or group affiliation (66.0%) Interests or expertise (16.3%)
Middling Tags N = 93	Less popular than Popular Tags, but occurring more than once	School, workplace or group affiliation (27.2%) Interests or expertise (23.9%) Hobbies (15.2%) Location (10.9%)
Uncommon Tags N = 147	Occurred only once	Interests or expertise (21.1%) Miscellaneous (15.6%) School, workplace or group affiliation (13.6%) Hobbies (12.9%)

Table II. User Ratings of How Accurate and Widely Known the Tag Buckets were, on 7-Point Likert Scale (1 = Very Inaccurate/Not Widely Known, 7 = Very Accurate/Widely Known)

	Popular Tags	Middling Tags	Uncommon Tags
Accurate	$\mu = 6.42$ $\sigma = 0.92$	$\mu = 5.83$ $\sigma = 1.39$	$\mu = 5.13$ $\sigma = 1.61$
Widely known	$\mu = 6.22$ $\sigma = 1.22$	$\mu = 5.21$ $\sigma = 1.58$	$\mu = 4.14$ $\sigma = 1.77$

4.4 Tag Accuracy and Popularity

Generally, the more popular the tag, the more accurate it was and the more well-known the fact. Survey participants rated all three classes of tags as accurate descriptors of themselves, and all but Uncommon Tags as known by many people (Table II, Figure 6). We ran one-way ANOVAs with tag bucket as independent variable and goodness of tag and expectations that others know the given facts as dependent variables. We found significant effects of tag bucket on goodness of tag ($F_{2,384} = 34.5$, $p < 0.001$, $\eta^2 = .15$) and expectation that others know the given facts ($F_{2,384} = 67.1$, $p < 0.001$, $\eta^2 = .26$). Pairwise posthoc comparisons using Bonferroni correction confirmed all factor levels were significantly different from each other in terms of accuracy and anticipated popularity.

We were surprised to find that even the Uncommon Tags were rated as fairly accurate descriptors, with a mean above neutral on the Likert scale. This result suggests that there is little inaccurate information in the Collabio tag database. Prior efforts, The Wisdom of Crowds [Surowiecki 2005] and Games with a Purpose [von Ahn 2006], claim that accurate data collection requires repeated independent verification of the same answer, and thus that that one-off answers should generally be discarded. However, we find that even the one-off answers in Collabio (the uncommon tags) are fairly accurate. It seems that Collabio's social incentives help to avoid serious misuse or off-topic tags.

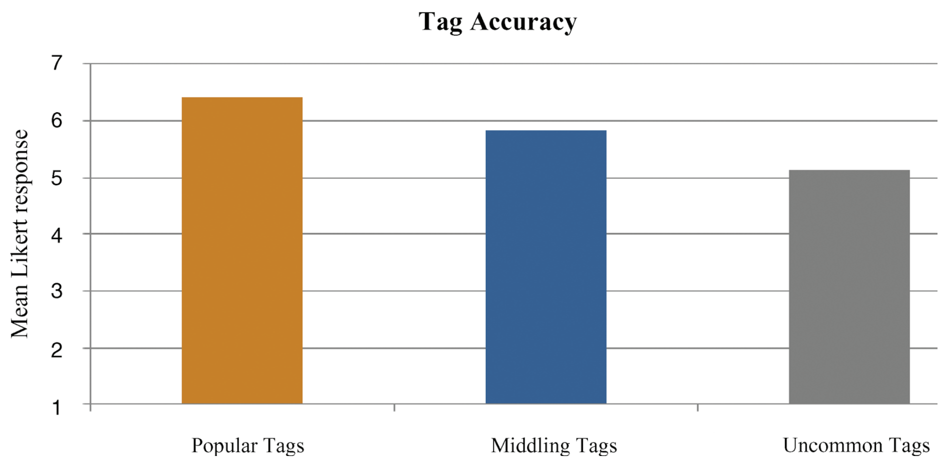

Fig. 6. A bar chart representation of Table II indicates that all three classes of tags were rated above neutral (4) as accurate descriptors.

4.5 Deleted Tags

Users commonly deleted tags. In fact, 18% of unique tag strings were deleted. For privacy reasons, we did not keep records of who produced and who deleted tags. However, few participants in our active user survey (11%) reported deleting tags from their own tag clouds. Thus, we believe that taggers perform most deletion, and that the following analysis applies primarily to the taggers themselves.

We saw deleted tags in negative spirit (e.g., *boring*, *arrogant*, *witch*, *blab*, *unreliable*, and *retarded*), humor (e.g., *married and unavailable Ha!*, *Croatian sensation*, and *Stanford Sucks*), sensitive information (e.g., *sexxxy*, *S&M,* and *smoke break*), and inaccuracies. Our social design thus causes taggers to bias against and retract negative tags, as we hypothesized it might. Players also commonly deleted typos (e.g., *adorabl*, *friendl;y*, and *girl scounts*).

Informal feedback suggests that lack of anonymity and concomitant social proof was also often responsible for deletion. Taggers did not like being the only person to apply a tag to an individual, especially if the tag was racy. Taggers reported trying a tag to see if others had guessed it, then retracting the guess if not. This behavior raises questions about the degree to which Collabio encourages groupthink rather than collects independent opinions. Modifying the design with regard to deletion and review by users may address and seek to optimize independent assessments versus collaboration on tags.

4.6 Tag Novelty: Rating Exercise

Our results suggest that Collabio generates accurate tags that are reasonably ordered by importance. However, if these tags are available elsewhere, we have not significantly advanced the state of the art. Could an algorithm or individual outside the social network just as easily create these tags by mining information available in users' Facebook profiles or the Web? Could these methods also reproduce the relative ordering of tags?

4.6.1 *Rating Study Method.* We randomly selected twenty survey respondents from the forty-nine who completed our previous survey. For each survey respondent we utilized the nine tags they had rated in the survey, as well as three *Fake Tags* that were false and thus should not appear anywhere associated with the individual. Fake Tags were chosen from the set of global Collabio tags: one from the top 5% most popular tags, one that occurred less than the 5% most popular tags but more than once, and one that occurred only once. Fake tags excluded any tags applied to the individual.

We recruited four native English speakers comfortable with Facebook and Web search, but who had never used Collabio and did not know any Collabio users, to serve as raters. We gave them a brief demonstration of Collabio. The raters' task was to find evidence for each tag on the user's Facebook profile and on the Web. For each target individual, raters were presented with the twelve tags in random order and asked to rate each on a 7-point Likert scale according to the following statement: "I can find strong evidence that the tag applies to this individual." Raters were trained to give a score of 7 if the tag appeared verbatim, a score of 1 if there was no evidence in support of the tag, and a score of 4 if moderate inference was required based on the available evidence (e.g., the tag was *Atlanta* but the only relevant evidence was that the person attended Georgia Tech); the other values on the ordinal scale captured in-betweens. Raters were trained on example tags and profile sets until satisfactory agreement on the scoring scale was achieved. We randomized the order that raters viewed individuals.

We tested whether our human raters, as a reasonable upper bound on machine inference, could find the tags on the Collabio users' profiles. Raters rated the set of tags under two scenarios: first using only the individual's Facebook profile available to friends, and second using only web search. In the web search scenario, raters were disallowed from concatenating the individual's name and the tag name into a search query (e.g., "john smith atlanta"), in order to better simulate a tag generation task with no prior knowledge of the tag. We believe this is a more difficult test for Collabio to pass than that undertaken by Farrell et al. [2007], who performed string equality tests to see whether tags existed on profiles, because human raters perform semantic inferences.

We also wanted to investigate whether our raters could determine how popular a tag had been, as verified by our survey data. For each individual, we asked raters to place each tag into its original bucket: Popular Tags, Middling Tags, Unpopular Tags, and Fake Tags. They were told that three tags came from each bucket.

4.6.2 *Rating Study Results.* Raters evaluated tag evidence on Facebook and the web for a total of 480 tags across the twenty individuals. Cronbach's alpha was calculated to measure agreement across the raters, producing an overall agreement score of .82.

Experts found more supporting evidence for the more popular tag buckets, on both Facebook and the Web (Table III, Figure 7). A two-factor ANOVA comparing the effect of tag bucket (Popular vs. Middling vs. Uncommon vs. Fake) and evidence type (Facebook vs. Web) on rating found a main effect of

Table III. Mean Ratings Applied to Tags, from 1 (No Evidence to Support Tag) to 7 (Tag Appeared Verbatim)

	Popular Tags	Middling Tags	Uncommon Tags	Fake Tags
Facebook Evidence	$\mu = 5.54$ $\sigma = 2.36$	$\mu = 4.20$ $\sigma = 2.68$	$\mu = 2.87$ $\sigma = 2.56$	$\mu = 1.56$ $\sigma = 1.76$
Web Evidence	$\mu = 5.72$ $\sigma = 2.29$	$\mu = 4.17$ $\sigma = 2.81$	$\mu = 3.04$ $\sigma = 2.65$	$\mu = 1.5$ $\sigma = 1.4$

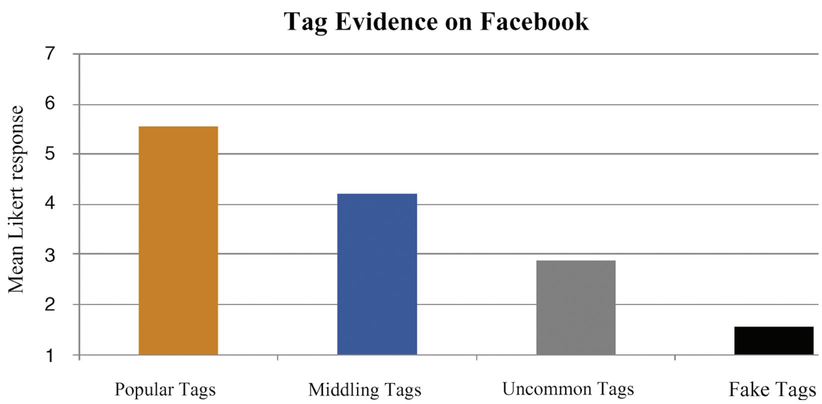

Fig. 7. A bar chart visualization of Table III, focusing on the Facebook condition. Popular Tags tended to have evidence available on the profile; Middling Tags and Uncommon Tags were much less likely to.

tag bucket ($F_{3,1915} = 270.0$, $p < 0.001$, $\eta^2 = .30$), and pairwise Bonferroni posthoc comparisons (all significant $p < 0.001$) suggested that the more popular a tag was, the higher rating it received and so the easier it was to find evidence for. Thus, the more popular the tag was, the more likely it occurred in a publicly visible area. We found no main effect of Evidence type, and inspection suggests that the scores between Facebook and the web are nearly identical.

In the bucket identification task, raters were the most reliable at identifying the extreme buckets: Popular Tags and Fake Tags (Table IV). Raters had the poorest performance on Middling Tags and Uncommon Tags, correctly recognizing only about 40% of each. Thus, beyond the most common tags, it is difficult for non-friends to reconstruct tag rankings.

Overall, raters found evidence supporting Popular Tags, but moderate inference was required for Middling Tags and very little evidence was available for Uncommon Tags. Our original survey respondents indicated that even Uncommon Tags were generally accurate, so we may conclude that Collabio is collecting accurate information with Middling and Uncommon Tags that would otherwise be difficult or impossible to acquire. Of the three categories, Popular Tags are fewest in number in the Collabio tag database, so most of the information Collabio collects is unique and thus complements existing public sources with typical online scraping techniques. Raters had considerable difficulty distinguishing Middling from Uncommon tags, and Uncommon from Fake Tags,

Table IV. Confusion Matrix of Rater Bucketing Decisions. Raters were Accurate at Identifying Popular Tags And Fake Tags, but Less so at Middling Tags and Uncommon Tags

		True Buckets			
		Popular	Middling	Uncommon	Fake
Rater Prediction	Popular	**151**	61	24	7
	Middling	63	**94**	50	30
	Uncommon	15	51	**103**	73
	Fake	11	34	63	**130**

so beyond the most obvious information it may also be difficult for a human, and certainly a machine, to recreate Collabio's tag ordering even coarsely.

5. PROOF-OF-CONCEPT APPLICATIONS

Friendsourcing is a useful tool for collecting explicit information about preferences and interests that may be difficult to obtain otherwise. Application developers seek out such information when they require user-specific data, for example in recommendation tasks, personalized search, and social network profiles. However, friendsourcing opens another avenue as well: applications which require *social* data; that is, applications which trade in information known to or relevant to only a small group. Yahoo! Answers cannot easily help with questions about the history of your small a cappella group or the way your friend delivered his marriage proposal; building applications on data such as the Collabio tags makes this possible.

We have created three illustrative prototypes utilizing the Collabio database: a tag cloud aggregator for tag visualization and exploration, an expert-finding question answering system, and a personalized RSS feed. We attempt two goals with this work: to demonstrate that friendsourced information can reproduce interactions built on asocial sources such as mining of user workstations, and that in some cases friendsourced data can provide new opportunities for interaction.

5.1 Collabio Clouds

Our system has learned thousands of tag clouds for users, so a straightforward first step is to consider tools for making sense of the tag space. Collabio Clouds allows users to compare themselves and other users of the system.

Collabio Clouds (Figure 8) aggregate tag clouds based on user queries. The user can query his or her own tag cloud as well as the aggregated tag cloud of friends, Collabio users, users tagged with specific Collabio tags (like *tennis* or *Adobe*), or users in Facebook networks or groups. Collabio Clouds allows users to explore questions such as: What do the tag clouds of members of the Penn State network look like? What other tags show up on individuals tagged with *machine learning*? What tags are most popular amongst all my friends?

Collabio Clouds uses a comparison tag cloud technique developed by ManyEyes [Viegas et al. 2007] to allow users to compare two groups. Thus, a user can compare his or her friends to all Collabio users, compare Michigan students to Michigan State students, compare people tagged with *football* to

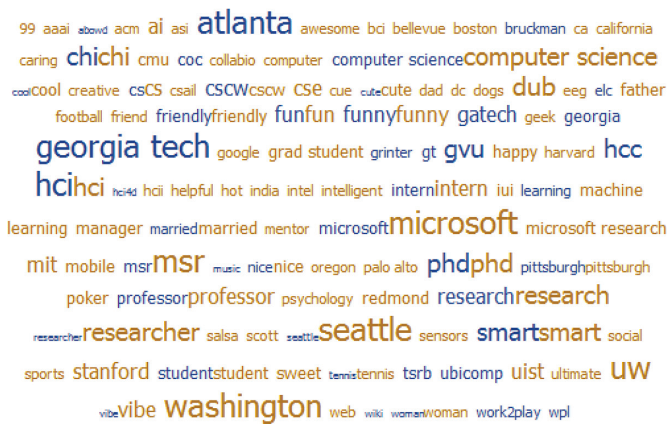

Fig. 8. A tag cloud comparing users tagged with *washington* to users tagged with *georgia* tech in Collabio Clouds.

people tagged with *baseball*, or compare Stanford members of the ACM SIGCHI group to Carnegie Mellon members of the ACM SIGCHI group.

Tag clouds are aggregated by number of members of the group who have a tag, so larger tags are more common in the population. To improve privacy, only tags that are applied to more than one individual are shown in the aggregate tag cloud.

5.2 Collabio QnA

Once we can aggregate tags, it makes sense to match people to each other in an expert-finding system. Question and answer (QnA) systems such as Yahoo! Answers[8] rely on a large community of answerers actively seeking out questions. Expert-finding algorithms can broaden the effectiveness of these tools by actively routing questions to users likely to know the answer. QnA systems with expert-finding components include Answer Garden [Ackerman and Malone 1990] and Aardvark[9]; Farrell et al. [2007] suggested that tags could be used for people-ranking.

As we did with Collabio, we embedded the Collabio QnA system (Figure 9) in Facebook. Users ask questions, and Collabio QnA searches over the collected Collabio tags to identify friends and friends-of-friends who are most likely to be able to answer the question. The user can then choose which friends to send the question to, and Collabio QnA provides a comment board for the answer thread.

Collabio QnA's expert-finding algorithm utilizes the Lucene search engine.[10] Each user's tag cloud is translated into a document in the search engine with terms weighted by number of friends who applied the tag. The user's question is then fed as a query to the search engine, and the ranked results are restricted to the user's friends and friends-of-friends. Lucene's default scoring

[8]http://answers.yahoo.com
[9]http://www.vark.com
[10]http://lucene.apache.org

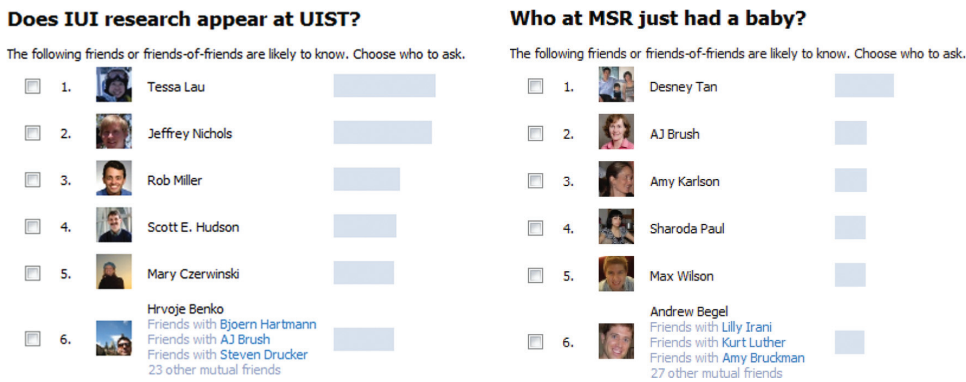

Fig. 9. Collabio QnA is a question and answer system that uses Collabio tags to find friends and friends-of-friends who can answer your questions.

function prefers short documents—in this context, users with fewer tags—so we utilize a length-independent scoring function to give all tag clouds equal scores regardless of size.

Collabio tags and the social network context provide the opportunity for our QnA system to route questions more highly relevant within the user's social network, such as *When is the next HCI group meeting?*, or *Who might be interested in starting an IM football team at Google?* These kinds of questions are difficult to answer using global QnA sites such as Yahoo! Answers.

5.3 Collabio RSS

Collabio QnA matches people to other people, but Collabio tags can also be used to match content to people. RSS (Really Simple Syndication) is a popular format allowing aggregation of web content, enabling users to subscribe to the feeds of web pages of interest. However, these feeds vary in relevance and can be overwhelming in number, making it difficult to identify the most relevant posts to read.

Collabio RSS is a personalized RSS feed of web content, powered by a user's Collabio tags. Collabio RSS builds on research in personalized web filtering (e.g., [Billsus et al. 2000; Brusilovsky et al. 2007]). It is unique from most content-based filtering algorithms in that its model is not implicitly learned from user behavior; the tag knowledge base enables a simple information retrieval approach to filtering and enhances scrutability of its results [Vig et al. 2009].

To produce the news feed, Collabio RSS indexes the title and text content of each feed item as a document in Lucene. When a user requests a personalized feed, it retrieves that user's Collabio tag cloud and performs a document-as-query search on the feed corpus: the weighted tag cloud is concatenated as an OR'ed search query and weighted by tag popularity. Tag weights are log-transformed to prevent the most popular tags from overwhelming the results. We filter the corpus using a sliding time window of the past day and a half to keep the feed's content fresh.

We crawled 2610 popular RSS feeds recommended as bundles by Google Reader,[11] indexing 68,069 items posted over 36 hours. As an example, ten randomly selected posts vary greatly in topic:

1. 2010 Pontiac Solstice GXP Coupe Test Drive: 28 MPG and Turbo Power, but Practicality Not So Much
2. The X-Files: Season 7: Disc 4
3. 26 Carriers Commit To Deploying LTE; Some Backers Look For Way To Make Voice Calls
4. 5 Reasons Why the PTR Sucks
5. 30 More Free Blog Icons, Website Icons, Symbol Icons
6. 2009 is the year of the comic book in Brussels
7. Superman Cartoons
8. 84th Precinct Crime Blotter
9. 1D-navigation using a scalar Kalman filter
10. 13 Tasteless Costumes Ever

However, Collabio RSS feed identifies items of much greater interest to one of the authors, containing items relevant to HCI, graduate school, and nerd culture in Boston.

1. Weekly Mashable Social Media & Web Event Guide
2. Job Offer: PhD Position in Information Visualization, Växjö University, Sweden, and TU Kaiserslautern, Germany
3. 6 Y Combinator Startups I Would Have Invested In Back Then
4. Mind Meld: Non-Genre Books for Genre Readers *[sci-fi books]*
5. Job: Postdoc in Visual Perception, Delft University
6. The Information School Phenomenon
7. Speaking of (and in) 2009 [*speaking schedule of HCI figure*]
8. Tonight: Video Game Orchestra at Berklee
9. Brain-Computer Interfaces: An international assessment of research and development trends
10. Exploring Siftables: the blocks that play back [*HCI research at author's university*]

The Collabio RSS feed has high relevance because Collabio collects so many tags related to professional and recreational interests. Affiliation-oriented tags, also popular, are responsible for returning news relevant to places and organizations the author has been associated with in the past.

5.4 Next Steps with Applications

These applications serve as proofs of concept, and their individual utility would be best demonstrated by field studies. However, our goal in this work is not to

[11]http://reader.google.com

evaluate the applications per se, but to demonstrate that friendsourced information can successfully power personalized applications in a context separate from their original capture. Thus we have opted to produce a variety of prototypes to explore the broad space of such applications. In depth analysis of the applications remains future work.

6. DISCUSSION AND LESSONS LEARNED

In this section we report open challenges with friendsourced designs and reflect on data mining the information that friendsourced applications such as Collabio will produce.

6.1 Friendsourcing Collects Accurate and Novel Information

Tying together the survey and the rating exercise we see that Popular Tags, which largely captured group affiliations, could in principle be generated by mining available information such as Facebook or the web, even though we know of no current system that can do this reliably. Middling Tags and Uncommon Tags, which users view as good descriptors of themselves, are difficult for others outside the social network to verify and by extension to generate. Thus, Collabio generates tags that are not available to typical Web mining methods and these tags cannot reliably be judged accurate by individuals outside the social network.

Even unverified, unpopular information is typically accurate in Collabio. This result suggests that guaranteeing accuracy may not be a major design concern for friendsourced systems. This benefit may have been carried over from crowdsourcing: only 1–2% of Wikipedia edits are dedicated to reverting vandalism [Kittur et al. 2007].

Friendsourced applications may be most useful, then, in producing a wide variety of non-generic information about its users. While the system may reward its users for producing popular information, the large majority of tags in our database are not popular. This large majority is the class of information that exhibits the most potential: it is both accurate and unavailable elsewhere. The Dogear Game makes clever use of this situation as well by focusing on incorrect answers as sources of information about the misattributed individuals [Dugan et al. 2007].

6.2 Friendsourcing Succeeds in Generating Information about Active Users; Convincing New Users to Join Is Harder

We have learned that Collabio's social design is successful at tagging users who join the application. Users who tagged at least one other friend were well-tagged by their friends in return: the median such user received 11 unique tags.

Social spread is a more difficult problem, and our approach could be improved with knowledge gained through the design process. A large majority of tagged users never accepted the application invite, suggesting that the social incentives appropriate for notifying active users differ from those appropriate for invited users, that the Facebook population is largely wary of unknown applications, or both. The problem is exacerbated by active Collabio users who were hesitant to send invitations to their friends. One user reported: *"I'm reluctant to start tagging people that haven't added it. If they were already on Collabio, I'd probably tag [them]. Since they're not though, I feel like it'd be annoying if I started tagging them."*

We pursued two designs targeting this challenge. The first design offered extra points to the user who was the first to tag each friend. However, this offer was not sufficient incentive for many users. While point scores are successful at motivating continued tagging to become a person's top friend, they do not motivate the first few tags when there is nobody to compete with. We replaced this design with one that automatically seeds the tag cloud with words from the profile, encouraging the first tagger to compete against the Collabio Bot. However, this design was likewise not enough to overcome spam concerns. This situation remains an open design challenge for Collabio and similar social tagging systems on Facebook. It is possible that publicly releasing our applications utilizing the tags would more strongly motivate new users to join.

6.3 Tag Exhaustion

Tag exhaustion occurs when a user "runs out of tags" and cannot invent any more tags for their tagging partners. When tag exhaustion happens, no notification is sent and neither player is prompted to return. This is the primary reason users stop using Collabio. The design challenge is that the information Collabio collects is static—new tags do not often invent themselves. We believe that information exhaustion will occur with any similar system collecting static information about friends. One way to overcome Collabio's tag exhaustion is to solicit different kinds of tags each day (e.g., about history, interests, hobbies, and so on) to encourage return visits; another would be to ask players to verify tags rather than create new ones. Dynamic information is another option: status updates ("What is Desney doing?") are always in need of updating, and thus the most dedicated users could continue to contribute.

Dynamic information needs seem especially promising to keep friendsourcing systems fresh. To explore this space, we have built a friendsourced news sharing system called FeedMe, embedded in the Google Reader RSS reader [Bernstein et al. 2010]. FeedMe extends the friendsourcing design space by actively encouraging users to do work beneficial to their friends, not just to collect information. FeedMe suggests that users consider each new blog post they read as information that a friend might be interested in seeing. The system simplifies sharing those posts with friends, thus gifting those friends with posts and information they might not have otherwise seen. Simultaneously, FeedMe learns what posts particular friends might be interested in seeing so that it can recommend the friends as recipients for future posts. Users have adopted FeedMe for near-daily use for months on end thus far, so this approach shows promise.

6.4 Controlled Study of Game Parameters

Earlier we presented an analysis of the social tagging game design space, but we did not report controlled study—rather, a highly iterative design process. The design space is of such high dimensionality that controlled study would both be challenging and leave out important higher-order interactions. Simulation and modeling can be used to test multiple dimensions simultaneously once theory has developed further [Ren and Kraut 2009]. Until then, we are interested in pursuing a controlled study of a subset of these dimensions, particularly the effects of tag anonymity and of directly encouraging uncommon information via novel point schemes.

6.5 Building Applications Using the Friendsourced Information

Collabio Clouds, Collabio QnA, and Collabio RSS have given us insight into techniques and challenges associated with mining the Collabio tags. Information retrieval techniques such as tf-idf are important means for normalizing out common tags such as *kind*, *beautiful,* and *nice*. Tag sparsity issues may have been expected, but we found that Collabio users typically tried several different versions of a single idea when tagging (e.g., *computer science*, *CS, comp sci*), so in practice this was not a major issue. In addition, the stemming that search engines apply to the tags often hashes together different conjugations of a tag. If sparsity becomes an issue for applications, collaborative filtering (people tagged with one tag were often tagged with another) could implicitly add likely tags.

We cannot distinguish the semantics of any given tag, so we do not know if a tag is appropriate for a given personalization purpose. In the future we intend to try targeting tagging activities more carefully in order to generate tags relevant to a particular application. For example, once a week we might encourage only tags related to college, or to favorite movies. We believe human users are best situated to make these hard semantic decisions, and we would like to leverage this fact. In addition, new tagging tasks might help keep the application fresh.

We believe that Collabio tags will complement existing data mining approaches to personalization. Collabio largely sidesteps the privacy and deployment issues that burden the data mining of private sources such as e-mail or web history. Furthermore, the generated information is guaranteed to be semantically meaningful to the user, whereas automated techniques often result in information that textually distinguishes a user but does not carry much meaning.

7. CONCLUSION

We have investigated the design space of friendsourced social applications: designs that collect information or execute tasks in a social context by mobilizing a user's friends and colleagues. Friendsouring enables support for previously difficult tasks such as personalization, upkeep of public information about inactive users, and recommendation. To explore this design space, we developed Collabio, a social network application that extracts information about peoples' interests and preferences by encouraging friends to tag each other with descriptive terms in a game. Collabio has been successful in motivating players to tag almost 4,000 people with tags that are both accurate and contain information not available elsewhere. The resulting data can power visualization and personalization applications, especially those requiring social knowledge.

ACKNOWLEDGMENTS

We thank the Microsoft Research VIBE and ASI groups, as well as the 2008 Microsoft Research interns, for helping us improve and spread Collabio. Many thanks also to Rob Miller and David Karger at MIT for their feedback and support.

REFERENCES

ACKERMAN, M. S. AND MALONE, T. W. 1990. Answer Garden: a tool for growing organizational memory. In *Proceedings of the International ACM SIGGROUP Conference on Supporting Group Work (GROUP'90)*. ACM Press, 31–39.

BEENEN, G., LING, K., WANG, X., ET AL. 2004. Using social psychology to motivate contributions to online communities. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW'04)*. ACM Press, 212–221.

BERNSTEIN, M., MARCUS, A., KARGER, D. R., AND MILLER, R. C. 2010. Enhancing directed content sharing on the Web. In *Proceedings of the Conferece on Human Factors in Computing Systems (CHI'10)*. ACM Press.

BERNSTEIN, M., TAN, D., SMITH, G., ET AL. 2009. Collabio: A game for annotating people within social networks. In *Proceedings of the Annual ACM Symposium on User Interface Software and Technology (UIST'09)*. ACM Press, 177–180.

BILLSUS, D., PAZZANI, M. J., AND CHEN, J. 2000. A learning agent for wireless news access. In *Proceedings of the International Conference on Intelligent User Interfaces (IUI'00)*. ACM Press, 33–36.

BRUSILOVSKY, P., KOBSA, A., AND NEJDL, W. 2007. *The Adaptive Web*. Springer-Verlag.

BRYANT, S. L., FORTE, A., AND BRUCKMAN, A. 2005. Becoming Wikipedian: transformation of participation in a collaborative online encyclopedia. In *Proceedings of the International ACM SIGGROUP Conference on Supporting Group Work (GROUP'05)*. ACM Press, 1–10.

CABRERA, A. AND CABRERA, E. F. 2002. Knowledge-Sharing Dilemmas. *Organiz. Stud. 23*, 5, 687–710.

CHA, M., KWAK, H., RODRIGUEZ, P., ET AL. 2007. I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system. In *Proceedings of the Internet Measurement Conference*. ACM Press, 1–14.

CRESS, U., KIMMERLE, J., AND HESSE, F. W. 2006. Information exchange with shared databases as a social dilemma: The effect of metaknowledge, bonus systems, and costs. *Commun. Resear. 33*, 5, 370–390.

DIMICCO, J. M. AND MILLEN, D. R. 2007. Identity management: multiple presentations of self in facebook. In *Proceedings of the International ACM SIGGROUP Conference on Supporting Group Work (GROUP'07)*. ACM Press, 383–386.

DUGAN, C., MULLER, M., MILLEN, D. R., ET AL. 2007. The dogear game: a social bookmark recommender system. In *Proceedings of the International ACM SIGGROUP Conference on Supporting Group Work (GROUP'07)*. ACM Press, 387–390.

FARRELL, S., LAU, T., NUSSER, S., ET AL. 2007. Socially augmenting employee profiles with people-tagging. In *Proceedings of the Annual ACM Symposium on User Interface Software and Technology (UIST'07)*. ACM Press, 91–100.

GILBERT, E. AND KARAHALIOS, K. 2009. Predicting tie strength with social media. In *Proceedings of the Conferece on Human Factors in Computing Systems (CHI'09)*. 211–220.

GOLDER, S. AND HUBERMAN, B. 2005. *The Structure of Collaborative Tagging Systems*. http://arxiv.org/pdf/cs.DL/0508082.

GOSLING, S., GADDIS, S., AND VAZIRE, S. 2007. Personality impressions based on Facebook profiles. In *Proceedings of the International Conference on Weblogs and Social Media (ICWSM'07)*, AAAI Press.

GOULDNER, A. W. 1960. The norm of reciprocity: A preliminary statement. *Am. Soc. Rev. 25*, 2, 161–178.

GRANOVETTER, M. 1973. The strength of weak ties. *Am. J. Soc. 78*, 6, 1360–1380.

GRANOVETTER, M. 1978. Threshold models of collective behavior. *Am. J. Soc. 83*, 6, 1420–1443.

GUTWIN, C. AND GREENBERG, S. 2002. A descriptive framework of workspace awareness for real-time groupware. *Comput. Supp. Cooper. Work 11*, 3, 411–446.

HARPER, F. M., XIN LI, S., CHEN, Y., AND KONSTAN, J. A. 2007. Social comparisons to motivate contributions to an online community. In *Persuasive Technology*. Springer, 148–159.

HOFFMANN, R., AMERSHI, S., PATEL, K., ET AL. 2009. Amplifying community content creation with mixed initiative information extraction. In *Proceedings of the Conferece on Human Factors in Computing Systems (CHI'09)*. ACM Press, 1849–1858.

HONG, L. AND CHI, E. H. 2009. Annotate once, appear anywhere: collective foraging for snippets of interest using paragraph fingerprinting. In *Proceedings of the Conferece on Human Factors in Computing Systems (CHI'09)*. ACM Press, 1791–1794.

KAMMERER, Y., NAIRN, R., PIROLLI, P., AND CHI, E. H. 2009. Signpost from the masses: learning effects in an exploratory social tag search browser. In *Proceedings of the Conferece on Human Factors in Computing Systems (CHI'09)*. ACM Press, 625–634.

KARAU, S. AND WILLIAMS, K. 1993. Social loafing: A meta-analytic review and theoretical integration. *J. Personal. Soc. Psych. 65*, 4, 681–706.

KERR, N. L. 1999. Anonymity and social control in social dilemmas. In *M. Foddy, M. Smithson, S. Scheider and M. Hogg, Resolving Social Dilemmas: Dynamic, Structural, and Intergroup Aspects*. Psychology Press, Philadelphia, PA, 103–120.

KIESLER, S., SIEGEL, J., AND MCGUIRE, T. W. 1984. Social psychological aspects of computer-mediated communication. *Ame. Psych. 39*, 10, 1123–1134.

KITTUR, A., SUH, B., PENDLETON, B. A., AND CHI, E. H. 2007. He says, she says: conflict and coordination in Wikipedia. In *Proceedings of the Conferece on Human Factors in Computing Systems (CHI'07)*. ACM Press, 453–462.

LAMPE, C. A., ELLISON, N., AND STEINFIELD, C. 2007. A familiar face(book): profile elements as signals in an online social network. In *Proceedings of the Conferece on Human Factors in Computing Systems (CHI'07)*. ACM Press, 435–444.

LATANE, B. AND DARLEY, J. M. 1970. *The Unresponsive Bystander: Why Doesn't He Help?* Appleton-Century Crofts, New York, NY.

LATANE, B., WILLIAMS, K., AND HARKINS, S. 1979. Many hands make light the work: The causes and consequences of social loafing. *J. Person. Soc. Psych. 37*, 6, 822–832.

MARKEY, P. 2000. Bystander intervention in computer-mediated communication. *Comput. Human Behav. 16*, 2, 183–188.

MARLOW, C., NAAMAN, M., BOYD, D., AND DAVIS, M. 2006. HT06, tagging paper, taxonomy, Flickr, academic article, to read. In *Proceedings of the ACM Conference on Hypertext and Hypermedia (HYPERTEXT'06)*. ACM Press, 31–40.

MILLEN, D. R., FEINBERG, J., AND KERR, B. 2006. Dogear: Social bookmarking in the enterprise. In *Proceedings of the Conferece on Human Factors in Computing Systems (CHI'06)*. ACM Press, 111–120.

MULLER, M. J. 2007. Comparing tagging vocabularies among four enterprise tag-based services. In *Proceedings of the International ACM SIGGROUP Conference on Supporting Group Work (GROUP'07)*. ACM Press, 341–350.

RASHID, A. M., LING, K., TASSONE, R. D., ET AL. 2006. Motivating participation by displaying the value of contribution. In *Proceedings of the Conferece on Human Factors in Computing Systems (CHI'06)*. ACM Press, 955–958.

REN, Y. AND KRAUT, R. E. 2009. *A simulation for designing online community: Member motivation, contribution, and discussion moderation. Orgniz. Stud. 28*, 3, 377–408. In press: Online at http://www.cs.cmu.edu/~kraut/RKraut.site files/articles/reno7-simulating OnlineCommunit-v4.6-rek.pdf.

RESNICK, P. AND ZECKHAUSER, R. 2002. Trust among strangers in Internet transactions: Empirical analysis of eBay's reputation system. *Adv. Appl. Microec. 11*, 127–157.

STECHER, K. AND COUNTS, S. 2008. Thin slices of online profile attributes. In *Proceedings of the International Conference on Weblogsand Social Media (ICWSM'08)*. AAAI Press.

SUROWIECKI, J. 2005. *The Wisdom of Crowds*. Random House, Inc.

TOOMIM, M., ZHANG, X., FOGARTY, J., AND LANDAY, J. A. 2008. Access control by testing for shared knowledge. In *Proceedings of the Conferece on Human Factors in Computing Systems (CHI'08)*. 193–196.

VAZIRE, S. AND GOSLING, S. 2004. e-Perceptions: Personality impressions based on personal websites. *J. Person. Soc. Psych. 87*, 1, 123–132.

VIEGAS, F. B., WATTENBERG, M., VAN HAM, F., ET AL. 2007. ManyEyes: a site for visualization at internet scale. *IEEE Trans. Visual. Comput. Graph. 13*, 6, 1121–1128.

VIG, J., SEN, S., AND RIEDL, J. 2009. Tagsplanations: explaining recommendations using tags. In *Proceedings of the International Conference on Intelligent User Interfaces (IUI'09)*. ACM Press, 47–56.

VON AHN, L. AND DABBISH, L. 2004. Labeling images with a computer game. In *Proceedings of the Conferece on Human Factors in Computing Systems (CHI'04)*. ACM Press, 319–326.

VON AHN, L. 2006. Games with a purpose. *Computer 39*, 6, 92–94.

Received July 2009; revised January 2010; accepted February 2010 by Starr Roxanne Hiltz