

# IBM Research Report

## Augmenting Employee Profiles with People-Tagging

**Stephen P. Farrell, Tessa A. Lau, Eric M. Wilcox**

IBM Research Division  
Almaden Research Center  
650 Harry Road  
San Jose, CA 95120-6099

**Michael Muller, Kate Ehrlich**

IBM Research Division  
One Rogers Street  
Cambridge, MA 02142



Research Division

Almaden - Austin - Beijing - Haifa - India - T. J. Watson - Tokyo - Zurich

# Augmenting employee profiles with people-tagging

## ABSTRACT

Employee directories play a valuable role in helping people find others to collaborate with, solve a problem, or provide needed expertise. Serving this role successfully requires accurate and up-to-date user profiles, yet few users take the time to maintain them. In this paper, we present a system that enables users to *tag* other users with key words that are displayed on their profiles. We discuss how people-tagging is a form of social bookmarking that enables the community to maintain part of the employee directory.

Analyses of the usage data from our system show that people have used tags for impression management and that they contribute distinctive information to the employee profile beyond the information available in the basic profile. We also report on initial feedback from users, which has been mostly positive. We conclude with a discussion of the social issues we have encountered based on our experience deploying a people-tagging application in a corporate environment.

## Author Keywords

social software, tagging, impression management, online directory, bookmark, computer-mediated communication

## ACM Classification Keywords

H.5.3 Group and Organizational Interfaces (CSCW)

## INTRODUCTION

Enterprise directories are an important tool for finding information about other people in the workplace to build awareness and to learn about them prior to contacting them. Our corporate directory, known as “BluePages”, receives over 350,000 hits per day and is widely regarded as the most successful Intranet application.

Users come to the directory not only for contact information but also for information about the person’s organizational and social context. In support of this behavior, BluePages has an elaborate profile template that allows users to specify their projects, expertise, experience, teams, education, and

other attributes. Despite the overall success of BluePages, the level of individual contribution to profiles is frequency cited as a problem. Current methods for populating and maintaining directory content have many limitations. Critical parts of the profile, built automatically from human resources databases (e.g., the organization chart), can be reliable and up-to-date but limited. Other parts, such as a biography, experience, teams, projects, and interests rely on each person filling out that information. Yet in the BluePages directory, only 37% of profiles have been updated in the past nine months and 17% have been updated in the past three months. These numbers reflect *any* changes to the profile; the number of employees who have provided comprehensive and up-to-date information is far lower.

Our prototype directory, BluePages+1, investigates a variety of approaches to enrich and extend the information available in the directory. One approach is to automatically integrate content created by users on other corporate systems, such as blogs or social bookmarks. While successful, this approach is still limited since relatively few people use those other tools. In this paper, we explore the question: can we empower the *community* to maintain the corporate directory as a whole?

“Web 2.0” technologies like blogs, social bookmarking, and wikis have begun to break down traditional barriers and give a community of users the power to create content on the web. Since the activity of users often fits a power law curve, this approach often yields compelling results where an active and motivated minority contributes information to a much larger passive user base. For such systems to work, the active users need to feel rewarded for their contributions, and the passive users need to trust the process and outcome. Ensuring that both these needs are met are the challenge of creating such systems. In the case of the directory there is an additional hurdle—the profile is *personal*, and so reputations and feelings are at stake.

We have taken a Web2.0 approach to maintaining the corporate directory, inspired by social bookmarking systems like del.icio.us<sup>1</sup> and Dogear [9]. Such systems enable users to bookmark a resource and associate additional metadata with it such as comments and *tags*. Tags can not only help the tagger recall information in the future, but can also help other people find interesting content.

---

<sup>1</sup><http://del.icio.us/>

Specifically, we have extended BluePages+1 to let *people tag people*. Users can tag each other with key words that are then automatically displayed on the taggee's profile. Our goal is to bring the benefits of tagging to the domain of contact management, and in the process distribute the work of profile management across the entire community.

## PRIOR WORK

Social bookmarking systems have emerged recently as popular tools for organizing and sharing information. The del.icio.us social bookmark manager was one of the first of this kind of application. Since then, other systems have incorporated tagging as a means for annotating information, such as photos<sup>2</sup> and blogs<sup>3</sup>.

The concept of tagging *people* as opposed to web resources has received relatively little attention. Tagalag<sup>4</sup> is one such system; it enables users to tag others based on their email address. Tagalag provides integration with web-based mail systems via a Greasemonkey script that adds tags to the web-mail interface. However, it seems to have had relatively little adoption. 43people<sup>5</sup> is another people-tagging service that focuses on tagging celebrities. To our knowledge, we are building the first general-purpose people-tagging application in widespread use.

One of the potential benefits of tagging is locating experts in an enterprise. Prior expertise-finding systems have examined the tags one applies to social bookmarks as a measure of one's expertise, which is a more indirect measure compared to tagging people directly. For example, John and Seligmann describe the ExpertRank algorithm for finding expertise based on bookmarks one has tagged [8] (see also [15]). The tag rank analysis we present below is based on tags applied directly to people, rather than tags applied to web pages.

One use of people tagging is to help maintain relationships with one's contacts. For example, the ContactMap system [11] was based on a study that found that people have difficulty "remembering the identities of people in their social networks, particularly those who are important but contacted infrequently." ContactMap mined email communications to determine one's social network, and provided a graphical tool to manage that network. We have similar goals of helping people manage their contacts, though the approach we have taken (distributed social tagging) is quite different.

Managing relationships is notoriously difficult in the "lean media" of computer-mediated communication in general (see [3] for review), and the especially lean medium of directories and tags. While CMC is increasingly being integrated into people's relationship management strategies [13], it is not yet clear how people construe tags in general, and of course we do not know how people construe tags *on persons*. Boyd notes that "the digital world requires people to

write themselves into being," [4], and tags (applied to objects or to persons) are likely part of the digital traces that people create as they perform this self-making. But how do people use tags for this purpose in an enterprise context? Turner et al. noted that organizations establish norms about media use, in which media serve as both communication channels and as signals [14]. New media may not have norms—their norms and their meanings have not yet been socially constructed, either through usage (bottom-up) or through policy (top-down). When we introduced people-tagging, we were introducing a new medium, and we are currently engaged in studying the bottom-up social construction of the meaning of that new medium for our employees.

## PEOPLE TAGGING SERVICE

We define a tag as a tuple of (tagger, tag, taggee). Each tagger may ascribe multiple different tags to the same taggee, but not more than one instance of each tag for the same taggee. Multiple people may tag the same taggee with the same tag.

We make the distinction between *incoming* and *outgoing* tags. Incoming tags are those ascribed to a person, while outgoing tags are those tags that the person has ascribed to others.

In order to learn more about people-tagging, we developed and deployed a web-based application. We begin with a discussion of the design principles we followed, and then describe various features of the web application we produced.

### Design principles

#### *Identity and transparency*

The first significant design decision was to allow users to tag one another without the express permission of the person being tagged. Many popular social networking services such as MySpace<sup>6</sup> and LinkedIn<sup>7</sup> require that users send and accept invitations to establish connections. Inasmuch as socially tagging others represents establishing a link, we decided to keep the system as open and unconstrained as possible. We based this decision on the fact that within the corporate firewall, every user's identity is concretely known and traceable. In the interface we expose who has entered any given tag. Our inclination has been to assume the best in people, and trust that people are less likely to contribute malicious or inappropriate tags when their professional identity is clearly tied to the words and people they tag.

#### *Low barrier to entry*

Our second design decision was to minimize the steps required for people to tag one another. One of the pain points in the existing corporate profiling system has been the amount of time and complexity needed to update personal content. With tags, we aimed to find a lighter-weight way to let people contribute content. Following from our first principle of openness and transparency, we set out to make tagging as simple as entering a handful of keywords into a text field. Users can immediately see the impact of their contribution.

<sup>2</sup><http://flickr.com>

<sup>3</sup><http://technorati.com>

<sup>4</sup><http://tagalag.com>

<sup>5</sup><http://43people.com>

<sup>6</sup><http://myspace.com>

<sup>7</sup><http://linkedin.com>

## Tag dialog

Each profile page follows a standard layout as shown in Figure 1. At the top of the left-most column is a small panel that provides an entry field for tagging a person (A). It suggests tags as the user types based on overall frequency. As each tag is entered by a user, the display dynamically updates. All the tags ascribed to the profile by the current user are shown in green above the entry field. The user can choose to discard tags by clicking and dragging them with the mouse to a discard area near the bottom of the panel. All tags are centrally stored and persist. Subsequent visits to the profile page will load any tags attributed to the profile by the current user and display them in this panel.

## Tag clouds

Beneath the tag entry panel are two tag clouds. The first, incoming tag cloud represents a collection of all the tags attributed to the current profile across all users. As is the convention, tags are sorted alphabetically and typographic weight and size are used to signify frequency for a given tag with more frequently used values represented in larger and bolder type. Hovering the mouse over any given tag provides a count and the list of users that ascribed the tag (B). The second, outgoing tag cloud follows the same interface conventions as the first. However, the contents here represent all tags used by the profiled person. In each case, a count of the number of people (number of people tagging in the first cloud, number of people tagged in the second) is called out at the top of the panel as a hyperlink. Clicking on either of these links navigates the user to a page that lists all the individuals attributed to the given cloud.

## Contact icons

A major theme in the interface aims to provide context to the current profile by displaying related people. As a base case, every profile contains one or more panels detailing the persons management chain. In the event that the user has tagged any of the individual listed, we augment the individuals name with a contact icon (C). Hovering over the icon displays the ascribed tag set. The icons themselves provide a visual mechanism to trigger relevance and provide a means to explore potential relationship channels.

## Pivoting and searching

Users can directly pivot-browse by clicking on any tag. A pivot results in showing all the people across the company tagged with the selected value as shown in Figure 2. The same results can be reached by searching via keyword in the top-right corner search field.

In addition to listing people with matching tags, a cloud of related tags is calculated and displayed in the left column. These tags are aggregated across all the people displayed in the list. Users can continue to pivot on these tags and further explore the tag network of individuals.

The single search box at the top of the page can be used to search either by name or by tag.



Figure 2. BluePages+1 search result page

## Buddylist import

In order to save users time, we also provided them the ability to import their instant-messenger buddy list. During the import, folders (names given to groups of buddies) were translated into tags, such that each buddy in a folder was tagged with the name of the folder. Though this feature was not widely advertised, a few dozen users have discovered it.

## INTEGRATING TAGGING WITH OTHER APPLICATIONS

References to people appear everywhere: in email, on the web, in instant messaging, in documents. We believe that one ought to be able to tag people wherever they are encountered: while reading an email, chatting over instant messaging, or talking on the phone.

To this end, we have provided a REST [6] API that enables developers to access and contribute tag data. This API has resulted in a handful of visualizations and mashups with other internal systems. Although not directly tied to the results presented in this paper, the openness of the tagging system was seen as yet another channel to encourage people to explore and leverage the social aspects of people tagging.

For example, one of our colleagues has added support for tagging into an internal Firefox extension called Tommy!. This extension scans each web page for person identifiers (typically an email address, a link to the employee directory, or a link to someone's internal blog). Right-clicking on such a link causes a popup menu to appear with an enhanced "business card" for that person, using information drawn from the employee directory. One of the tabs in the business card displays a person's tag cloud (Figure 3), along with other visualizations of their tags. A user can also add or remove tags using the business-card interface. This plugin enables a user

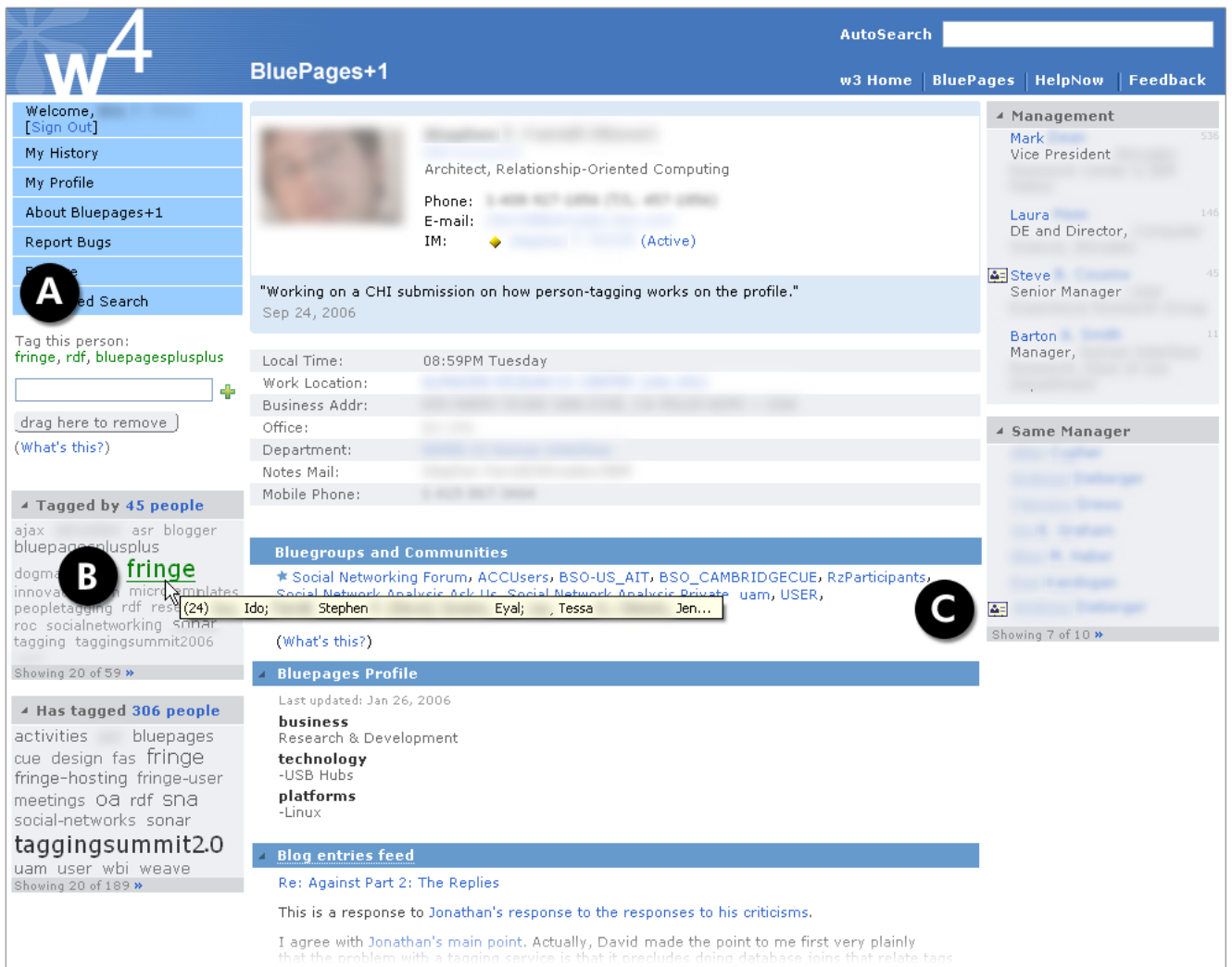


Figure 1. BluePages+1 profile page (details obscured for blind review)

to quickly discover information about people they encounter while surfing the web, simply by right clicking on their name or email, including tag information.

We have also developed a plugin for an instant messaging application that enables users to tag others, and view their tags, from within the context of a chat conversation (Figure 4). When a new conversation window is opened, the system displays the tags you have applied to your buddy, and the tags he/she has applied to you. During the conversation, the plugin provides commands to add a tag to the person, delete a tag, replace the set of tags with a new set, and display the person's tag cloud (list of all tags ascribed to this person by anyone, ordered by frequency). Changing someone's tags results in a message being generated in the chat conversation visible to both partners; this practice encourages users to spread awareness of people-tagging.

The plugin also provides a tag-based buddylist. The system retrieves the set of people the user has tagged, and displays them as a dynamically-generated buddy list. Hovering over

the name of each individual in the list brings up a tooltip with details about the contact, including your tags for that person. We also provide the ability to search for and initiate conversations with people based on tags. We are intrigued by the possibility of using tagging to manage one's contacts, and plan to study the use of these features in more depth.

We have also developed a plugin for the Thunderbird email client to integrate with people tagging (Figure 5). One can right-click on any email address to see the tags one has applied to this person, or to add new tags to this person. In the future, we envision using tags to define ad hoc mailing lists by sending email to everyone one has tagged with a particular tag. We would also like to provide the ability to look up people by tag while addressing an email message. For example, one should be able to address a message to "finance-person" and have the email client resolve that tag to the person currently in the role of financial analyst.

The large majority (70%) of tagging actions have been done from within the BluePages+1 interface, and a significant frac-

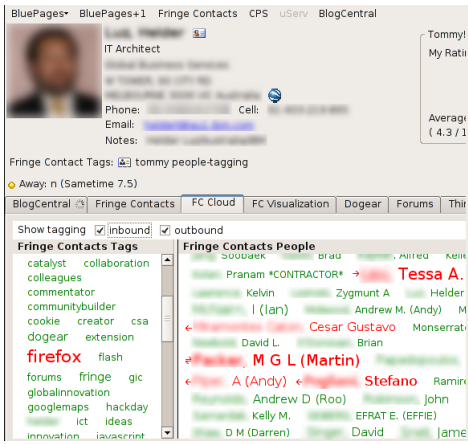


Figure 3. Tommy!, a Firefox extension that displays an enhanced business card, including tagging information

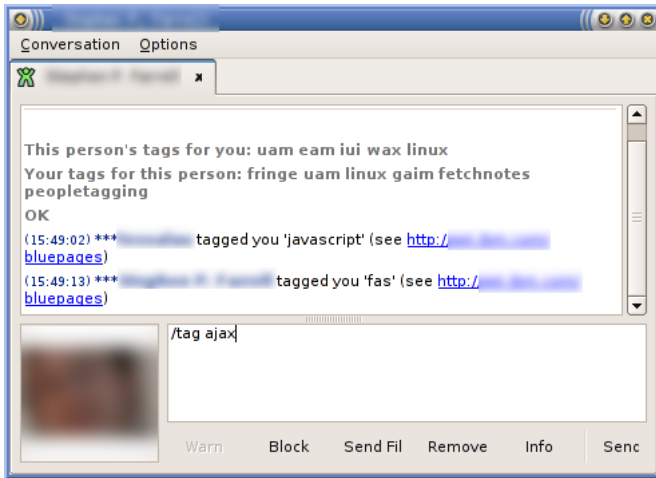


Figure 4. Tagging a chat buddy from within an instant messaging conversation

tion (6%) have been logged as coming from Tommy!. We also plan to integrate people tagging with the predominant email and instant-messenger clients in our company.

### TAG USAGE STATISTICS

In order to understand how tags were being used, we conducted an in-depth analysis of the tagging data based on usage logs recorded as of 21 July 2006. At that time, there were 10,273 records created by 716 taggers. Each record took the form of the tuple described above: (tagger, tag, taggee). Some tags were used only once. Other tags were used many times, either by the same tagger or by different taggers. Figure 6 shows how the usage of the system has grown steadily over time. In the figure, we call out the distinction between self-tagging behaviors (applying a tag to one's own directory record) and other-tagging behaviors (applying a tag to someone else's directory record); these behaviors will be discussed in more detail below.

We removed the data from three taggers, who were members of the original research team, in case their usage patterns

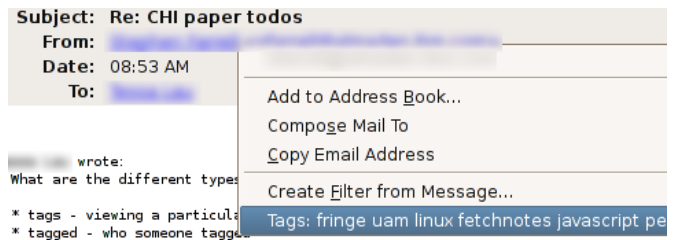


Figure 5. Tagging someone from email

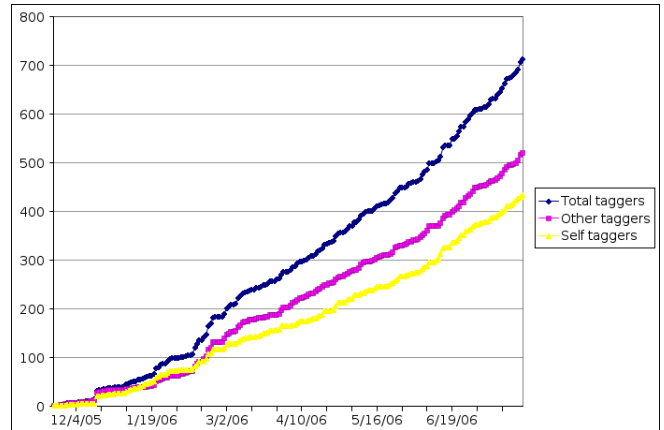


Figure 6. Growth in number of taggers over time

might have been influenced by their hypotheses about what the data would show. These restrictions produced a dataset of 9816 records created by 713 taggers, who used a total of 2992 unique tags to describe 7601 taggees. Removal of the research team data resulted in a 4% reduction of the dataset.

Table 1 presents a high-level view of the tagging behaviors. There are two ways to summarize the data: simple ratios applied to total numbers in the dataset, and aggregations of per-person statistics. In normally-distributed datasets, these two summaries would be similar. However, social-software phenomena are usually distributed according to the power law. We therefore present both types of summary in Table 1.

As a simple ratio across the entire dataset, each tagger was responsible for 4.20 tags, applied to 10.66 taggees. Each

	Basics	
Tuples	9816	
Taggers	713	
Taggees	7601	
Unique Tags	2992	
	Ratios	Per-Person Statistics (Mean (Median) Range)
Tags/Tagger	4.20	8.32 (3) 1-245
Taggees/Tagger	10.66	13.77 (2) 1-484
Tags/Taggee	2.54	1.97 (1) 1-59
Taggers/Taggee	0.09	1.29 (1) 1-39

Table 1. Summary of usage data as of 21 July 2006



Figure 7. Number of Self-only, Self+Other, and Other-only taggers

taggee was described by 2.54 tags created by 0.09 taggers.

Because of the power-law distribution, and because of redundancies in tag usage, the summary looks quite different when computed as aggregated statistics on a per-person basis. In the per-person analysis, each tagger was responsible for a mean of 8.32 tags (median 3, range 1-245) applied to a mean of 13.77 taggees (median 2, range 1-484). Each taggee was described by a mean of 1.97 tags (median 1, range 1-59) created by a mean of 1.29 taggers (median 1, range 1-39).

These statistics suggest that a small number of active taggers are responsible for tagging people in the community at large. Due to the public nature of the tagging system, the work done by the small group of active taggers creates content that benefits all.

### TAGS FOR IMPRESSION MANAGEMENT

We were surprised to find that 434 taggers (61% of the users) tagged *themselves*. As shown in Figure 7, these 434 taggers can be further analyzed in terms of people who tagged themselves only (192 taggers, 27% of the sample) and people who tagged *both themselves and others* (242 taggers, 34% of the sample). We called these two groups Self-only and Self+Other, respectively. There was of course a third group of people who tagged other people only (279 taggers, 39% of the sample). We called this group Other-only.

Because we had not anticipated a strong phenomenon of self-tagging, we analyzed the data in greater detail. As a rough measure of the effort that people put into tagging, we compared the mean number of tags-per-taggee that were used by each tagger in the three groups (Self-only, Self+Other, Other-only in Figure 7). Figure 8 shows a significant overall difference ( $F(2, 710) = 43.100, p < .001$ ). People in the Self-only group used the largest number of tags-per-taggee (2.74 tags-per-self). People in the Other-only group used the smallest number of tags-per-taggee (1.16 tags-per-each-other). All differences between groups were significant according to the Least Significant Difference test at  $p < .05$ .

These between-subjects results suggest that people put more effort into tagging themselves than into tagging others. To pursue this question in greater detail, we conducted a second, within-subjects analysis for the 242 people who had tagged both themselves and others (the Self+Other group in Figure 7), as shown in Figure 8b. The motivation for

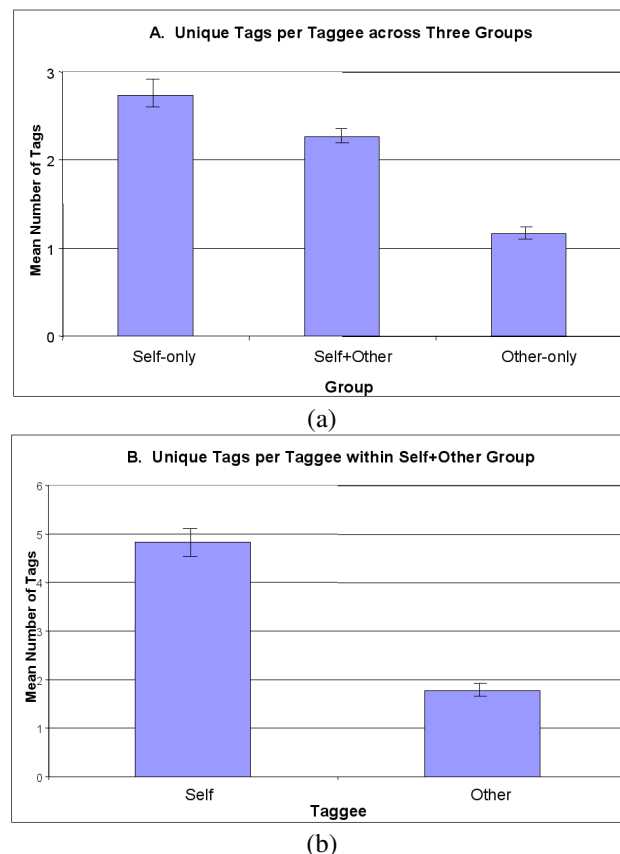


Figure 8. Comparison of Tags-per-Taggee. (a) Between-subjects analysis on three groups (Self-only vs. Self+Other vs. Other-only). (b) Within-subjects analysis on Self+Other group (Self vs. Other).

this analysis was to control for *differences among the persons doing the tagging* (i.e., each person in this group had tagged both self and other), so that we could focus on *differences in behaviors*. We analyzed tags-applied-to *self* vs. tags-applied-to *other* for each tagger in this group. Again, the mean number of tags-per-taggee measure showed significant differences ( $F(1, 241) = 107.31, p < .0001$ ), with more than twice as many tags applied to Self (4.83) as compared with tags applied to each of the Others (1.77).

Did people use the same words when tagging self as contrasted with tagging other? To answer this question, we again focused on the 242 people in the Self+Other group. We performed within-person comparisons of the tags used to characterize self vs. the tags used to characterize other. The mean number of tags in common was 2.37 (median 1, range 0-30), or (averaged across Self and Other for each tagger) 44% of the tags. Thus, there was substantial overlap in the tagging vocabularies across Self and Other. The differences that we described above (in the number of tags used per taggee for Self vs. Other) appears to be due to decisions about how much effort to put into tagging, and not to differences in tagging vocabularies.

These results may be interpreted as showing that people put more work into self-tagging than they did into tagging oth-

ers, as measured by both between-subjects (Figure 8a) and within-subjects (Figure 8b) analyses. These results are consistent with the phenomenon of impression management described by Goffman [7]. Goffman proposed an analysis of individuals' behaviors in organizations in terms of dramatic performances, in which the organizational actor was concerned in part with how s/he was perceived by other members of the organization. In this perspective, self-tagging appears to be an asynchronous component of impression management, through which one member of an organization attempt to influence how other members view her or him. The theme of impression management has continued in contemporary CSCW research, as reviewed above [2, 3, 4, 13].

### AUGMENTING PROFILES WITH TAGS

We have also investigated how tagging contributes to enhancing an employee's profile. Each employee at our corporation can update their profile (in the online employee directory, which is widely used) at any time using a simple web-based application. However, not all employees have chosen to populate their profile with meaningful prose, nor does everyone keep their profile up to date with their most recent projects or expertise. Our hypothesis is that people-tagging, in which others contribute meaningful words to describe people they know in the organization, can help populate employee profiles with more timely information about their interests and expertise, and thus enable people to find other people more effectively within a large organization.

We conducted this analysis with a larger dataset, collected from usage logs as of 12 September 2006. This dataset included a total of 25,904 unique (tagger, tag, taggee) tuples, created by 1158 taggers. We did not remove tags created by the original research team for this analysis.

### Comparing tags to prose

We observed several differences between the tags used to characterize people and the prose in their profiles. First, taggers often conjoin multiple words with dashes to form a single tag, such as "it-specialist" or "websphere-portal". These compound tags are unlikely to appear literally in plain text, but rather as phrases such as "IT specialist" or "WebSphere Portal". Thus we opted to split all text (tags and prose) on non-alphanumeric characters, as well as lower-casing all strings, to facilitate comparisons across the two different systems.

Second, we noted that different word forms are often used to express the same information. For example, someone may be tagged as a "designer" or with the term "design". As these different forms are conceptually similar, we used a stemmer [12] to canonicalize tags and words, and we base our comparison on the stemmed forms.

For clarity, we refer to the part of a tag that has been split on non-alphanumerics and stemmed as a *tag fragment*. After this preprocessing, our dataset contained 33889 unique (tagger, tag fragment, taggee) tuples.

### Datasets

For each employee, we collected the words visible on their profile. The profile includes text that the employee provided, such as a biography or statement of interests. It also includes organizational information not written by the user, such as job title and department name. The profile also includes a limited amount of data drawn from internal data sources such as the titles of recent publications and patents. All of this text typically appears on each employee's profile page in the current generation of our employee directory.

We denote the set of words on an employee's profile as  $T_{BP}$ . If  $frag(t)$  denotes the set of fragments of tag  $t$ , then let  $F_{BP}$  be the set of tag fragments generated from  $T_{BP}$  after tokenizing and stemming each tag in  $T_{BP}$ . In other words,  $F_{BP} = \bigcup_{t \in T_{BP}} frag(t)$ .

For the tags, we computed the set of words with which each employee had been tagged (this included self-tags—tags the employee ascribed to him or herself). The resulting set we denote as  $T_{in}$ , for incoming tags. Similarly, we also compute  $T_{out}$ , the set of outgoing tags this employee has used on himself or others. Let  $F_{in}$  and  $F_{out}$  be the corresponding sets of tag fragments generated from each of these tag sets.

### Tags contribute information

Our first analysis aimed to measure how much additional information is added to each person's profile through the use of tags. For each person who had been tagged in BluePages+1, we calculated  $F_{novel} = F_{in} \setminus F_{BP}$ , that is, the set of tag fragments in their incoming set that were not present in their profile text. These words might represent new information or things a person is known for that they do not choose to present on their profile, but that might be useful to others in trying to locate expertise.

Of the 12,566 people that had incoming tag fragments, 10,368 (83%) of them had a non-empty  $F_{novel}$ . On average, including people whose tags did not contribute any novel words, people acquired 1.65 novel tag fragments via people-tagging (median 1, stddev 2.27, range 0-55). While it is difficult to interpret the meaning of tags without asking the tagger or having familiarity with the subject matter, a cursory scan of the novel tag fragments reveals internal project names, attributes such as "mentor" or "coach", technical skills such as "eclips" or "ajax", locations such as "frankfurt" and "nyc", and roles such as "blogger".

Though our data are still preliminary, we believe this evidence shows that people tags already provide distinctive information about a person beyond what that person has provided in their profile. And as people tagging becomes more ubiquitous, this value will continue to grow.

### Tags provide ranking

A tag represents one person's wish to associate a label with another person. In this way, tags are conceptually similar to hyperlinks on the web. Just as Google's PageRank [5] uses links as "votes" to indicate the value of a page, we imagine using tags as votes to indicate the strength of a connection





Tags might be incorrect, or might have less-than-current information that a person might not want to be associated with.

One colleague blogged on the topic of incoming tags in BluePages+1 and solicited a number of responses. One respondent drew a distinction between intrinsic and extrinsic properties of objects, and suggested that blurring the distinctions could cause problems:

I think it's confusing (on Flickr as well as in BluePages+1) that tags given by random people to a particular resource are so closely associated with the resource itself, as though they were properties of the resource. Tags, to me, belong to the people who created them, and are aggregators, creators of collections.

Another reminded us that we've been living with tags or labels all along:

I suspect however that this is nothing new. We have forever grouped other people into categories (and multiple categories) such as 'friend', 'idiot', 'family', 'artist' etc. Sometime these categories have also been expressed both verbally and in written form. With social networking, this expression of tags is simply more efficient. Just as people have not liked their tags in the past, some will also not like some of their tags in the future.

## DISCUSSION

One of the interesting uses of tags is to collaboratively define a group of people. We observed this in action earlier this year: our colleagues used tags to keep track of who from our company was attending the CHI 2006 conference. Tagging was primarily done by one person in our department, but as word spread, others participated as well. Seven people self-tagged themselves with "chi2006". Two others besides the person in our department tagged an additional two people with the tag. All told, these eight people tagged a total of 38 people with this tag, illustrating a real collaborative use of tags to define a group.

A similar phenomenon has occurred among players of SecondLife, a virtual reality game. 115 employees have been tagged (or tagged themselves) "secondlife" by 47 others. Curious about the spread of this tag, the authors discovered a wiki page that served as a roster for those interested in this technology. This page included a link to the BP+1 tag, as well as the SecondLife identity of each person with a link to their BP+1 profile.

Our experience with a people-tagging system raised several social issues surrounding the use of tags to contribute to employee profiles. For instance, some of our colleagues have expressed concerns with the incoming tags—those that others have given to them. We have yet to investigate the general awareness of tagging in our employee population. 1,158 people have used the system to tag 12,567 people. How many of those taggees are aware of being tagged? How many agree with their tags, or would rather not have a par-

ticular tag associated with their profile? Although we have received no complaints from people wishing to have tags removed from their profile, we suspect that this may arise in the future as people tagging becomes more widespread.

The concerns seem to stem from the expectation that a user has ultimate control over his profile. The simple solution—to only show the outgoing tag cloud—is undesirable since about ten times as many profiles have incoming tags as have outgoing, and one of our goals is to leverage an active minority contributing data to many profiles. Therefore, we have been trying to understand the concerns that our colleagues have in depth so we can design features to compensate.

The specific concerns that have been stated range from slander (will someone tag me "jerk"?), to philosophical (incoming tags are extrinsic properties that should not appear as part of the profile). The concern of slander is largely mitigated by the combination of professional work environment, inability to tag anonymously or under a pseudonym, and the open display of the tagger's identity. Since occurrences of slander would probably undermine people-tagging, it's unlikely to work in environments that lack such controls.

A less easily dismissed concern is that incoming tags might offend the taggee despite positive intentions by the tagger. For example, one user might tag another "helpful" not knowing that the taggee is sensitive to being perceived as subservient. Similarly, some tags might describe a skill that the taggee doesn't want to disclose. For example, being tagged "windows-guru" could lead to unwanted solicitations for help. Even if the user is happy to disclose their association with a topic at one point in time, he might change his mind later, perhaps after switching projects. Even worse, undesired disclosure could be more personal and even run into institutional privacy guidelines. For example, the tag "brunette" might be inappropriate, especially when used to pivot (show everyone tagged "brunette").

The authors have discussed various ways to address the concerns about people-tagging. The first, implemented a few months after the system was deployed, is to ensure transparency: hovering over a tag lists the people who have used it. As suggested above, this has mitigated fears of slanderous uses of tags. Another approach is to change the presentation of the incoming tags to make it clear that they do not necessarily represent the skills or attributes of the person represented by the profile.

Another real issue we have observed is that there may be a need for tags to decay over time. For example, a tag representing the name of a project may be less relevant after the project ends. Tags associated with events that happen at a particular point in time, such as a summit or workshop, may assume less importance over time. In the current implementation, each tag retains equal importance regardless of age. But as tags accumulate, the tag cloud associated with a person may become more polluted with tags that are no longer relevant. An open research question is how to downweight the importance of some tags while retaining the tags that rep-

resent long-standing interests or expertise.

### CONCLUSION AND FUTURE WORK

We have described the use of people tagging to enhance our BluePages+1 employee directory. We have shown how people tagging enhances the basic profile information in the directory, contributing novel information about employee interests and expertise beyond what they have written about themselves. The collaborative nature of tagging distributes the work, making it possible to leverage the collective wisdom of the company to describe people's interests and expertise. The distributed nature makes it possible to use tags as a form of ranking, where each tag consists of a vote towards associating a person with a particular topic.

One of our interesting findings was extensive self-tagging, which we interpreted as a form of impression management. While this kind of impression management might appear to be a matter of vanity or self-promotion, we understand it as a kind of organizationally responsible behavior for people in a knowledge-intensive company (e.g., [1]). Knowledge work often involves finding opportunities to contribute to collaborative work, and one way of finding those opportunities is to "advertise" one's skills to other members of the organization. Self-tagging would support the creation and refinement of such a public persona. Our results so far encourage us to look for other uses of tagging (e.g., of objects [9]), for evidence of similar attention to how the tagger may be perceived by colleagues. The next steps in our research will examine that hypothesis in other collaborative media.

We see many directions for future work. Our empirical analysis, based on usage data, only tells part of the story for how people tagging affects the enterprise. We plan to use surveys to investigate how people use, perceive, and benefit from people tagging. We also anticipate comparing the tagging of people in BluePages+1 with tagging behaviors in other enterprise-oriented tagging systems, such as webpages and resources [9] and activities and their components [10].

People tagging is one component of a generalized social networking architecture called Sonar that we are developing to help manage relationships within the enterprise. Sonar defines an API for social network data sources to provide information about how strongly two people are connected. Future work will investigate how data sources such as people tagging, social bookmarking, email communication, and file commonalities can be integrated in a common framework, and used to facilitate relationship management.

### REFERENCES

1. Ackerman, M. and Halverson, C. Sharing expertise: The next step for knowledge management. In V. Wulf and M. Huysman, eds., *Social Capital and Information Technology*, MIT Press, Cambridge, MA, USA, 2003.
2. Beard, J., ed. Greenwood, Westport, CT, USA, 1996.
3. Becker, J. and Stamp, G. Impression management in chat rooms: A grounded theory model. *Communication Studies*, 56, 3, 243–260.
4. Boyd, D. Identity production in a networked culture: Why youth heart myspace. Talk at AAAS 2006 (part of panel: "It's 10PM: Do You Know Where Your Children Are ... Online!"). St. Louis, Missouri: 19 Feb, 2006.
5. Brin, S. and Page, L. The anatomy of a large-scale hypertextual web search engine. In *Proceedings of the seventh international conference on World Wide Web 7*. 1998, 107–117.
6. Fielding, R. *Architectural Styles and the Design of Network-based Software Architectures*. Ph.D. thesis, University of California, Irvine (2000).
7. Goffman, E. *Presentation of Self in Everyday Life*. Doubleday, New York, 1956.
8. John, A. and Seligmann, D. Collaborative tagging and expertise in the enterprise. In *Workshop on Collaborative Web Tagging, WWW 2006*. 2006. [Http://www.rawsugar.com/www2006/26.pdf](http://www.rawsugar.com/www2006/26.pdf).
9. Millen, D., Feinberg, J., and Kerr, B. Dogear: Social bookmarking in the enterprise. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM Press, New York, NY, USA, 2006, 111–120.
10. Moore, M., Estrada, M., Finley, T., Muller, M., and Geyer, W. Next generation activity-centric computing. *Demonstration at CSCW 2006*.
11. Nardi, B., Whittaker, S., Isaacs, E., Creech, M., Johnson, J., and Hainsworth, J. Integrating communication and information through contactmap. *Commun. ACM*, 45, 4 (2002), 89–95.
12. Porter, M. An algorithm for suffix stripping. *Program*, 14, 3 (1980), 130–137.
13. Spitzberg, B. Preliminary development of a model and measure of computer-mediated communication (cmc) competence. *Journal of Computer-Mediated Communication*, 11 (2006), 629–666.
14. Turner, J., Grube, J., Tinsley, C., Lee, C., and O'Pell, C. Exploring the dominant media: How does media use reflect organizational norms and affect performance? *Journal of Business Communication*, 43, 3 (2006), 220–250.
15. Xu, Z., Fu, Y., Mao, J., and Su, D. Towards the semantic web: Collaborative tag suggestions. In *Workshop on Collaborative Web Tagging, WWW 2006*. 2006. [Http://www.rawsugar.com/www2006/29.pdf](http://www.rawsugar.com/www2006/29.pdf).