Bloggers' Expectations of Privacy and Accountability: An Initial Survey

Fernanda B. Viégas
Media Laboratory
Massachusetts Institute of Technology

Go to a section in the article:

Abstract

This article presents an initial snapshot, based on an online survey of weblog authors, of bloggers' subjective sense of privacy, and of their perceptions of liability. The findings suggest that the social norms of bloggers are emergent and self-imposed. When confronted with questions of defamation and legal liability, respondents in the survey expressed contradictions between their actions and their knowledge of how the technology works. They generally believed that they were liable for what they published online, although they were not concerned about the persistence of their entries. In general, bloggers do not feel as if they know their audiences. For the most part, blog authors have no control over who accesses their entries, and this inability to define their audiences leads them to make a number of assumptions about who their readers are.

Introduction

Blogs, more formally known as weblogs, have captivated the media and, by extension, the public's imagination. The number of blog writers and readers has grown consistently in the last few years.[1] Moreover, blogs are permeating most niches of social life, addressing a range of topics from scholarly and political issues (Glenn, 2003) to family and children's daily lives (Turnbull, 2004). As blog writers become increasingly prolific, however, they are likely to encounter issues of privacy and liability. For example, accounts of bloggers hurting friends' feelings or losing their jobs because of materials published on their sites are becoming more frequent (Bray, 2004; Nussbaum, 2004; O'Shea, 2003; Pax, 2003; Phillips, 2003; Sarnataro, 2003; St. John, 2003; Whitworth, 2003).

Because blogging has the power to affect not only the lives of bloggers themselves but also of the people, companies, and products that are "blogged," it is important to understand how accountability and privacy expectations function in this emergent arena. This article reports on the results of a survey of blog authors' subjective sense of privacy and their perceptions of liability as regards the content of their weblogs. Its primary goal is to introduce privacy concerns into the ongoing academic discussion of

blogging practices. An additional goal is to describe bloggers' current expectations as regards privacy and liability for purposes of comparison with future phases of blogging activity.

The findings of this survey suggest that self-imposed social norms are emerging among weblog authors. For instance, bloggers are beginning to create informal guidelines for publishing the names of people and companies in their blog entries. For the most part, however, authors have no way of knowing who their readers are—a situation that has serious implications for both privacy and accountability. At the time of this study, most blog hosting services—with the exception of LiveJournal, which is discussed in the survey—did not provide authors with mechanisms for articulating distinct audiences for different entries. A significant portion of problems reported by the participants in the survey stems from lack of control over who has access to their writings.

The next section of this article provides background information on privacy issues in relation to blogging. The following sections explain the methodology used in the online survey and describe the results from the study. A discussion section follows, where the results are interpreted in light of the existing literature on blogging practices. Finally, the concluding remarks summarize the major findings and point to the need for further work in this area.

Background

Privacy in Networked Environments

In *The Unwanted Gaze*, Jeffrey Rosen (2000) highlights the contextual basis of privacy violations, stating that disclosure of personal information is a highly circumstance-sensitive matter. When taken out of context, the same information can be severely misjudged by others. In blogging practices, blog sites arguably serve as the context for the entries they contain. There is no guarantee, however, that individual entries will not be plucked out of their original context and exposed to radically different forums in the future.

Grudin (2001) refers to this "loss of control" as the steady erosion of *clearly situated action*. "We are losing control and knowledge of the consequences of our actions, because if what we do is represented digitally, it can appear anywhere, and at any time in the future. We no longer control access to anything we disclose" (p. 11). This "desituating" phenomenon can easily be applied to blogs. If bloggers experience the authoring environment as an intimate, secluded space—and the evidence suggests that many do view it as "their" space—it may not occur to them that, once their thoughts are published on the Web, they automatically become part of the most public, fragmented environment in existence today.

Palen and Dourish (2003) propose that privacy in networked environments is a dynamic, dialectic process of negotiation that is conditioned by people's own expectations and experiences and by those of others with whom they interact. In agreement with Rosen (2000) and Grudin (2001), Palen and Dourish note

that these negotiation processes are fundamentally dependent on people having control over their information and over the contexts in which that information is presented. How much control do bloggers have over the information they disclose? As will become clear, unfortunately, most blog authors today have little or no control over who gets to read their blog entries.

Technologies such as mass storage and search engines enable the misuse of persistent personal information. If the past is any guide, bloggers should be mindful of what information they publish online. The history of an earlier online social space, Usenet newsgroups, tells a cautionary tale. DejaNews (http://groups.google.com/googlegroups/deja_announcement.html) started archiving Usenet newsgroup discussions in 1995, providing a wake-up call for many newsgroup users. Before DejaNews existed, users knew that their conversations were public and able to be archived, but nobody paid much attention to those features. As Grudin (2001) points out, "newsgroup discussions felt like conversations in a room, ephemeral chatter among like-minded people" (p. 13). People felt safe in the obscurity of their newsgroups, and non-mainstream communities flourished in the Usenet environment. When DejaNews was launched, however, it sent a chill throughout Usenet. People realized that their most intimate discussions were suddenly aggregated, searchable, and available to anyone who typed their username into DejaNews. Employers started searching for the names of interviewees before making hiring decisions. Some previously thriving online communities collapsed, as users realized that it was not safe to express intimate thoughts and feelings on the open Internet. Suddenly, many people became painfully aware of what it meant to have personal data persistently available in a computer network.

Blog authors voluntarily make extensive archives of their personal thoughts and interests readily available to anyone surfing the Web. Unlike the problems that arose with the advent of DejaNews, here it is the author him- or herself who chooses to present personal information in one place, aggregated and easily searchable. However, not all blog authors may be conscious of the potential repercussions of these actions.

Blogs from a Privacy Perspective

By their very nature, blogs raise a number of privacy issues. On the one hand, they are persistent and cumulative. At the same time, they are easy to produce and disseminate, resulting in large amounts of sometimes personal information being broadcast across the Internet. Each of these points is discussed below.

One of the fundamental differences between blogs and other web-based publishing venues such as home pages is that blogs are typically archive-oriented. Instead of substituting new materials for old ones, as is normally done on regular web pages, the idea with blogs is to *add* postings frequently, creating an ever-growing compilation of entries. This simple feature, the ability to append without erasing any of the previous content, makes blogs fundamentally different from other kinds of web sites.[2] The mounting compilation of postings serves as context for readers of blog sites. By allowing readers to examine earlier

postings at will, blogs provide an informative backdrop against which to situate more recent writings.

Readers can return to a blog frequently with the expectation of looking at new posts by the same author. Regular readers thus can get a sense of the identifying "voice" behind the posts on the site. More than with any other kind of web publishing, authorship is central to blogs. On a blog, every word is clearly and inescapably associated with its author. Even on group blogs, where multiple authors post entries, the tendency is for the author of each entry to identify him/herself—group blogs usually identify individual authors by attaching a "posted by" stamp next to each entry. Over time, a blog archive can read very much like an evolving portrait of its author's interests and experiences. This reliance on authorship also means that blogs are fertile ground for liability disputes.

In addition to the persistent and cumulative nature of blogs, there is the ease and speed with which bloggers can add new text to their sites. Because the cost of publishing is low, and changes can be broadcast instantly over the web, authors tend to add entries to their blogs frequently and spontaneously, creating highly dynamic sites. As Evan Williams, one of the creators of Blogger web authoring software, points out, "the blog concept is about three things: Frequency, Brevity, and Personality" (Turnbull, 2001). This spontaneity means that, at times, bloggers publish impulsive, ill-considered messages. Some authors generate unusually high volumes of output—typically with a noticeable decrease in the quality of content—a phenomenon known to bloggers as "blogorrhea" (Havilland, Cronin, & Amon, 2004).

This combination of manifest authorship, spontaneous publishing, and automatic broadcast is potent, and has a number of implications for online privacy, freedom of speech, and liability. The survey conducted for this study explores the tension between the publishing freedom experienced by bloggers and the potentially problematic consequences of public, persistent blog entries.

Previous Blog Research

For the most part, blog culture has been documented by the media and by bloggers themselves (Blood, 2000; Dibbell, 2002; Turnbull, 2002). Recently, however, a growing body of academic work has started to analyze the makeup of the blogosphere (the collective term encompassing all weblogs). This section briefly reviews the academic literature as it relates to the four axes of inquiry in this survey: content, identity management, audience, and persistence.

Content

According to Blood (2000), the earliest blogs were link-driven sites: "Editors present links both to little-known corners of the Web and to current news articles they feel are worthy of note [...]. These weblogs provide a valuable filtering function for their readers. The Web has been, in effect, pre-surfed for them." Such filter sites, however, have ceased to account for the majority of weblogs. Herring and colleagues (Herring, Scheidt, Bonus, & Wright, 2004) coded a random sample of blogs for overall use based on the nature of the content posted; counter to the popular view, they found that the

overwhelming majority of blogs (70.4%) were of the personal journal type: "in which authors report on their lives and inner thoughts and feelings" (p. 10). Although filter blogs—those that focus on external, public events—are usually assumed by the media and members of the blogging community to be representative, they accounted for only 12.6% of the blogs in their sample. Knowledge blogs[3] accounted for even less, with 3% of the sample. The study also found that the common belief that blogs are extensively interlinked, and interactive, is unfounded. The average blog entry in the sample received only 0.3 comments, and the majority of entries received none. Moreover, one-third of blog entries contained no links at all. When links were present, they rarely led to news sites or other blogs as is usually assumed.

In a subsequent study, Herring, Kouper, Scheidt, and Wright (2004) looked at gender- and age-based differences in the content of blog sites. Their findings suggest that one of the main reasons for the disconnect between popular perceptions of blogs and reality is the disproportionate attention paid by the media and the blogging community to filter blogs, which are primarily produced by adult males.[4] The researchers found that women and teenagers—who make up over half of the blogging population—tend to write personal journal-style blogs that end up not being part of the public discourse about weblogs.

Identity Management

In a small, qualitative study, Nardi, Schiano, and Gumbrecht (2004) found that most bloggers were surprisingly open on their sites, some of them choosing to reveal very personal information. Many of the bloggers in their sample provided accurate identity information and even contact information (via links to personal home pages). These findings echo those of Herring, Scheidt, et al. (2004), where 92.2% of the sampled blogs included explicit personal information on their first pages. Full names were found in 31.4% of sites, first names on 36.2%, and pseudonyms appeared on 28.7% of blogs. More than half of the weblogs in the sample provided other kinds of explicit personal information such as age, occupation, and geographic location.

Bloggers also have the option of publishing anonymously. Sites such as invisiblog.com (http://www.invisiblog.com) offer services that allow authors to anonymously create and maintain a blog site. Fusing existing privacy technologies with a tool for blogging, these programs make it easy to broadcast thoughts and views incognito, as Invisiblog states on its site:

> Unlike other blogging and hosting services, Invisiblog doesn't have access to any potentially sensitive information about you—not even your email or IP address. All contact between bloggers and invisiblog.com is via the Mixmaster anonymous remailer network, which uses encryption and 'mixnet' techniques to hide the source of an email. (invisiblog.com/info/faq/)[5]

As of July 2004, Invisiblog listed 188 blogs on its directory and it had no reports of court orders or subpoenas for content or log files.

Different forms of anonymous blogging behavior have also emerged in other venues. Bandlow and Jensen (2002b) report on a recent trend in LiveJournal (http://www.livejournal.com) entries:

> Instead of really writing anything of interest yourself, you invite people to make anonymous comments to this post. They can post whatever they want, stuff they are too embarrassed to discuss with anyone, stuff they have never told anyone before. IP logging is turned off, and as long as the people remember to make an anonymous post (an option you have to check if you are a LiveJournal member), it is an anonymous post. (p. 7)

The authors of the study state that they saw about five LiveJournal sites that hosted such anonymous entries, each one containing well over 300 posts; some users apparently found this to be a truly cathartic experience while others likened it to going to confession.

Audience and Control Features

Bloggers' relationship with their audiences seems to be one of the high points of blogging. Nardi et al. (2004) have characterized blogs as a "studied minuet between blogger and audience" (p. 4), and have distinguished two kinds of audiences: the bloggers' own, known social networks, and a larger audience beyond the author's friends and family. The study found several reasons why authors blog and, significantly, every reason was associated with the existence of an audience:

- building chronicles of everyday events that serve to update others on bloggers' lives
- expressing opinions to influence others
- seeking others' opinions and feedback
- thinking by writing: The study points out that "while 'thinking' might seem a solitary activity, or one not quite social, in blogging the presence of the audience and the writer's consciousness of the audience clearly introduce the social into an individual's thought process" (Nardi et al., 2004, p. 6)
- releasing of emotional tension: Even in an activity like this, where the purpose of blogging seems private, the same attention to audience was present; it was emotional tension *with an audience* that was found to be especially powerful for bloggers.

Nardi et al. (2004) also found that, even though bloggers delighted in their audiences, there was a clear desire to keep the audience at arm's length: "interactivity was valued, but only in controlled small doses" (p. 6). A related article, reporting on the same study and set of participants, has characterized blogs as "protected space" (Gumbrecht, 2004). Drawing on the limited interactive nature of blogs first noted by Herring, Scheidt, et al. (2004), Gumbrecht claims that because readers' comments are subservient to the blogger's posts, blogs create a protected arena in which authors feel safe to express emotions and experiences. The study also states that authors use ambiguous language and references in order to protect themselves and select their audience. When posts are ambiguous enough, only those readers who know what is happening in the blogger's life are able to understand the meaning of what has been published.

LiveJournal blogs give authors more control and flexibility than most other blogging services. LiveJournal

allows authors to create lists of friends and manage the access to specific entries via group filters and locks. As Bandlow and Jensen (2002a) note:

> Group filters can be used to help manage the readability of a large friends list or in conjunction with locks do add privacy by controlling who can read posts. If a post is locked and you aren't allowed to read it, it simply won't appear for you, so you don't know you're missing a post. If you are allowed to read it, you will see a small lock icon next to it. (p. 10)

Persistence

Studies have found that bloggers tend not to express concern about the persistent nature of blog entries. Nardi et al. (2004) found that the prospect of having personal information floating publicly on a networked environment did not bother the bloggers in their sample. For the most part, they were indifferent to it: "the bloggers in our sample seemed fairly unconcerned about privacy" (p. 7). The study identified three levels of publishing control available to bloggers: password-protected weblogs, unlisted blogs (not listed by the blogging service's directory), and listed sites (easily found by search engines). Password-protected blogs—the most "protected" kind of blogs—were almost nonexistent in the sample studied. Only one of the participants had a password-protected blog, which he used for sharing pictures of his daughters with relatives. This same blogger also kept a public blog.

Bandlow and Jensen's first study of LiveJournal (2002a) also reports that users were aware of and comfortable with the possibility that their journal postings would continue to exist for a long time. Their second study (Bandlow & Jensen, 2002b), however, reports that bloggers who had bad experiences became acutely aware of the potentially negative consequences of having public, persistent entries on the Web.[6] This finding suggests that there is a disconnect between the way users say they feel about the privacy settings of their blogs and how they react once they experience unanticipated consequences from a breach of privacy.

Methodology

The findings presented here are taken from a survey conducted between January 14 and 21, 2004. During that time, 492 bloggers responded to an online questionnaire, which consisted of 34 multiple-choice Likert-scale questions, plus 16 open-ended essay questions to allow for further qualification of answers (see Appendix A). Out of the total number of responses received, 486 were selected for data analysis; the other six responses were incomplete and were disregarded.

The respondents to this survey were not randomly selected but were found through a variant of the snowball-sampling strategy. Announcements for the online survey were posted to mailing lists in a major university in the New England area as well as on a few high-traffic blogs published by people known to the author of this article. The viral nature of blogs meant that the links to the survey page quickly spread to several other blogs. However, the resulting population of participants does not qualify as a

random sample and, as such, the results from this survey cannot be generalized to the entire blogging population. Rather, the findings are representative of certain niches of the educated, English-speaking blogging world.

Participants answered questions about their blogging practices and their expectations of privacy and accountability when publishing online. Authors also answered questions about whether they had experienced any problems due to materials posted on their blogs.

The survey focused on four key areas:

1. **Content**
   What kinds of contents do authors publish?
   Are authors getting in trouble—legally or personally—because of materials published on their blogs?
2. **Identity management**
   Do authors identify themselves on their blogs?
   Do authors identify others on their blogs?
   Is identity disclosure management dependent on content? (i.e., do authors refrain from revealing people's identities when writing about sensitive matters?)
3. **Audience and control features**
   Do authors change their blog settings to limit who has access to their entries?
   Do bloggers have a good idea of who their readers are?
4. **Persistence**
   What are authors' attitudes towards the persistent nature of their published entries?
   Do authors ever edit posts after they have been published?

These four aspects of blogging are not handled independently by authors. For instance, identity disclosure is highly dependent upon bloggers' expectations of who might be reading their sites, and upon their perception of how persistent entries are. Once an author realizes that her blog entries are persistent and easily searchable, she might change the way she discloses certain aspects of her identity. Likewise, if a blogger has no good way of knowing who his readers are, he might decide to refrain from publishing details about his private life. Therefore, even though the questions on the survey were designed to cover four distinct areas, the results from one area have implications for the other areas as well.

Results

Study Population

As Table 1 indicates, the majority of respondents in this study (63%) are male. Even though a few of the popular blogging sites attract mostly teenagers,[7] our respondents tended to be older, with almost half of them between 21 and 30 years of age (Table 2). The overwhelming majority of participants were Caucasian (Table 5). Perhaps not surprisingly, given that the survey questionnaire was available only in English and that announcements for the survey were posted to mailing lists in American universities, 67% of the respondents reported they lived in the United States (Table 6). The education level of respondents

was quite high: 59% had been to four-year colleges or universities, and an additional 31.1% had attended graduate school (Table 3). Finally, most participants in this study (67%) have been blogging for over one year, with 36% having blogged for over two years (Table 4).

In general, these demographic characteristics resonate with findings from other blog surveys (Herring, Scheidt, Bonus, & Wright, In press) as well as with other public Internet social environments such as discussion forums (Herring, 2003) and personal homepages (Döring, 2002). As Herring and colleagues have pointed out, participants in these spaces "tend to be young adult males residing in the United States" (Herring, Scheidt, et al., 2004, p. 5).

Sex distribution of respondents	
Female	36%
Male	63%
Undisclosed	1%

Table 1

Age distribution of participants	
below 20	10.5%
21 - 30	46.3%
31 - 40	28.2%
41 - 50	11.3%
51 - 60	3.1%
above 61	0.4%

Table 2

Education level of participants	
elementary school	0.2%
high school	9.7%
college/university	59%
masters/professional degree	26.2%
Ph.D.	4.9%

Table 3

Length of time blogging	
< 3 months	6%
3 - 6 months	9%
6 months - 1 year	18%
1 - 2 years	31%
> 2 years	36%

Table 4

Ethnicity of respondents		Current place of residency	
African	0.4%	U.S. total	67.5%
African-American	1.0%	Non-U.S. total	30.4%
Asian	4.1%	Undisclosed location	2.1%
Asian-American	3.3%		
Australian	1.4%		
Caucasian	78.6%		
Latino/Hispanic	2.1%		
Native American	0.6%		
Pacific Islander	0.4%		
Other	7.8%		

Table 6 (appears beside the right table)

Table 5

Content

In agreement with results from previous studies, this survey found that the so-called "filter blog" does not seem to be the dominant genre of blogs. Most respondents (83%) said their entries could be characterized as "personal musings" whereas a smaller portion of people (20%) described their entries as being mostly compilations of links.

When asked how often they had posted highly personal materials on their blogs, 25% of respondents said they had done so fairly often. Only 19% of respondents said they had never posted anything highly personal on their blogs.

Respondents also revealed that the question of whether certain materials were too personal to blog about weighed heavily in their decisions of what to publish. Most respondents (62%) said they had considered whether certain topics were too personal to write about on their blogs fairly often. This suggests that the question of where to draw the boundaries between publishable and non-publishable materials is of concern to bloggers at present.

When asked how often they had gotten in trouble because of things they had written on their blogs, 64% of respondents said they had never gotten in trouble. The rest of the respondents (36%) said they

had gotten in trouble to some extent, with 6% saying they had frequently gotten in trouble for materials published on their blogs. In addition, 34% of respondents said they knew other bloggers who had gotten in trouble with family or friends because of content published on their blogs.

The survey results also reveal that the frequency with which a blogger writes highly personal things is positively and significantly correlated to how often they have gotten in trouble ($r = 0.3$, $p < 0.01$). This means that the more personal materials a blogger publishes, the more likely he/she is to get in trouble with others.

When asked to explain the kinds of problems they had encountered because of materials published on their blogs, respondents described the following situations:

> A message that I posted insulted a co-worker at a non-profit where I work. The non-profit suspended me from my work for three months.
> - a male blogger from New York

> I got 7 extra weekend duties for criticizing an army officer, even though I didn't name him, but merely alluded to him on my blog.
> - a male blogger from Singapore

> A TV station that I worked for [�] found my writings about people at the station, from a personal standpoint and said that I had spoken to some in a (subjectively) non-flattering light. After seeing these public entries in my blog, I was fired.
> - a male blogger from North Carolina

These stories echo those found on newspapers and magazines about bloggers who have lost their jobs because of materials they published on their sites (Bray, 2004; Phillips, 2003; Sarnataro, 2003; St. John, 2003). Clearly, the notion of what defines "socially acceptable behavior" on blogs has yet to be collectively defined.

Even though respondents of both genders characterized their entries as "personal musings" in this survey, female respondents tended to do so more often: 92.5% of females described their postings as "personal ramblings" as opposed to 77.5% of males. This tendency confirms findings from Herring, Kouper, et al. (2004).

Identity Management

Identification of People (Self and Others)

For the most part, respondents identified themselves on their blogs. 81% of participants said they used some form of self-identification: 55% used their real names, 20% used some variant of their name (only first name, a nickname that friends knew, initials, etc.). This tendency for self-identification has also been found in other studies (Herring, Scheidt, et al., 2004; Nardi et al., 2004) and it suggests that authors are

not overly concerned with preventing their genuine identities from being disclosed on their blogs.

Bloggers write not only about themselves but often also about other people with whom they interact. When asked whether they sought other people's permission to blog about them, 66% of respondents almost never asked permission, and only 3% said they always asked permission first. Interestingly, only 9% of the survey respondents said they *never* blogged about people they knew personally. Thus the great majority of respondents write about people they know but most of them never ask their permission to do so. Furthermore, no correlation was found between how private the content of an entry is and whether authors ask permission to write about the people they know. In other words, respondents who tend to write about more private situations or experiences do not ask permission from their friends or acquaintances any more frequently than those who do not write personal entries.

When asked whether they identified other people by name on their blogs, 21% of respondents said they almost always revealed names, whereas 42% said they almost never did. This indicates that most authors were somewhat sensitive to issues of privacy when blogging about friends and family. When asked to further explain why they chose to reveal (or not) people's names on their blogs, several respondents clarified that they had encountered trouble with acquaintances in the past because they had disclosed their names on blog entries. After having gotten in trouble, respondents became more sensitive to the issue of identification of others in their entries.

The responses to the open-ended questions about personal identification of self/others revealed an array of guidelines that authors followed when deciding how to blog about others. Some quotes of respondents' typical behavior follow:

Bloggers versus non-blogger identification:

> If they are not bloggers, I will not reveal their names. Bloggers have put their lives out
> there already, but my friends/acquaintances have not done so. I'll give them nicknames or
> refer to them as "J" or "T" rather than give their real names.
> - a female blogger from Tennessee

> If they are bloggers, I know I can go ahead and mention their names without permission.
> If they are non-blogging personal friends, I would never mention their name. The names
> of bloggers and others, I consider to be in the public domain anyway, i.e. from news
> reports, articles, academia etc.
> - a female blogger from England

Content-based decision:

> If it is a friendly blog, I use their name; if it is someone I hate (and know personally), I
> won't use their name but will refer to them somehow.
> - a male blogger from Delaware

> If I'm writing about someone personally and I think it could come back to bite me in the
> rear, they receive a stage name. If the content is innocuous—like, I went out to eat with
> Johnny and we saw a funny dog—then I use real names.
> - a male blogger from California

As can be seen from these quotes, some authors are constructing social guidelines for what they deem appropriate blogging behavior as regards the identification of others. Below is a summary of the main directives described by respondents:

Strategies for not fully revealing the names of friends and family members included disclosing:

- a person's first name only
- initials or first name and an initial for the person's surname
- pseudonyms or nicknames
- the nature of relationship, e.g., "my daughter", "my father", etc.

Other naming conventions included:

- revealing the person's name if they have an online persona; for instance, if someone is a blogger, respondents will link back to the person's blog and use the same name or handle the blogger has chosen to use for themselves. This is also true of people who are not bloggers but who have some other kind of Web presence (home pages, for instance)
- several participants explained that, if they have good things to say about their friends and family friends they will, for the most part, reveal their names; if, however, they post negative comments about a person, they will often conceal the person's name

Identification of Commercial Institutions and Products

In addition to deciding whether to identify individuals, bloggers also have to make decisions about whether to identify public entities such as companies and products when writing about them. Blogging about commercial institutions makes one potentially liable for defamation in costly legal suits. For this reason, respondents were asked about the decisions they make when they publish entries about public institutions.

Only 7% of respondents said they never blogged about companies or products. However, the overwhelming majority of respondents (84%) never ask permission of companies to blog about them; only 1% of respondents said they always asked permission before blogging about companies. Unlike the high degree of sensitivity displayed in relation to people they knew personally, most participants in this study (71%) said they almost always revealed the names of companies and products about which they blogged.

Here, too, respondents were asked to further explain their publishing decisions in open-ended questions. These answers revealed the emergence of additional self-imposed social norms that address why and when to identify companies and products. One of the main issues that surfaced from these answers is the question of whether to blog about the company that employs the blogger. For the most part,

respondents explained that they treat their current employer differently from other companies; for several bloggers, the current employers are off limits, so there is no mention of the company on their site. Other people felt comfortable blogging about their employer as long as they followed two guidelines: (a) They do not say anything bad about their employer, and (b) They do not identify the company by name. Furthermore, some of the respondents explained that they had gotten in trouble with employers in the past for things they published on their blogs and had, since then, changed their blogging practices so that there was no further mention of their employers on their sites.

Below are some quotes that reflect respondents' typical attitudes regarding public institutions and commercial products:

Good content vs. bad content:

> Depends again on the nature of the post. Good things - always show the name, to promote what they are doing. Bad things - depends :)
> - a male blogger from Finland

For other people's information benefit:

> I always make sure to inform my readers of good products, so that they can buy/use them too, and warn them (while trying to avoid libel) of very bad products, so they know what to stay away from.
> - a male blogger from Norway

Other companies are fair game except for the place where one works:

> Companies and products are public entities. The exceptions are the company I work for now and the one I worked for just before this one; they are never mentioned by name.
> - a male blogger from the southern part of the U.S.

> I always revealed the names of companies mentioned in the blog if, for nothing else, context; UNTIL I nearly got fired for a rather oblique mention of my place of employment. Now, I never write about the office directly, but do not hesitate to use the real names of any other company.
> - a male blogger from California

> For reasons of intellectual property and competitive intelligence, I obscure the names of my employer, my parents' company, and any companies I have job interviews with. Other than that, I have few qualms recommending or bitching about companies I do business with.
> - a female blogger from Massachusetts

The reasons respondents gave for revealing or concealing the names of companies and products on their blogs can be summarized as follows:

Reasons for revealing the names of companies on blog postings:

- providing a service to other people; that is, informing potential consumers about products
- if the products are good, the company deserves the credit

Reasons for concealing the name of a company:

- if the company is the blogger's current employer
- if the company is a former employer
- if the blogger is interviewing with the company

One of the issues that these strategies for identification of companies seem to overlook is the fact that companies ones blogs about today could become potential employers tomorrow. In that case, would the company in question consider hiring a person who has blogged about them in an unflattering manner? Will companies read candidates' blogs before making hiring decisions, similar to what happened with Usenet newsgroups in the past? One of the respondents in the survey alluded to this kind of situation:

> Early on in my job hunt, I applied for a low-paying job and then agonized [on my blog] over whether to take it. The next day, I noticed several hits to my journal from that employer, and they never called me back. I have no way of knowing for certain, but I suspect my blog post may have cost me that job.
> - a female blogger from Massachusetts

Control Features

When asked whether they do anything to limit who gets to read what they post, 76% of respondents said no. Several of the respondents who said they limited readers' access to their entries maintain blogs on LiveJournal. Unlike other blogging applications at the time of this study, LiveJournal has special features that allow one to exercise a finer degree of control over who reads what on one's site. The LiveJournal site defines its privacy features (http://www.livejournal.com/site/about.bml) as follows:

LiveJournal Authors Control:

- Who can see each individual entry
- Who can see your contact information
- Who can text message you
- Who can leave comments, and whose comments are screened pending your approval

In the open-ended responses it became clear that access control is of major importance to bloggers, even if most of them do not have tools available to manage access in the way they would like to. The following examples illustrate the situations bloggers have faced due to lack of control features:

Creating multiple blogs for multiple audiences:

> Well, my boyfriend found and read a blog I created specifically for the purpose of having an open forum to write in that he didn't know about. If your best friend is your boyfriend, who are you going to talk about your boyfriend with? But, he found it, and it probably helped push us down the road to a breakup [...].
> - a female blogger from West Virginia

> I am always aware that what I write other people read. Currently I have an ex-boyfriend who stalks my blogs. Although I don't write particularly personal things, I find myself constantly upset that he would find out these details of my life, even though I don't mind strangers knowing them. He is a virtual peeping Tom and I'm now constantly questioning the kinds of things I write about. And I thought I was careful before!
> - a female blogger from California

Safety through obscurity:

> With work-related entries, I'll sometimes use only a first initial [to identify people]. I'm not overly concerned since the people I write about don't know of my blog's existence and aren't particularly net-savvy.
> - a female blogger from Canada

Audience

Respondents in this survey said they learn who their audience is mostly in three ways: by reading the comments that people leave on their sites, by looking at the access logs (logs of all the IP addresses that have accessed the blog site), and by checking trackbacks and referrer links. Interestingly, there was no correlation between feeling like one "knows" one's audience and the amount of time one has been blogging. This means that even authors who have kept blogs for a long time do not necessarily feel they know their audience well.

Strategies for Discovering One's Audience

Looking at access logs. Almost half of the respondents (46%) said they felt they knew their audience very well, whereas 22% felt the opposite way. Interestingly, 47% of the respondents said they checked access logs fairly often. Looking at these numbers, one might conclude that the people who look at their access logs are the same people who responded that they felt they knew their audience fairly well; such is not the case, however. There was no correlation between how often respondents check access logs and how well they feel they know their audience. This is surprising, in that access logs are one of only a few mechanisms that provide bloggers with tangible indications of their audience. This lack of correlation suggests that frequently looking at access logs does not necessarily lead to better knowledge of one's audience.

Reading comments. The fact that, in general, readers can leave comments on entries gives rise to "conversational" exchanges and creates highly interactive communities around blog sites. Not all blog sites allow comments, however. The overwhelming majority of respondents (92%) in this survey said that leaving comments was an option on their blogs. Most respondents whose blogs allow comments to be posted said they look at the comments to gain an awareness of the identity of their audience. As will be explained below, this can be problematic because bloggers sometimes focus solely on the comments to gain an understanding of who makes up their audience.

Checking trackbacks and referrer citation links. Bloggers have the ability to link to each other's entries, creating what is known as a "trackback." For instance, if a blogger sees an entry on someone else's blog that she finds interesting, she might link to it and discuss it on her own blog; by doing so she creates a trackback to the other blogger's entry. Referrer links are used to see who is linking to the referring page, i.e., blog authors can put a link on their main page that automatically enables them to see who links to them. Looking at trackbacks and referrer links is an additional strategy that respondents in this survey use for gaining a better sense of their audience:

> I check my referrer logs every day, which helps me keep a handle on where people are
> coming from.
> - a male blogger from Florida

> I use a statistics package and a referrer log, and I also enable comments. I don't know
> all of my audience, and I make no effort to track IP addresses, but a fair proportion of
> my readers choose to make themselves known either by linking or commenting.
> - a female blogger from Massachusetts

It should be noted that authors who keep LiveJournal blogs make up a special group in this case. Because they have the ability to assign specific audiences to each entry, they have a much better idea of who their readers are. This is described in the following quotations from LiveJournal authors:

> I restrict access to my journal. I have designated a "friends" list, and those are the only
> people who can read my entries. I have about 50 designated friends, most of whom I
> know in real life; there are only 3 I haven't met. These people I know either from
> message boards, or from friends of friends. Otherwise I know everyone pretty well. I'm
> very careful about who can read my journal.
> - a female blogger from Utah

> LiveJournal has an option to limit access to certain entries; thus, for the more private
> entries that I have access-controlled in this fashion, I know the audience extremely well.
> - a male blogger from California

Perceptions of One's Audience

There is a considerable gap between reading the "traces" left by readers on one's blog and feeling that one knows one's audience. Nevertheless, most bloggers must rely on limited indicators of past actions (access logs, comments, and trackbacks) in order to form a mental picture of who is reading their posts. This paucity of clues indicating identity and presence can cause distorted views of readership to emerge. For one thing, bloggers may begin to perceive the people whose presence is more tangibly obvious (e.g., commenters) as their entire audience.

Core versus periphery. Respondents' open-ended essay answers reveal that a significant portion of the people who replied that they felt they knew their audience "fairly well" actually meant they knew their "core" audience well, that is, they knew those few people who are frequent readers and who, in many

cases, leave comments on their blogs. They also noted that, for the most part, they did not know who the rest (i.e., most) of their audiences were. Although respondents were well aware that their "core" audience comprised a minor segment of their entire readership, they tended to think about their entire audience in terms of this small group of people. This has significant implications for privacy in the sense that, once people start thinking about a small part of their readership as the whole of it, they may customize their postings for that particular group of people. For instance, if all the comments a blogger gets on his site come from close friends, he might forget that his actual readership is broader and might start blogging about things that he would only talk about with close friends.

The quotations below illustrate the perception of core versus periphery; paradoxically, these quotes come from people who responded that they feel they know their audience fairly well:

> Some [readers] I know quite well and correspond with but that's just a handful. I've got hundreds of readers who I don't know at all.
> - a male blogger from Georgia

> Many of my readers leave comments, and I have quite a few readers whose websites I read myself. Others are family or friends. I bet, though, that at least two-thirds of my readers are people I've never heard of.
> - a female blogger from Texas

> I have a base readership, some of whom I know personally; but the majority of my site visitors come via Google.
> - a male blogger from Arizona

> About half of the readers are people I know very well, and about half of the readers are complete strangers coming in off the net.
> - a male blogger from California

> For the most part, I have no idea. There is a minority that are personal friends or acquaintances, but for the rest, I can only go by the small percentage that post comments, leaving a majority of anonymous readers I know nothing about.
> - a male blogger from Colorado.

Performing for an audience. In the present survey, various respondents expressed the pressure that comes with having an audience. Even in its current elusive form, a blog's audience can become such a powerful element of the interaction that it can affect what the blogger chooses to write about. One in five respondents (20%) said that the comments people leave on their blogs affect the entries they write. For that reason, audience interests and reactions can help shape (or change) what a blog is all about.

> My friends from offline, mostly non-technical, read my blog out of personal interest (or to see the photos), and ignore the technical stuff. But the vast majority of strangers finding my blog through Google do so looking for my Movable Type tutorials or tech commentary. It's an interesting division, which keeps me writing for both halves.

> - a male blogger from California

> My audience has been with me for over two years and I'm continually getting emails and comments from people I've grown to know by screen name, as well as in person. I know what they like to read about, what they don't, and I tailor my site to match while still giving me the freedom to rant about my world.
> - a female blogger from Washington

Content-based, self-selective audiences. Sometimes authors perceive their audience as being comprised solely of people with similar interests. They assume that, because they focus their writings on a certain topic (or a certain group of topics), their readers must be people who are interested in those topics.

> I assume they're an audience of fellow web designers/bloggers, plus a few personal friends, but I don't have any real evidence.
> - a male blogger from Virginia

> It's a self-selecting group, people who are interested in the same things I am interested in.
> - a female blogger from the U.S. (unspecified region)

> I write mainly from a faith-based perspective about faith issues and living out a Christian life, so most (if not all) of my readers are Christians themselves, on the same faith journey.
> - a female blogger from Malaysia

Persistence

The vast majority (84%) of respondents in this survey said that they were not bothered by the fact that their blog entries could be available online for a long time. This attitude is somewhat surprising considering that 36% of respondents said they had gotten in trouble for something they had posted in a blog entry. Based on the guidelines identified above for revealing names of companies, it seems that respondents are highly mindful of their current, past, and potential employers (companies where they are currently being interviewed). Respondents did not, however, express concern about how their persistent postings might affect their employment chances in the long-term future.

When asked whether they had erased or edited content they had published on their blogs, 75% of respondents said yes. In the essay portion of this answer, respondents explained that they usually edited typos and grammatical errors. Nevertheless, 35% of respondents said they had edited for content as well: entries they later decided were too personal or "mean." Some respondents also mentioned having revisited entries to obfuscate peoples' names. The following responses illustrate these editing situations:

> I have written some things politically that I have subsequently removed for fear of how it might be interpreted in the future. But with Google's caching I'm not so certain my methods will be that effective.
> - a male blogger from Nevada

> I often edit old entries, mainly because I don't want to be held accountable for something mentioned in them.
> - a female blogger from Washington D.C.

No correlation was found between how often authors disclosed highly personal information and how concerned they were about persistence. Given that respondents who write more personal materials tend to get in trouble more often, this is an unexpected finding. Additionally, there was no correlation between the amount of time a person has been blogging and how concerned they are with the persistence of their entries.

When asked how liable they thought they were for materials published on their blogs, 58% of respondents said they thought they were highly liable. Only 10% of respondents did not believe they were liable at all. Most respondents (75%) felt that others could not sue them for what they had written on their blogs.

Discussion

In the survey presented here, respondents were asked to self-report on a number of attitudes, habits, and behaviors in relationship blogging and privacy. Unlike other studies where researchers accessed participants' blogs and conducted content analysis of posts (Herring, Scheidt, et al., 2004; Nardi et al., 2004), here it is the authors' subjective sense of privacy and liability that is revealed. This self-disclosure approach has two important implications: (1) There can be disparities between stated privacy attitudes and actions, and (2) Participants' perceptions of their blogs might differ from those of outside observers and researchers.

It is well documented that people's perceptions of their own behavior can differ from how they actually behave (Whyte, 1990). This phenomenon may help explain why, for instance, respondents in the present study reported being comfortable with the persistence of their posts while at the same time admitting to having edited the content of entries on their sites. In addition, because of the self-reporting nature of this study, accuracy is difficult to verify. For example, no external validation was conducted on the sites of participants who described their blogs as being mostly "personal musings." It is possible that some of the "personal musings" sites reported here might fit the "filter blog" definition found in other blog studies (Herring et al., 2004), or vice versa. Therefore, comparisons between the present and previous findings should be made with these caveats in mind.

It is also important to keep in mind that the respondent sample in this study may not be representative of the entire population of bloggers. As mentioned above, participants in this survey were not randomly selected. The demographic profile of respondents here differs from what we know from studies and popular blogging sites in two main attributes: age and gender. On popular blogging engine sites, such as LiveJournal, where usage statistics are available, teenagers account for the majority of the blogger

population whereas most of the respondents in this survey were between 21 and 40 years old. Moreover, previous blog studies have indicated that women maintain as many blogs as do men (Herring, Kouper, et al., 2004), yet males accounted for 63% of the respondent sample here. The results from this survey might thus have been different had the pool of respondents been randomly sampled.

Keeping in mind the potential biases introduced by the self-report methodology and the respondent sample in this survey, it is nonetheless interesting to note parallels between the findings presented here and those of previous studies. Most respondents in this study described their blogs as the personal musings/journal type. This trend echoes those found in other blog studies, which reported that personal journal blogs comprise the majority of sites (Herring, Scheidt, et al., 2004; Nardi et al., 2004). Moreover, participants in this study described tactics for keeping certain information private even when it is publicly published—such as using friends' initials instead of their full names—that supports an earlier study describing how authors use "common ground" to manage privacy (Gumbrecht, 2004). Authors in this study were willing to publish the names of friends and companies/products when they deemed the contents of an entry to be either positive or innocuous. This tendency mirrors results from a previous study showing that people are inclined to value privacy less the more socially desirable a piece of information is (Huberman, Adar, & Fine, 2004).

Despite the emerging privacy strategies described in this study, authors reported having difficulty negotiating privacy boundaries under certain circumstances. The workplace is one setting where such problematic situations regularly occur. Bloggers' privacy boundaries in the workplace have yet not been clearly established, either socially or legally. As the quotations in the previous sections illustrate, this is one area of conflict that greatly affects bloggers, at times resulting in authors being fired from their jobs. It is likely that, while disagreement over what constitutes acceptable blogging material persists, bloggers will continue to be reprimanded and employers will continue to be frustrated by their employees' blogging activities.

One recommendation that emerges from the findings of this study is that companies should share the responsibility of articulating what constitutes acceptable blogging behavior and what they see as problematic practices. Some companies have already started to enunciate blogging guidelines for employees. A few companies have posted written policies concerning personal blogs on their Web sites, including clear, point-by-point suggestions addressing issues that are sensitive to the company but that may not occur to employee bloggers when they choose to discuss matters related to the company's technology or business.[8] Such policies could serve as the first step in a broader process of negotiation between employers and employees as blogging practices continue to evolve.

More recently, a new type of blog has emerged: the company-wide blog. Examples of this type include the Google Blog (http://www.google.com/googleblog/) and Microsoft's Channel 9 (http://channel9.msdn.com/). On company-wide blogs various employees post entries, ranging from

company-related news to interesting information about the field in which the company operates. Because this kind of blog is patently associated with a company's image, concerns about what is fit to publish are naturally greater. So far, however, neither Google nor Microsoft has published any company-wide blogging policy (Foley, 2003). It remains to be seen whether these sites will become popular, opinionated blogs, or whether they will conform to their companies' PR protocols.

Summary and Conclusion

With few exceptions, most blogs exist in a fully public arena, which means that, once published, entries are readable by anyone on the Web. The results from this survey suggest that bloggers are starting to come up against a range of privacy-related issues varying from minor embarrassments with family and friends to termination of their employment. The findings also indicate that many respondents are developing strategies for minimizing potential problems with others when posting their entries online.

In summary, the findings suggest that asking permission of others before blogging about them tends not to be the rule. Moreover, even though authors thrive on the interactive nature of feedback from readers, bloggers do not have a good idea of who their entire audience is. Authors count on activity traces—IP addresses and trackbacks—to piece together an imaginary map of their readers. However, this study found no correlation between checking one's access log files and feeling like one knows one's audience well. The relative lack of identity and presence cues can result in imprecise assessments of readership that have serious implications for privacy.

A positive correlation was found between the frequency with which respondents reported publishing highly personal materials and how often they had faced problems because of their postings. A significant portion (36%) of respondents stated having gotten in trouble because of things they had written on their blogs, and 34% of respondents knew other bloggers who had faced negative reactions from family and friends. When asked about the persistent nature of their entries, 75% of respondents said they had edited previously published postings, with 35% of respondents having edited content (as opposed to merely correcting typos). The fact that the overwhelming majority of respondents are not concerned about the persistent nature of their personal postings (nor by the fact that the contents are cacheable by search engines such as Google) is troubling. Even though authors in this study know they are liable for the materials they write online, it seems that most have not yet realized the potential for misuse of information in this data-persistent, heavily searchable environment.

More encouragingly, the findings also suggest that blogging is a world in flux where social norms, most of them self-imposed, are starting to emerge. When confronted with questions of defamation and legal liability, however, respondents painted a conflicting picture. In general, they believed that they were liable for what they published online. At the same time, authors in this study did not believe others could sue them for what they had written on their blogs.

Taken as a whole, these results point to three general areas in need of improvement: clearly articulated social norms for blogging in sensitive environments such as the workplace; more sophisticated access control over how blog entries are presented, when, and to whom; and more prominent markers of the presence of readers on blog sites so that authors are constantly reminded of the full scope of their audience.

As they continue to permeate society in their many forms, weblogs are bound to raise various privacy and liability issues. The results presented here are simply the tip of the iceberg, an initial snapshot of how a segment of the blogging community deals with issues of privacy today. More research is needed in order to better understand the privacy concerns inherent in the practice of blogging and how to address them. It would be useful, for example, to examine whether there are different expectations of privacy and liability among the various blogging communities—political, technological, scholarly, family, and teenage bloggers. Studies taking a longitudinal approach to privacy and how it is experienced by authors as they continue to blog over time would also be an important contribution to both weblog and online privacy research.

Notes

1. A recent national survey by the Pew Research Center (Pew, 2004) reports that 2% of Internet users in the survey maintain web diaries or weblogs. An earlier survey indicates that, even though only a small number of users are writing blogs, a slightly larger number of Internet users are visiting them: 11% of Internet users report visiting blogs written by others (Pew, 2002). Blog readers most frequently visit the online diaries of friends (56% have done so), strangers (46% have visited the blogs of people they have never met) and family members (a quarter report visiting family blogs).

2. Admittedly, bloggers have the option of going back to published entries and editing their content. The survey results indicate that some authors take advantage of these editing capabilities.

3. Knowledge blogs, or k-logs, are usually theme driven. Researchers, for instance, may keep knowledge blogs where they make observations and record relevant references about a particular topic or field.

4. Nussbaum (2004), an exception to this trend, focuses on how teenagers use blogs.

5. This technology is already being embraced by bloggers who feel they have things to say in public but who, because of the nature of their posts, cannot afford to be identified. In July of 2003, the *Village Voice* ran an article entitled "The Sharer of Secrets" (O'Shea, 2003), on a blogger who publishes under the name Hasidic Rebel and whose comments range from musings about the Hasidic lifestyle to stinging indictments of the community to which he belongs. The blog has angered Orthodox Jews, many of whom view the writings as a betrayal. At the same time, the blog has also sparked an outpouring of affirmations and brought other doubting Jews to the surface.

6. The reported bad experiences include one blogger who found out that his boss was using his

LiveJournal to keep tabs on where he was and what he was working on, and an author who received bothersome emails and instant messages because of her LiveJournal entries.

7. LiveJournal and DeadJournal are prime examples of popular blog services that appeal to teenagers. Both services publish statistics on the age distribution of their subscribers: http://www.livejournal.com/stats.bml, http://www.deadjournal.com/stats.bml

8. Groove Networks, a company founded by Ray Ozzie (the creator of Lotus Notes), has posted such a policy on their corporate Web site. The policy was designed to address four areas of concern: (1) that the public would consider an employee's entries to be official company communications, rather than expressions of personal opinion; (2) that confidential information—the company's own or a third party's—would be inadvertently or intentionally disclosed; (3) that the company, its employees, partners, or customers would be disparaged; and (4) that quiet periods imposed by security laws or other regulations would be violated (Suitt, 2003). Here is an excerpt from their blogging policy:

> "Please make it clear to your readers that the views you express are yours alone and that they do not necessarily reflect the views of the company. To help reduce the potential for confusion, we would appreciate it if you put the following notice—or something similar—in a reasonably prominent place on your site (e.g., at the bottom of your "about me" page):
> *The views expressed on this website/weblog are mine alone and do not necessarily reflect the views of my employer.*
> If you do put a notice on your site, you needn't put it on every page, but please use reasonable efforts to draw attention to it—if at all possible, from the home page of your site."

References

Bandlow, A., & Jensen, P. (2002a). *LiveJournal: Personal thoughts in a public forum.* Unpublished midterm paper for "The Design of Online Communities" class (CS6470). Georgia Institute of Technology.

Bandlow, A., & Jensen, P. (2002b). *LiveJournal: Augmenting the community.* Unpublished final paper for "The Design of Online Communities" class (CS6470). Georgia Institute of Technology.

Blood, R. (2000). *Weblogs: A history and perspective.* Retrieved February 17, 2005 from http://www.rebeccablood.net/essays/weblog_history.html

Bray, H. (2004, January 4). Job blogs hold perils, opportunities. *The Boston Globe,* p. G1.

Dibbell, J. (2002). Portrait of the blogger as a young man. In J. Rodzvilla (Ed.), *We've Got Blog: How Weblogs are Changing Our Culture* (pp. 69-77). Cambridge, MA: Perseus Publishing.

Döring, N. (2002). *Personal home pages on the web: A review of research. Journal of Computer-Mediated Communication, 7* (3). Retrieved January 5, 2005 from http://jcmc.indiana.edu/vol7/issue3/doering.html

Farivar, C. (2003, August 27). UC-Berkeley students find a new voice in blogs. *Daily Californian,* via U-Wire.

Foley, M. J. (2003, June 16). Microsoft blog policy coming down the pike? *Microsoft Watch.* Retrieved January 5, 2005 from http://www.microsoft-watch.com/article2/0,4248,1128705,00.asp

Glenn, D. (2003, June 6). Scholars who blog. *The Chronicle of Higher Education, 49* (39), A14.

Grudin, J. (2001). Desituating action: Digital representation of context. *Human-Computer Interaction, 16* (2-3), 269-286.

Gumbrecht, M. (2004). Blogs as "protected space." *WWW2004 Workshop on the Weblogging Ecosystem: Aggregation, Analysis and Dynamics.* Retrieved January 5, 2005 from http://www.blogpulse.com/papers /www2004gumbrecht.pdf

Havilland, P., Cronin, A., & Amon, D. (2004). Retrieved 05 January 2005 from http://www.samizdata.net /blog/glossary_archives/001964.html

Herring, S. C. (2003). Gender and power in online communication. In J. Holmes & M. Meyerhoff (Eds.), *The Handbook of Language and Gender* (pp. 202-228). Oxford: Blackwell Publishers.

Herring, S. C., Kouper, I., Scheidt, L. A., & Wright, E. (2004). Women and children last: The discursive construction of weblogs. In L. Gurak, S. Antonijevic, L. Johnson, C. Ratliff, & J. Reyman (Eds.), *Into the Blogosphere: Rhetoric, Community, and Culture of Weblogs.* University of Minnesota. Retrieved January 5, 2005 from http://blog.lib.umn.edu/blogosphere/women_and_children.html

Herring, S. C., Scheidt, L. A., Bonus, S., & Wright, E. (2004). Bridging the gap: A genre analysis of weblogs. *Proceedings of the 37th Hawai'i International Conference on System Sciences (HICSS-37).* Los Alamitos: IEEE Computer Society Press. Retrieved January 5, 2005 from http://www.blogninja.com /DDGDD04.doc

Herring, S. C., Scheidt, L. A., Bonus, S., & Wright, E. (In press). Weblogs as a bridging genre. *Information, Technology & People.*

Huberman, B. A., Adar, E., & Fine, L. R. (2004). Valuating privacy. *IEEE Security & Privacy Magazine.* Retrieved January 5, 2005 from http://www.hpl.hp.com/research/idl/people/eytan/deviance.pdf

Mortensen, T., & Walker, J. (2002). Blogging thoughts: Personal publication as an online research tool. In T. Morrison (Ed.), *Researching ICTs in Context.* Retrieved January 5, 2005 from http://www.intermedia.uio.no/konferanser/skikt-02/docs/Researching_ICTs_in_context-Ch11-Mortensen-Walker.pdf

Nardi, B., Schiano, D., & Gumbrecht, M. (2004). Blogging as social activity, or, would you let 900 million people read your diary? *Proceedings of Computer Supported Cooperative Work 2004.* Retrieved January 5, 2005 from http://home.comcast.net/%7Ediane.schiano/CSCW04.Blog.pdf

NITLE Blog Census (2004). Retrieved January 5, 2005 from http://www.blogcensus.net/weblog/

Nussbaum, E. (2004, January 11). My so-called blog. *The New York Times, Late Edition, Section 6,* 33.

O'Shea, W. (2003, July 22). The sharer of secrets. *Village Voice,* p. 55.

Orlwsky, A. (2003, November 6). Pseudonymous blogging safe (for now). *The Register.* Retrieved January 5, 2005 from http://www.theregister.co.uk/content/6/33837.html

Palen, L., & Dourish, P. (2003). Unpacking "privacy" for a networked world. *Proceedings of the ACM CHI.* Ft. Lauderdale, FA. Retrieved January 5, 2005 from http://www.cs.colorado.edu/~palen/Papers/palen-dourish.pdf

Pax, S. (2003, September 9). I became the profane pervert Arab blogger. *The Guardian,* p. 2.

Pew Research Center For the People & Press. (2002, September 5). One year later: September 11th and the Internet. *Pew Internet and American Life Project.* Retrieved January 5, 2005 from http://www.pewinternet.org/pdfs/PIP_9-11_Report.pdf

Pew Research Center For the People & Press. (2004, February 29). Content creation online: 44% of U.S. Internet users have contributed their thoughts and their files to the online world. *Pew Internet and American Life Project.* Retrieved January 5, 2005 from http://www.pewinternet.org/pdfs/PIP_Content_Creation_Report.pdf

Phillips, J. (2003, July 10). In this computer world, be careful what you blog. *The Record,* p. F09.

Rosen, J. (2000). *The Unwanted Gaze: The Destruction of Privacy in America.* Vintage Books: New York.

Sarnataro, B. R. (2003, July 3). Dear digital diary. *Arkansas Democrat-Gazette,* p. 37.

Schiano, D. J., Nardi, B. A., Gumbrecht, M., & Swartz, L. (2004, April). Blogging by the rest of us. *Proceedings CHI 2004.* Vienna.

St. John, W. (2003, July 22). Perilous universe of online blogging. *Chattanooga Times Free Press,* p. E1.

Suitt, H. (2003). A blogger in their midst. *Harvard Business Review, 81* (9), 30-40.

Turnbull, G. (2001). The state of the blog. Interview with Evan Williams. *Writing the Web.* Retrieved January 5, 2005 from http://writetheweb.com/Members/gilest/old/106/view

Turnbull, G. (2004, June 14). The seven-year-old bloggers. *BBC News World Edition.* Retrieved January 5, 2005 from http://news.bbc.co.uk/2/hi/uk_news/magazine/3804773.stm

Walther, J. B. (1996). Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research, 23* (1), 3-43.

Whelman, D. (2003, July 1). In a fog about blogs. *American Demographics,* p. 6.

Whitworth, D. (2003, June 30). The (not so) secret diary of a blogger. *The Times (London),* p. 4.

Whyte, W. H. (1990). *City: Rediscovering the Center.* New York, N.Y.: Anchor. (Original work published 1988.)

Appendix: Online Survey Form

Blog Survey

We are interested in finding out how bloggers think about issues of privacy and liability as they publish online. To that end, we are asking bloggers to fill out this survey. If you are interested in learning the results from our survey, please indicate so at the end of the questionnaire.

If you have any questions about this survey, please send email to: blog-survey@media.mit.edu Thanks for your help!

1. Please list the URLs of up to five blogs that you read. (optional)

2. Please list the URLs of blogs to which you are the sole author. (optional)

3. Please list the URLs of any blogs that you co-author with other people. (optional)

4. Do you identify yourself on your blog (i.e. is your real name clear on the site?)
    a. Yes
    b. No
    c. on some blogs, but not all
    d. it is a bit more complicated. please explain:

5. How long have you been blogging?
    a. less than 3 months
    b. 3 to 6 months
    c. 6 months to one year
    d. one to two years
    e. over two years

6. How would you characterize the kinds of entries you publish on your blog(s)? Please check all that apply:
    a. personal musings
    b. academic brainstorm
    c. professional brainstorm
    d. political opinions
    e. hobby
    f. gossip
    g. other:

7. How private are the things you write about on your blog? ("private" means either personal or confidential)
  extremely 1 2 3 4 5 not at all

8. How often have you considered whether something was too personal to write about in your blog?
  always 1 2 3 4 5 never

9. In general, when you write things about people you know personally in your blog:
  a. do you ask them permission to do so?
    always 1 2 3 4 5 never I never write about people I know personally
  b. do you reveal their names?
    always 1 2 3 4 5 never I never write about people I know personally
    Please explain why:

10. In general, when you write things about companies/products in your blog:
  a. do you ask them permission to do so?
    always 1 2 3 4 5 never I never write about companies/products
  b. do you reveal their names?
    always 1 2 3 4 5 never I never write about companies/products
    Please explain why:

11. How well do you feel you know your blog's audience?
  extremely well 1 2 3 4 5 not at all
  Please explain:

12. How often do you look at the log of who has accessed your blog?
  extremely often 1 2 3 4 5 never logs not available

13. If you were aware of all the people who read your blog, how likely is it that you would become more careful about what you write?
  very likely 1 2 3 4 5 not likely at all N/A

14. How often have you written highly personal things on your blog?
  all the time 1 2 3 4 5 never

15. How often have your gotten in trouble for anything you wrote on your blog?
  all the time 1 2 3 4 5 never
  If you have gotten in trouble, briefly explain the situation:

16. How much do the comments that people write on your blog affect the entries you write?
  extremely 1 2 3 4 5 not at all leaving comments is not an option on my blogs

17. Are you surprised when someone you meet in person says they have read your blog?
  extremely 1 2 3 4 5 not at all it has never happened to me

18. Does it bother you that the things you publish on your blog will be available online for a long time?

    extremely 1 2 3 4 5 not at all

19. How liable do you think you are for the things you write on your blog?

    extremely 1 2 3 4 5 not at all

20. Do you think people could sue you for things you have written on your blog?
    a. Yes
    b. No
       If yes, please explain:

21. Have you ever erased (or edited) something you had published on you blog?
    a. Yes
    b. No
       If yes, please explain:

22. Do you know any bloggers who have gotten in trouble with friends or family for things they posted on their blogs?
    a. Yes
    b. No
       If yes, please explain:

23. Do you know any bloggers who have gotten in trouble with the law (have been sued, or had their entries subpoenaed by lawyers) because of their blogs?
    a. Yes
    b. No
       If yes, please explain:

24. Have your blogging habits changed over time?
    a. Yes
    b. No
       If yes, please explain:

25. Do you do anything to limit who can read what you post?
    a. Yes
    b. No
       If yes, please explain:

26. If you would like to clarify any of your responses or provide more detail, please do so here. Also, feel free to share information about blogging aspects that you think might be relevant to this survey.

DEMOGRAPHICS

27. Age

28. Sex
    a. male
    b. female
    c. prefer not to disclose

29. Education
    a. Elementary school
    b. High school
    c. College / University
    d. Masters / Professional Degree
    e. Ph.D.

30. Ethnicity:
    a. African
    b. African-American
    c. Asian
    d. Asian-American
    e. Australian
    f. Caucasian
    g. Latino/Hispanic
    h. Native American
    i. Pacific Islander
    j. Other

31. Where were you born?

32. Where do you currently reside?

33. Email address (optional)

34. The results from this survey will be available on this site after the collection and analysis of data are done. Are you interested in being notified about the results from this survey?
    a. Yes
    b. No
    If yes, please make sure you enter a valid email address

About the Author

Fernanda B. Viégas is a doctoral student in the Media Laboratory at the Massachusetts Institute of Technology. Her research focuses on the visualization of the traces people leave as they interact online. Some of her projects explore email archives, newsgroup conversations, and the editing history of wiki pages. She is the creator of the award-winning Chat Circles program, an abstract graphical interface for communicating online. Her interest in issues of online privacy stems from her work with large archives of online social data. She received her master's degree in Media Arts and Sciences from MIT, and her

bachelor's degree in Graphic Design and Art History from the University of Kansas. More information is available at http://web.media.mit.edu/~fviegas/

Address: MIT Media Lab, E15-391, 20 Ames Street, Cambridge, MA 02142 USA

© 2005 Journal of Computer-Mediated Communication