

By Seth P. Berman, Lam D. Nguyen & Julie S. Chrzan

## Web 2.0: What's Evidence Between "Friends"?

If you are over thirty, you may well have never visited Facebook, MySpace, or Bebo, and the idea of "Twittering," or posting mundane aspects of your life in short and incessant online posts may seem odd, but such social networking sites attract millions of users. They are no longer the sole domain of technologically savvy teenagers. Such Web 2.0 applications host and post untold amounts of personal and professional information about their users, and often serve as the primary means of recorded communication amongst their users. While Web 2.0 pioneers see a fun way to connect to their peers, lawyers are increasingly finding a rich new source of potentially discoverable information. In short, where people go, litigation must follow.

### What is Web 2.0?

So what exactly is "Web 2.0"? It may sound like a new version of the World Wide Web, but it doesn't actually refer to any specific new technology. In fact, Web 2.0 doesn't focus on the technology at all, but rather the participatory nature of how a website's content is created and delivered. Web 2.0 applications offer a framework in which its users interact with and even create content themselves, rapidly sharing the information across users and encouraging other users to respond in kind.

This open and dynamic exchange of information radically reduces the level of centralized control that a website owner has over the user experience. It is this very concept that has garnered millions of users rushing to these websites every day. The easiest way to understand this is by contrast to Web 1.0 content. Web 1.0 content is produced, edited, and maintained by the website creator. Thus, Yahoo.com, for example, creates or provides the content available on its web pages. Though Yahoo users had the ability to navigate through that content, they generally lacked the ability to alter it. Web 1.0 is interactive much in the way a television is interactive – the user can change channels, finding content that he or she likes, but the user can not do much to change, comment on, or affect the content he or she finds there.

Web 2.0 applications are designed to give far greater editorial control to users. A simple example of Web 2.0 is a web-based "wiki," the most famous of which is Wikipedia. In wikis, the content is created by users and is dynamic – users are empowered to undo and redo each other's work. Thus, unlike traditional encyclopedias, whose contents are created by a team of experts, Wikipedia entries can be created or modified by any user. Accordingly, if you look up "Boston" in an encyclopedia, you will learn what facts a team of experts thinks is important about the City of Boston. If you look up "Boston" in Wikipedia, you would learn of what an unknown number of individuals – probably many of whom live in or have some interest in Boston – think is relevant and interesting about the city. One significant advantage of the Wikipedia model is that far more data can be compiled through a loose collaboration than could ever be done by a centralized process. Thus, there is a Wikipedia entry for the Boston Bar Association, providing details of the organization, its history and mission, and even a photograph of the historical building it occupies.

Another way in which individuals utilize the interactive nature of Web 2.0 applications is by blogging and commenting on blogs. A surprising number of people publish blogs on every conceivable topic, often revealing quite intimate details of their lives, thoughts, or actions. Though blogs are often written under pseudonyms, some bloggers are quite open about who they are. The distinction could prove important in litigation if the writer's identity is relevant to the issues being litigated. The same analysis applies to people who routinely comment on blogs or in response to articles, even if they themselves do not have a blog.

Perhaps the most talked about Web 2.0 applications are the social networking sites alluded to earlier. Users of Facebook, MySpace and other such sites post comments, pictures, and details about their interests and activities. The contents of these posts can then be shared with "friends," a term that refers to other users who are allowed to access a user's page. Some Facebook users have tens of thousands of



Seth P. Berman serves as Managing Director and Deputy General Counsel of Stroz Friedberg's Boston office.



Lam D. Nguyen is the Director of Stroz Friedberg's Digital Forensics lab in Boston.



Julie S. Chrzan is a Vice President in the Boston office of Stroz Friedberg.

such “friends.” Additionally, it is possible to track the “friends” of your “friends,” thereby indirectly linking hundreds of thousands of individuals through small degrees of separation. Oftentimes, users are able to see some of the data on pages of the friends’ friends, meaning that information on such social networking sites, though at first appearing closed to all but permitted users, often becomes quite widely available. As the user base of these sites has expanded, the sites are no longer limited to talk of college life and twenty-something parties. There is an ever growing array of networking sites that target professionals, such as LinkedIn and Plaxo. All of these sites can have interesting, relevant, embarrassing or just plain titillating information about millions of people, much of which might well be relevant to ongoing litigation.

Another aspect that Web 2.0 sites tend to have in common is that they promote very informal means of communication. As was true for the early days of email (and to a large extent is still true even for emails), electronic communication has a spontaneity that makes it seem impermanent. Thus, people write things in electronic communication that tend to mimic casual spoken conversation rather than formal, written communication. It is for this reason that discovery often unveils truly damning email communications, introduced with statements like “Don’t tell anyone about this, but...” This informality is even truer of Web 2.0 applications. Blogs and journal entries on social networking sites are often stream-of-consciousness statements. The perceived anonymity frequently produces much harsher condemnation in response to blogs and articles than is typically acceptable in face-to-face communication. In short, Web 2.0 applications may record people’s thought processes and impressions in unguarded moments, exactly the sort of evidence that can be invaluable during litigation.

### **Potential Value of Web 2.0 Content for Lawyers**

The vast amount of user-created content of Web 2.0 applications is a growing resource for lawyers. Many employers routinely conduct web searches on their potential employees as part of background checks. This often turns up the sort of information that more formal background checks would likely miss. For example, an employer might find that an applicant whose criminal record checks and credit history are pristine, may nevertheless be hiding something important in his or her background – such as membership in a neo-Nazi organization. Similarly, in a recent criminal case, pictures on MySpace were credited with being a deciding factor in the sentencing of a then 22-year-old student who received five years and four months in prison after she drove under the influence of alcohol and got into an accident that killed her passenger. Despite a recommendation for probation, the judge decided in favor of a prison sentence after reviewing pictures from the defendant’s MySpace page that showed her drinking with friends in the months following the car crash.<sup>1</sup>

These kinds of embarrassing revelations are not limited to people in their twenties. Though it ultimately led to nothing

more than a few months of bad publicity, one Web 2.0 application led to an SEC investigation of a prominent CEO. In 2007, the CEO of a popular grocery chain was investigated by the SEC when it became known that he had posted thousands of messages on Yahoo Finance Board promoting himself and his organization (and denigrating competitors and their CEOs) using a pseudonym.

### **Limiting Discovery of Web 2.0 Information?**

Electronically Stored Information (“ESI”) decisions in other contexts have established that courts will not hesitate to sanction litigants who fail to turn over ESI because it has been accidentally destroyed or otherwise overlooked.<sup>2</sup> On the other hand, courts want to tamp down on unrestricted “fishing expeditions” amongst the vast expanses of electronically stored information. For example, the defendant in a recent sexual harassment suit sought to compel production of emails from two MySpace accounts, alleging that messages from those accounts might contain evidence that the plaintiff engaged in consensual sexually-related email communications with other persons, of the same type she characterized as sexual harassment in her complaint. In rejecting this argument, the court stated that although the evidence did establish that the two MySpace accounts were indeed the plaintiff’s accounts, the defendant had no information concerning the identities of the persons with whom the plaintiff was supposedly exchanging email messages, or the subject matter or content of those messages.<sup>3</sup> However, those seeking to limit discovery of social networking site information shouldn’t be overly heartened by this decision, the strong implication of which is that if the defendant had had more evidence that the contents of the websites might have been relevant to the issues, the court likely would have enforced the subpoena.

### **Collecting and Introducing Web 2.0 Content in Evidence**

Once issues such as ownership and scope of the data have been sorted out, attorneys must determine how to introduce evidence derived from Web 2.0 content in court. Although the end result will be different, the steps to be followed for introducing electronic evidence are no different from those for dealing with paper. The correct source of the data must be determined, then the data must be preserved in a defensible manner, authenticated, and deemed relevant to the particular case at bar.

Web 2.0 data, like all data on the World Wide Web, typically can be found in two places. The first is on the server of the website where the data was posted. Thus, one place to obtain data posted on Facebook would be from Facebook itself. Another often overlooked source may be the computer of the user who accessed the website. Indeed, these two sources can work together, with one source augmenting or authenticating the information obtained from the other.

Because the web is dynamic, a website’s data cannot be

relied upon for its longevity. Data from websites routinely are altered or deleted altogether. In order properly to preserve a webpage, it must be captured at the exact moment in time in which the preservation is desired. It may also require multiple captures to produce a series of “snapshots,” showing how a site changed over time. This can be especially useful in instances where the very fact of habitually updating a website, or the dates and times of the updates, are of evidentiary significance. Once content on a social networking site, blog or wiki has been identified as relevant, it must be preserved in such a way that it can later be authenticated for use as evidence in litigation, as well as to show that a complete preservation was performed. Preservation must be done completely to ensure that all website information – including that which might be contradictory to the stance of a party – is captured.

Web-based evidence can be collected in a variety of ways. With text-based web pages, this can be as easy as pressing the print button and producing the printed pages as “best evidence.” For the most part, a fact witness can authenticate the evidence by explaining that she went to the Internet, found the website, and printed what she found. The printed pages can then be entered in evidence without requiring technical expert testimony to introduce complex web concepts.

Of course, websites are rarely restricted to simple text-based pages. Web 2.0 sites rely on rich multimedia content comprised of pictures, video and even sound. This data, especially when user-created, can be a treasure trove of evidence. The preservation of this data, however, is far from simple. It often requires significant technical expertise to faithfully reproduce the information as it would have appeared on the World Wide Web. Will it be sufficient to print stills from a video, or will it be necessary to show the video in its entirety? Can you produce snippets of text from a web posting, or is it only pertinent when shown within the context of the entire website? Can a posted audio file be transcribed, or do you need to introduce it as aural content? From opposing counsel to judges and jury, this data will have a wide audience. It will not only need to be preserved in its original form, but also portable enough to survive the discovery process and to ensure that the relevant parties will be able to review the materials deemed significant.

Additionally, website content such as posted images, documents, and other file attachments may also contain metadata that is crucial in an evidentiary capacity. For example, consider a situation in which an image of illegal activity is found on a group of users’ MySpace pages. It may be necessary to determine which of the users originally took the photo. Forensic analysis can reveal metadata in a file that may show details such as the make and model of the camera, the date the picture was taken, or other relevant information, provided that this easily lost data is promptly preserved.

Ultimately, the best evidence will be ineffective if a correlation cannot be shown between the website and the investigative target. Consider a case where an individual is disclosing non-public financial data to a public website. How should a lawyer connect these internet postings to a particular individual or company? While the website may keep a record of the

screen name used to make the post, screen names rarely reflect anything more than an anonymous designation. A forensic expert may be able to establish the necessary link between an online posting and a real world individual.

To understand this, it is necessary to understand IP (Internet Protocol) addresses. While a thorough discussion of IP addresses is beyond the scope of this article, it is enough to know that every computer connected to the Internet is designated with a unique numeric address. This is very similar to the phone numbering system. An internet service provider must assign an IP address to a computer before it can access the Internet. Computers require an IP address to determine how to route and transmit data across the World Wide Web. The IP address thus becomes a unique identifier that can pierce the seeming anonymity of the Internet. Knowing how to look for this IP address and knowing whom to subpoena for this information is crucial to authenticate data found on the web. Indeed, attorneys should consider having his or her expert assist in drafting the subpoena to the website to ensure that the necessary information is requested, such as the IP address of the individual who posted the information. An analysis of that website’s response will determine who the IP address service provider is, and allow the issuance of a subsequent subpoena to determine the specific computer assigned that IP address at the relevant time.

Depending on the factual circumstances of the case, identifying the computer used to conduct the relevant activity may only be the beginning of the investigation. Assuming that one obtains the appropriate legal authority to search the computer, the question then becomes whether it will still have any useful information. The answer is almost always yes. Since a user unwittingly leaves an evidentiary trail on her computer simply by using it, her computer will provide evidence of her web usage. The user’s activity log, time-date stamps of relevant documents, and even deleted files may be able to augment or corroborate evidence located on the website server. What can ultimately be found on a computer is wholly dependent on how often the computer was used, what programs were installed, and how the user interacts with the computer. Computer data is volatile; therefore a forensic exam is far more likely to reveal what a user did on his computer yesterday than what was done three months ago.

Web 2.0 technologies provide fertile grounds for learning more about parties, witnesses and others involved in litigation. Lawyers and judges must know how to find, capture, understand, and utilize this rich new source of potentially discoverable data. ■

(Endnotes)

1 [www.dailynexus.com](http://www.dailynexus.com); Published February 1, 2007.

2 See, e.g., *Zubulake v. UBS Warburg LLC*, No. 02 Civ. 1243 (SAS), 2004 U.S. Dist. LEXIS 13574 (S.D.N.Y. July 20, 2004); *Zubulake v. UBS Warburg LLC*, 217 F.R.D. 309 (S.D.N.Y. 2003); *Zubulake v. UBS Warburg LLC*, 216 F.R.D. 280 (S.D.N.Y. 2003).

3 *Mackelprang v. Fidelity Nat’l Title Agency of Nevada, Inc.*, 2007 U.S. Dist. LEXIS 2379 (D.Nev. Jan. 9, 2007).

This article has been reprinted/used with permission from the *Boston Bar Journal*, a publication of the Boston Bar Association.