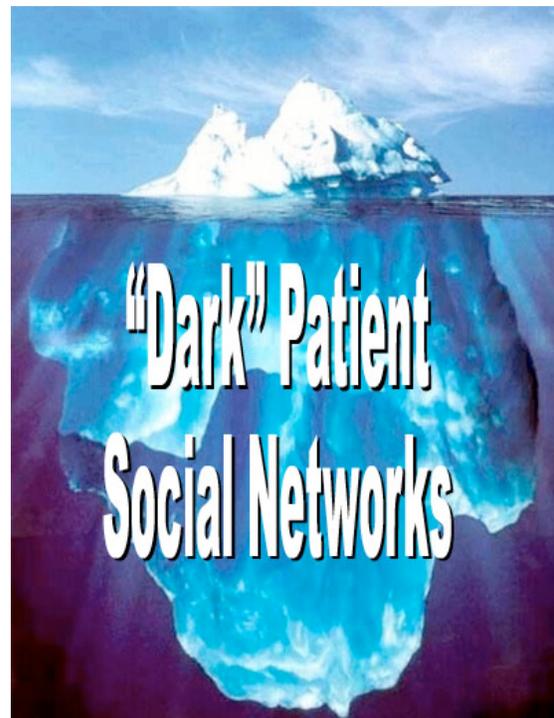## Article

Data Mining in the Deep, Dark Social Networks of Patients

Advice to Pharma: *Caveat Emptor*
Author: John Mack

PMN95-03

Find resources cited this article online at:
http://tinyurl.com/5hgxra

This article is part of the May 2010 issue of *Pharma Marketing News*.

For other articles in this issue, see:
http://www.news.pharma-mkting.com/PMNissueMay10archive.htm

At 1:50 PM on May 20, 2010, I posted the following tweet to my Twitter account:

```
PatientsLikeMe blocks "scraper."
Is this a trend in SM? Pharma can't
mine patient msgs & learn? See
http://bit.ly/9lWSIT #fdasm #hcsm
```

A few minutes later, I received an email message from Jamie Heywood, Co-founder of PatientsLikeMe (PLM), who said "Let's connect." A few minutes later, we were on the phone discussing the "richer story" beneath the surface. Before getting to that, here's how Jamie's partner, Ben Heywood, described what was going on viz-a-viz the "scraper" incident:

"Recently, we suspended a user who registered as a patient in the Mood community. This user was not a patient, but rather a computer program that scrapes (i.e. reads and stores) forum information. Our system, which alerts us when an account has looked at too many posts or too many patient profiles within a specified time interval, detected the user. We have verified the account was linked to a major media monitoring company, and we have since sent a cease and desist letter to its executives.

"While this was not a security breach, it was a clear violation of our User Agreement (which expressly forbids this type of activity) and, more significantly, a violation of the community's trust. Your Account Information (e.g. your names and emails) was NOT in danger of being stolen. It is likely that the forum information that was 'scraped' would be sold as part of that company's Internet monitoring product. In fact, we sell a similar service, PatientsLikeMe-Listen™, to our clients so they better understand the voice of the patient" (see "Transparency, Openness and Privacy"; http://bit.ly/dqAv3p).

Patient Information for Sale

A major issue for PLM is that the media monitoring company—probably employed by an unnamed pharmaceutical company—was not an authentic patient and violated PLM's User Agreement, which states "You may not use any robot, spider, scraper, or other automated means to access the Site or content or services provided on the Site for any purposes." Since PLM is using its own "scraper" software to troll its closed communities to create reports for pharma clients (eg, see "PatientsLikeMe Reports High Rate of Adverse Event Reporting Among Its Members"; box, page 4), it has a vested interest in preventing rouge pharma companies from hiring "scraper" agents to mine the PLM site for the same data it is selling its own pharma clients.

The Dark Depths of "Closed" Patient Communities

"Closed communities" like the member areas of PLM are "dark" to search engine spiders, which are forbidden to troll these areas of PLM and index the content. There are technical ways of doing this, but most search engines agree to abide by requests that certain domains not be indexed.

This is something known as the Dark Net or Deep Web. According to wikipedia, "searching on the Internet today can be compared to dragging a net across the surface of the ocean; a great deal may be caught in the net, but there is a wealth of information that is deep and therefore missed. Most of the Web's information is buried far down on dynamically generated sites, and standard search engines do not find it. Traditional search engines cannot 'see' or retrieve content in the deep Web -- those pages do not exist until they are created dynamically as the result of a specific search. The deep Web is several orders of magnitude larger than the surface Web."

Most pages in forums like those on PLM are dynamically generated. However, it is possible to flip a "switch" that allows the page to be indexed (don't ask me how this works; I only know it's possible because I use such a switch on the Pharma Marketing Network Discussion Forums).

What this all means is that, potentially, much of the best patient-generated information found on social networks is "dark" to pharma companies unless the owners of these communities flip that "switch" or allow pharma marketers access (paid or otherwise). Hopefully, however, these sites might also perform a FREE public (see "If Patients Know Best, then Patient Social Networks Can Help Capture and Report AEs"; http://bit.ly/brUOI0).

PLM community members, BTW, seem to be OK with the sale of their information to third parties: "I can live with you selling the information as long as you continue to reinvest in the infrastructure of the site and keep it more than just up to date," said one person in a comment to Ben's post. "Cutting edge is what I have found here and I expect you will still provide this."

How does PatientsLikeMe make money? It's FAQ (see http://bit.ly/bRQrJg) about that states:

> "We take the information patients share about their experience with the disease, and sell it in a de-identified, aggregated and individual format to our partners (i.e., companies that are developing or selling products to patients). These products may include drugs, devices, equipment, insurance, and medical services. We do not rent, sell or share personally identifiable information for marketing purposes or without explicit consent. Because we believe in transparency, we tell our members exactly what we do and do not do with their data.

> "By selling this data and engaging our partners in conversations about patient needs, we're helping them better understand the real world medical value of their products so they can improve them. We are also helping companies accelerate the development of new solutions for patients. Our end goal is improved patient care and quality of life."

Pharma Trolls Social Media

Searching online patient social networks for key words and discussions about products is something pharmaceutical companies have been doing for quite some time. A recent Pharma Marketing Blog post described one way this is done.

The post described how a BzzAgent agent (aka "operative") was trolling social networks on behalf of Johnson and Johnson (see http://bit.ly/uZfWB). BzzAgent is a leading word-of-mouth (WOM) marketing company that solicits consumers to be their agents in exchange for free samples and other considerations. One "operative" admitted that he took a survey through BzzAgent for Johnson & Johnson, "which basically was more of a 'contract' where if chosen, I agreed to notify J&J if I became aware of any negative talk about their products."

BzzAgent is not engaged in the kind of "scraping" that PLM objects to. First of all, BzzAgent operatives may be legitimate patients. Secondly, monitoring is done manually by a real person. This may be a neat way of getting around User Agreements of "closed" patient communities like PLM's, but it is not nearly as

efficient or effective as using software tools that are the staple of data mining firms. But, if J&J enlisted the help of thousands of real people to monitor patient communities, it could be pretty effective.

By the way, BzzAgent operatives monitoring PLM communities may still be in violation PLM's User Agreement, which forbids individuals from using information collected manually or otherwise on PLM "in connection with any commercial endeavors." This has repercussions for drug companies that indirectly hire such agents, as discussed below.

Transparency, Openness and Privacy

This whole "scraping" incident raises the issue of "Transparency, Openness and Privacy." Anyone can join PLM and claim to be a patient—we are ALL patients at one time or another. Therefore, when I joined, I was not violating PLM's User Agreement, which also states "To become a member and access the area on this Site reserved for members (the 'Member Area'), PatientsLikeMe requires that you are either (a) a diagnosed patient of the particular community you are joining, (b) a caregiver for a patient eligible to join such community, (c) a health care professional (e.g. doctor, nurse, health researcher, etc.), or (d) a guest as authorized by a PatientsLikeMe member or employee." That includes practically everyone in the known inhabited universe. It would be difficult to prove, therefore, that pharmaceutical company agents were "posing as patients" on PLM. In any case, it's easy to hire real patients who are open about their medical conditions on health discussion boards.

"Privacy" is also not the over-riding issue; at least not for members of PLM who are well aware that whatever they say in PLM's "closed" communities can be revealed to third parties either by PLM itself (de-identified) or by other PLM members (as part of a non-commercial endeavor).

Beware of Stolen "Copper"

"Transparency" appears to be the major issue here, and one that pharmaceutical companies should be wary of when they hire agents to monitor patient communities. Heywood likened it to a situation in the real world where his church's copper gutters were stolen and the copper resold to people having repairs done on their homes. The end users of the stolen copper may not have known how it was obtained, nor did they probably care. Drug companies, however, cannot be so blase—they have a lot more at stake if they accept stolen "copper" and get caught. As Jamie said, you wouldn't want to see that story in the *Wall Street Journal* (or here on this blog).

PatientsLikeMe Reports High Rate of Adverse Event Reporting Among Its Members

PatientsLikeMe (PLM)—an online community for people with "life-changing conditions"—submitted comments to the FDA about its experience with adverse events reported by its members. In short, PLM found that 7% of 500 randomly selected posts from the 364,000 posts contributed by patients within the PatientsLikeMe Forum during 2009 incorporated all four elements required for reporting an adverse event (an identifiable patient, a specific medication, an identifiable reporter and a reaction). This is 35 times the rate in a frequently cited Nielson study, which found only one such message in its random sample of 500 posts.

PLM said its study data "would extrapolate to approximately 20,000 events per year in our Forum alone. If we were to include data collected from all other sources (e.g. symptoms, treatment evaluations, side effects, annotations, private messages, research surveys) within PatientsLikeMe then the volume would be much higher. Given our current membership of approximately 55,000 patients representing only 15 diseases among hundreds that we plan to build, the contribution to drug safety and public health could be substantial."

PLM also reported that if patients were given the right tools that facilitate access to their personal medical data, the time required to report an AE to the FDA could be reduced by as much as 44% (from 36 minutes down to 20 minutes). In a pilot program in its 10,000 plus member Multiple Sclerosis community, 195 AERs were generated in one year. PLM found that "75% of these voluntary patient reports contained the four required elements for an adverse event and 24% reported a serious outcome."

Is this good or bad news for pharma companies hoping to cash in on the social media bonanza? Surely, many were bouyed by the Neilson data, which held out hope that AE's reported by patients on company-controlled or sponsored SM sites were not an insurmountable problem. The PLM data could bash those hopes.

"Social media environments that engage in promot-ing FDA-regulated medical products should take responsibility in partnership with pharmaceutical clients to proactively provide effective adverse event data to FDA as well as enhance consumer understanding of adverse events and drug safety issues," said PLM.

"Sites that integrate data collections systems and processes to capture the richness of patient experiences with drug products have an unprecedented opportunity to improve safety and outcomes for everyone who uses these products," said PLM in its comments to the FDA.

Biopharma company UCB and PatientsLikeMe have partnered to create an online, open epilepsy community that captures real-world experiences of people living with epilepsy in the U.S. Part of this partnership includes a pharmacovigilance program to monitor the site for adverse events and report direct-ly to the FDA adverse events associated with UCB products (see "Finally, a Drug Company Embraces Social Media, AEs Included!"; http://bit.ly/RwXyi).

For more about adverse events and social media, see "Solving the Social Media Adverse Event Reporting Problem" (PMN Reprint 93-04; http://bit.ly/dpQRmQ).

The lesson for pharma companies is "caveat emptor." When hiring a third party to monitor "closed" patient social networks, be sure they do not violate the policies of these communities and to do it in an open and transparent manner. For communities like PLM, the only way is to hire PLM to do the scraping for you (ie, see "UCB & PatientsLikeMe: Embracing Social Media, Adverse Events Included!"; http://bit.ly/vNpn2).

The ePatient Perspective

Susannah Fox, Associate Director, Digital Strategy, for the Pew Research Center's Internet & American Life Project and principal author of the Project's survey reports on e-patients and online health, commented on the PLM scraping incident on Pharma Marketing Blog:

"I have been watching the Facebook controversy unfold, waiting for it to come home to health care -- and here it is," said Fox. "[The] analogy of petty copper thievery reminds me of Janice McCallum's comment over on e-patients.net. I think we are talking not only about motivation, but also about scale. I think scale is part of what is tripping up Facebook, for example, and yet scale is what they need to be as useful as they are. Public health records need to be collected & maintained on a grand scale to be useful, but that's also why the gov't must protect them and have clear policies about who can access them and for what purpose. PatientsLikeMe is in the middle -- a dot-com doing the work of a public-health entity."

In a blog post, Fox urged her readers to start "a conversation about health privacy that includes an open dialogue about the risks and benefits of sharing" (see "A New Conversation About Health Privacy: Who's In?"; http://bit.ly/aYeQU8). In response to that, Giles Frydman, founder of the Association of Online Cancer Resources (ACOR), said:

"First, I am not sure I understand what damage the robot user did to PatientsLikeMe's users. I am, OTOH, clear about its real potential damage to the PLM business model. So did the robot get suspended to protect PLM or its users? This is a non-trivial question and I think the answer is as important to this conversation as any of the comments made so far. Don't take me wrong. I love PLM. The Heywood brothers have implemented one aspect of health data aggregation and sharing that will generate many ideas. But they face the same problem that ACOR has faced for many years: it is easy to create robots that will scrape sites that are supposedly protected from this kind of activity. It creates real business issues and the solutions are very hard (if not impossible) to find.

"Too many times (and maybe in the majority of cases) ownership protection masquerades as privacy protection. Danah Boyd says it best: 'fundamentally, privacy is about having control over how information flows.' This becomes particularly true in this fast expanding networked world where the value is NOT in the individual data but in the metadata generated by sharing.

"Last year I agreed to be one of the co-writers of the declaration of health data rights (see box). If we had to rewrite the declaration today I would certainly insist on adding the concept of control of the personal data flow. In a perfect world a better future will bring a high level of granularity for individual control of that personal data flow. Just as anything taken from my body belongs to me, I should have a personal stake in the metadata I help generate. The new economy will have to come up with solutions. Otherwise, I am afraid social upheaval could happen online, just as it has happened offline many times before. Just as close to 1/2 billion people have used Facebook they could decide to terminate their account."

Declaration of Health Data Rights
(http://www.healthdatarights.org/)

In an era when technology allows personal health information to be more easily stored, updated, accessed and exchanged, the following rights should be self-evident and inalienable.

We the people:

- Have the right to our own health data
- Have the right to know the source of each health data element
- Have the right to take possession of a complete copy of our individual health data, without delay, at minimal or no cost; if data exist in computable form, they must be made available in that form
- Have the right to share our health data with others as we see fit

These principles express basic human rights as well as essential elements of health care that is participatory, appropriate and in the interests of each patient. No law or policy should abridge these rights.

What Are Your Social Media Principles?

Part of every drug company's social media principles should be a pledge not to violate the policies of patient social networks. AstraZeneca, for example, might consider adding this to its "five important principles for online dialogue," which are:

- Truth and Accuracy: Content that is truthful, balanced, accurate, and not misleading

- Be Respectful: Respect the interests of patients, caregivers, and health care providers, particularly related to matters of privacy and the primacy of the patient/physician relationship

- Protect and Advance Patient Health: Provide accurate and timely reporting on medicine safety

- Transparency: Any sponsorship should be entirely transparent as to the role of product sponsors as participants in online discussions

- Respect the Views of Others: Acknowledge other opinions but don't censor or limit these views; rather, add the product sponsor's views to the ongoing discussions

(See " Principles for product-related online dialogue"; AZ Health Connections Blog; http://bit.ly/aT0O19.)

What's Your Plan?

Before you find yourself in a social media crisis that may result from inappropriate monitoring of online patient discussions, you should have a social media implementation plan. But what are the action items for implementing your plan? A recent *Pharma Marketing News* survey presented a number of actions that may be appropriate and respondents voted on how important these or other items might be in developing a plan. The choices were:

- Become a dialogue company - learn how to listen and respond, not just push messages out

- Create a social media "Tsar" position to oversee all the company's SM projects to assure compliance with guidelines

- Develop a moderation strategy

- Develop guideposts, internal and external standard operating procedures

- Get everyone - including marketing, regulatory people, corporate communications, C-level execs - on board
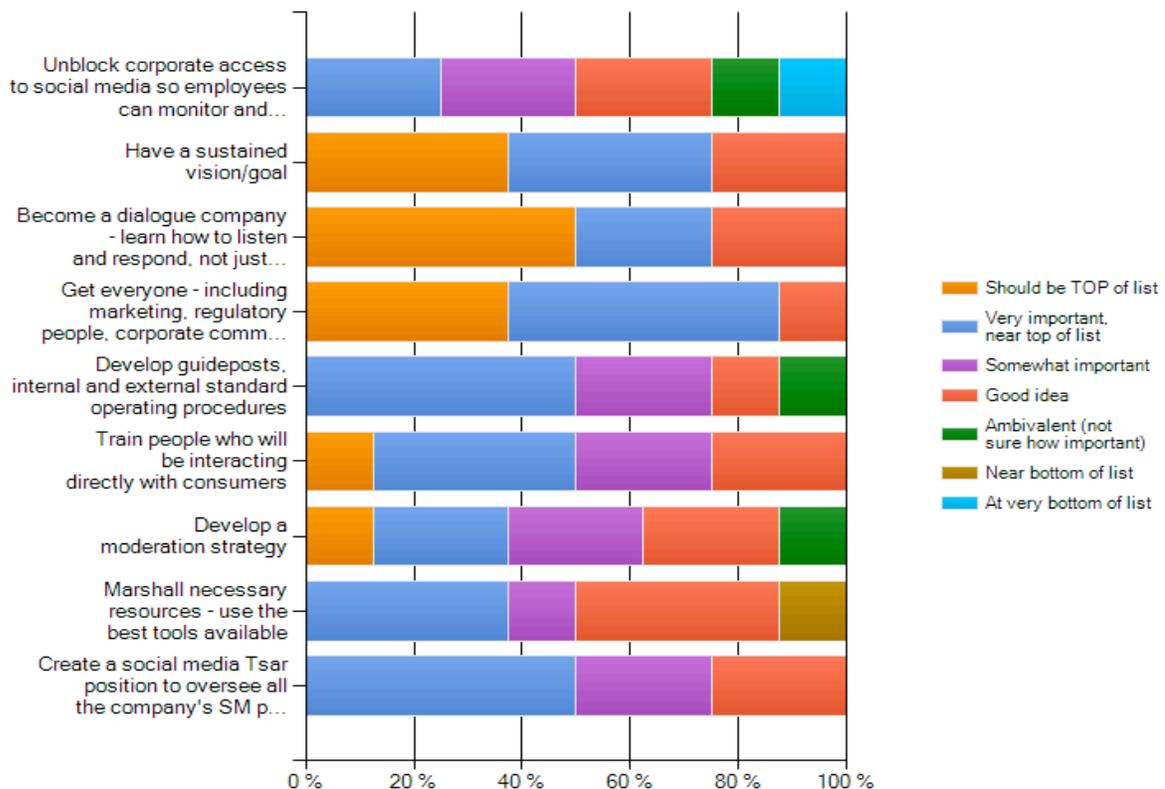
- Have a sustained vision/goal

Continues…

Figure 1. Respondents were asked to rate the importance of the social media implementation plan action items according to importance.

- Marshall necessary resources - use the best tools available
- Train people who will be interacting directly with consumers
- Unblock corporate access to social media so employees can monitor and use applications such as Facebook while at work

The preliminary results, based on only a handful of responses, is shown in Figure 1, page 6.

"Become a dialogue company" and "having a sustained vision/goal" are most important followed by "getting everyone on board." Surprisingly, however, "developing guideposts" did not rate very high with respondents. What do you think? The survey is still open; see http://bit.ly/dxRHaE.

 "Unblocking access" was one of the action items that seemed important to respondents. How much personal experience one has with using social media is critical measure of your readiness to engage in online conversations with patients and consumers. The ongoing Social Media Marketing Readiness Self-Assessment Survey (see http://bit.ly/ahmFll) shows that a substantial though decreasing percentage of pharmaceutical marketers shave little or no personal experience using social media (see Figure 2, below).

Aligning Your Message with Patient Needs

There are millions of health-related conversations on social media sites. Who are having these conversations and what are they saying? Should you monitor these conversations? What will you learn that can help you refine your marketing message to address the needs of your customers?

Pharma marketers can benefit from monitoring online patient conversations. By learning what patients are saying about their medical conditions, treatments they have tried, valuable resources they have used, and even adverse events they have experienced, drug companies can not only improve their marketing messages but also improve their products and value to patients. But it is important for the drug industry to obey the rules specified by both open and closed online patient communities and ensure that its agents also obey the rules lest they be sold "stolen copper."

Pharma Marketing News

Pharma Marketing News

Pharma Marketing News is an independent, free monthly electronic newsletter focused on issues of importance to pharmaceutical marketing executives

Subscribe Online • **Download Media Kit** •
Request a Rate Card

Personal Use of Social Media
(Pharma Respondents, N = 197)

- Non-User: 7%
- Infrequent User: 14%
- Occasional Use: 41%
- Frequent User: 39%

% Respondents

Pharma Marketing News. A Pharma Marketing News chart. (c) 2010. All rights reserved.

Figure 2. Social Media Marketing Readiness Self-Assessment Survey Results. Q: Have you ever personally used social media (i.e., read an online forum or posted a message to an online forum, submitted comments to a blog or written a blog post, edited a wiki, etc.)? Data from March 2007 through April 2010. More recent data covering the period between April 2009 and May 2010, suggests that only about 12% of respondents say they never or only infrequently use social media.