

Mining the Blogosphere

There's a whole new world of consumer opinion emerging on the internet. It's called the "blogosphere" and it is ripe with opportunity for your brands.

by
Chris Hoyt
 President
 Hoyt & Company

If you haven't tapped into the blogosphere yet, you're not alone. The results of a November *Reveries.com* survey on blogs indicate that almost 62% of the 262 marketers who participated in the survey are not currently monitoring blogs.

In addition, only 35% of survey participants think that it is either "important" or "critically important" for their companies or brands to track what is being said on blogs. Another 41% view this type of intelligence as "nice to know" but do not see it as a priority. Finally, 11% of respondents note that they are only "vaguely familiar" with blogs and do not yet know enough to comment on the potential of blogs, either as a market research tool or as a marketing opportunity.

None of this is surprising. The fact is that until recently—defined as *this year*—both the number of blogs and the people who wrote them represented a relatively small slice of the general population.

In 2002, there were only 2.2 million blogs, according to Perseus Development. Of these, 92% were written by Gen Y teenagers—a lucrative target for some—such as soft drink, music and games manufacturers—but certainly not for the marketing community in general. Then again, who's really going to bother with a discreet target of only 2.2 million bloggers when there are so many other opportunities to pursue?

TABLE 1 **Blog Demographics: 2005 vs. 2002**

	2002	2005 (Nov)	2005 General Population
Blogs	2.2MM	23MM	
Gen Y (1981-1995)	92%	46%	30.3%
Gen X (1965-1980)	6%	31%	37.8%
Boomers (1945-1964)	2%	23%	31.8%
Male	44%	57%	49.1%
Female	56%	43%	50.9%
Bloggers under 30	92%	48%	-NA-

Source: Perseus Development Corp., 2002; Pew Internet & American Life Project, *Intelliseek*, 2005, U.S. Census, 2004

TABLE 2 **Influence of Blogging**

To what degree would your decision to purchase a product or service be influenced by:

	Responses	Mean Index
Positive word-of-mouth from someone you know personally	7.7	134
Negative word-of-mouth from someone you know personally	7.0	121
A negative news story on TV or radio or in a newspaper or magazine	5.7	99
A TV or radio commercial	4.7	82
An advertisement in a newspaper or magazine	4.6	80

Source: *Intelliseek 2005 Consumer-Generated Media (CGM) and Engagement Study Oct 1, 2005*

Exponential growth

The blog situation has completely changed with lightning speed over the past three years. As of the day we write this (December 13, 2005), Technorati — a blog search engine — reports that there are now 22.9 million blogs — forecasted by industry experts to grow to 25 million by the end of December. Think this is a fad? Check out the following:

- ◆ If the 25MM prediction holds true, between Jan 1 and December 31, 2005, the number of blogs will have more than quadrupled from approximately 5.5MM to 25MM — an increase of 455%.
- ◆ The number of blogs has consistently doubled every 5 months for the past 36 months.
- ◆ Presently, 80,000 new blogs are added daily — up from approximately 20,000 new blogs daily in December 2004.
- ◆ Over twice as many people read blogs as write them. In fact, *Ad Age* recently estimated that in 2005, corporate employees will “waste” the equivalent of 551,000 years reading blogs.
- ◆ ComScore reports that blog readership at the end of Q1, 2005 was estimated at 50,000,000 souls or 33% of the U.S. adult population — up 45% from Q1, 2004.
- ◆ Approximately 55% of all blogs are active and 13% are updated weekly. Technorati reports an average of 1.2MM posts per day during November, 2005.

In addition to warp-speed growth, the other aspect of the blogosphere important to marketers is that its demographics appear to be quickly evolving to be more representative of the general population. While still heavily weighted towards Gen Y, the indications are that this imbalance will gradually correct itself as the blogosphere continues to grow (*Table 1*).

Priority or not, the reason that marketers can no longer afford to ignore this phenomenon is that, in publicly sharing their personal experiences and opinions, chances are now reasonably good that in any given week, at least several thousand of these 23MM consumers will mention your brands or services — their likes, dislikes, reasons why and recommendations to friends, and so forth.

Keep in mind that while these comments may be primarily directed to a relatively small circle of friends for each blogger, they can be read by anybody doing a blog search on a specific subject — e.g., Xbox 360, cell phone companies’ customer service, Wal-Mart, Nike or Harley-Davidson, etc. In fact, Jupiter reports that 26% of top search results for the world’s 20 largest brands are *consumer-generated media* — i.e., blogs.

As everyone reading this already knows, these kinds of comments — when aggregated — can take on a life of their own. Because they are spontaneous, unsolicited, unadulterated, frequently passionate and always immediate, they carry far more weight in influencing consumer purchase decisions than do traditional forms of advertising.

To put a dimension on this, Intelliseek — a company in the business of mining blogs — has done frequent surveys on this subject. The most recent, published Oct. 1, 2005, shows that positive word-of-mouth from bloggers has more than 50% more influence on consumer purchase decisions than any of the major forms of traditional advertising — TV, radio or print (*Table 2*).

When the blogosphere comprised only 2.2MM consumers — mostly consisting of teenyboppers talking about last week’s date — this wasn’t a concern. Now, however, with 23MM consumers — of which 54% comprise either Gen X or Boomers — it would appear to be time to start figuring out how to put this intelligence to good business use.

An opportunity, not a threat

In its November 14, 2005 cover article entitled “*Attack of the Blogs*,” *Forbes* built an 8-page article on a series of anecdotal, isolated instances of companies, products and people who had recently been ruined by loose-cannon bloggers. According to the article, this “fear of attack” is the primary reason why everyone should monitor blogs. “Some of these bloggers have just one goal and that is to do damage. It’s evil,” quotes *Forbes* in big, bold type. *Forbes*’ advice: “Spot blog smears early, before they can spread, and stamp them out by publishing the truth.”

Reveries survey respondents took a much broader view. Although some (22%) did indeed cite “early warning signals on →

TABLE 3 For What Purposes Would Your Company Monitor Blogs?

Rank	Purpose	# Responses (out of 262)	% Totals
1	General consumer understanding	62	28.7%
2	Early warning signals on issues	47	21.8%
3	Specific consumer likes/dislikes on products or services	46	21.3%
4	Overall business or brand health monitoring	27	12.5%
5	Competitive intelligence	13	6.0%
6	Product/service improvements	10	4.6%
7	Campaign tracking	6	2.8%
8	WOM sentiment of spokespeople/affiliations	5	2.3%
TOTALS		216	100.0%

Source: www.reveries.com: "Blog attack" survey verbatims, 11/15/05

← issues" as the key reason they would monitor blogs, the other 78% view blog monitoring as a positive endeavor — *i.e.*, to do things like improve general consumer understanding, learn more about consumer likes and dislikes concerning their products or services and continually monitor the health of their businesses or brands. (Table 3).

To excavate this type of intelligence, 22% of *Reveries* survey respondents report that they have somebody inside monitoring blogs on a regular basis. The other 16% say they do monitor blogs but usually only on an occasional basis.

The surprise is that only *five* respondents — or less than 2% of the total — noted that they use a professional third-party blog monitoring service to perform this task. In addition, 71% indicate that they are unaware that these types of services exist.

Three Key Issues

Given the size and growth of the blogosphere today, three key issues have emerged with respect to monitoring blogs internally: spam, lack of coverage breadth and the inability to turn results into quantifiable business intelligence.

Spam is a huge problem. Independent industry experts (Natalie Glance and Sundar Kadayam) estimate that spam — defined as consumers paid to optimize search en-

gine rankings or to push products under the guise of neutral objectivity — now comprise about 30% of all postings on blogs, message boards, forums and review sites — and up to 50% of all mentions or comments.

The "free" (no cost) blog search engines like Technorati, BlogPulse and Google's Blog Search — the principal tools available to companies who monitor internally — do not filter for spam and therefore leave it up to the analyst to decide which postings are genuine vs. bogus — an obviously impossible task when one is confronted with 23MM blogs. Technorati — the search engine mentioned most frequently by *Reveries* respondents — plays down the importance of spam by claiming that only 2-8% of all blog postings are spam. But, of course, Technorati doesn't filter for spam.

Inadequate Coverage is an issue for those who monitor internally because even though the popular blog search engines might cull 4,500 blogs out of 23MM that are relevant to the analyst, the analyst is still left with the task of personally reading 4,500 individual blogs. On this point, it is important to understand that the free blog search engines only produce the URL containing the words for which one is searching: they leave it up to the user to do all of the remaining footwork.

Questionable Results: The sheer practical inability of human beings to tabulate and roll-up the results of these types of analyses into quantifiable, usable insights stems directly from the coverage issue. To do this, one would have to create a spreadsheet and manually identify every posting, mention, positive or negative sentiment, reasons why for each sentiment, the gender and generation of the speaker for each blog *and* keep a running total of the results over time. For 4,500 blogs at 12 blogs an hour, 40 hours per week, it would take more than nine interns to do this tracking — week in and week out!

The net of the internal, free blog monitoring route is that one is confronted with all of the problems noted above and yet can produce little at the end but "gut feel" analysis. This, indeed, falls into the "nice to know" category but is obviously of little use from an actionable standpoint.

Blog Monitoring

To fill this void, an entirely new industry of third-party blog monitoring services has recently emerged. These have the capacity to scan at least 95% of all blogs, siphon the intelligence for which their clients are looking, quantify this intelligence and then format it into graphs and charts that facilitate informed business decisions.

While there are approximately 8 or 9 of these services currently available, the three that appear to be frontrunners — defined as having the best technology and the most accurate read capabilities — are VNU’s BuzzMetrics, Intellisearch’s BrandPulse (not to be confused with its free sister service BlogPulse) — and Umbria, Inc.

What separates these companies from the free search engines is that all three have developed highly sophisticated text-mining capabilities in addition to being blog search engines. In essence, these companies are in the business of quantifying the qualitative input in the blogosphere and formatting it in ways that provide their clients with customizable, usable intelligence on a broad variety of blogosphere-driven subjects, including (but not limited to) all of the subjects that *Reveries* survey respondents say are important to them as illustrated in *Table 3*.

Unlike the no-cost search engines, all of these companies charge for their services. Based on a recent *Fortune* article “Blogging for Dollars” (December 2005) — fees range from \$60K a year to fees which “can easily run into seven figures.” Whether or not these fees are accurate (we do have difficulty believing any company would be willing to pay seven figures for blog intelligence at this point in the blogosphere’s development), the fact is that there are other options.

Based on the verbatims accompanying the *Reveries* survey, it is clear that not many respondents have used these types of sophisticated text-mining services. Consequently, most of these verbatims mirror precisely the problems we cited earlier when referring to the common, everyday free blog search engines like Technorati, BlogPulse or Google’s Blog Search:

- ◆ “None of these services analyzes what is being said with any depth.”
- ◆ “Still requires a lot of footwork for the person doing the tracking.”
- ◆ “Lack of ability to monitor writer sentiment.”
- ◆ “No filtering or ranking — hard to source out the solid blogs from the fluff or ranters.”

◆ “Too many spam blogs.”

◆ “You still have to dig through too much unuseful information.”

This is really unfortunate because it appears the free blog search engines have done little but to turn off the marketing community to the benefits of blog mining in general. In contrast, all the text mining services mentioned above — BuzzMetrics, BrandPulse and Umbria — do address these issues but with significant differences in capabilities and approach:

- ◆ All, for example, quantify the number of posters on a particular topic and provide positive/negative sentiment analysis, topical depth analysis and influencer analysis.
- ◆ All provide dashboards and trend results
- ◆ Umbria, apparently is, the only service of the three that currently filters for spam, counts mentions in addition to posters and provides gender and generation data.
- ◆ On the other hand, Umbria’s approach is primarily automated so it does not offer the interpretive consulting which is a strong suit for both BuzzMetrics and BrandPulse.

A Competitive Edge

With 23MM bloggers at present and — if current growth rates hold up — the possibility of 45-50MM bloggers by mid-2006, the blogosphere offers marketers an entirely new and exciting op-

portunity to gain intelligence about their products, services and consumers that no other market research tool currently provides. Because this intelligence is passive — *i.e.*, listening to consumers rather than asking them questions — it provides a quality and integrity of input that no marketer can afford to ignore.

The flip side of this is that — if the *Reveries* survey results are any indication — the speed of the growth of the blogosphere appears to have caught most marketers unaware. Not only do the majority of respondents not monitor blogs at all but the majority of those who do attempt to do so internally which, for reasons noted previously, rarely delivers usable results.

Because of the level of detail, accuracy and reliability of the intelligence provided by third-party professional blog monitoring services such as BuzzMetrics, BrandPulse and Umbria, the (few) *Reveries* marketers who use these services are way ahead of the pack. There is no way internal monitoring can begin to emulate the capabilities of these types of services.

The good news is that because so few marketers are using these types of services, most everyone has an equal chance to begin now and clip off a genuine competitive edge.

If you still have doubts, get on a blog search engine now and pull up either your own or your competitor’s brands and see what your vulnerabilities and opportunities are. ■



Chris Hoyt is President of **Hoyt & Company**, a packaged goods training and consulting organization based in Scottsdale, AZ. Previously, Chris was president of the sales consulting division of Glendinning Associates and national sales manager of the the Clairol division of Bristol-Myers. A Fulbright Scholar and Phi Beta Kappa and Magna Cum Laude graduate of Princeton University with a degree in mathematics and philosophy, Chris began his career in sales at the Procter & Gamble Distributing Company. He can be reached at chrishoyt@hoytnet.com