



NEWS ANALYSIS April 6, 2007, 12:00AM EST

Cleaning Messy Message Boards

Companies are turning to a spate of new filtering tools to keep online conversations from devolving into either hate-filled arguments or meaningless drivel

by [Catherine Holahan](#)

The story was sure to provoke contentious debate among readers. Digg users had posted a piece on the design characteristics of Christian Web sites. Reader comment sections and message boards frequently turn hostile when far less polarizing subjects than religion are discussed.

Yet the conversation on Digg was surprisingly civil. Sure, there were some sarcastic comments. "Good design is created, not evolved," for example. But the back-and-forth didn't devolve into calling the sites "stupid" or less flattering names.

FILTERS CAN BE FACILITATORS

The reason for the civility may well be the design. Digg employs community filtering tools that let users promote comments to prominent positions and demote—or "bury," in Digg parlance—those deemed inappropriate, useless, or just plain mean. Frequently "buried" comments disappear from view. Digg co-founder Jay Adelson says the solution isn't perfect but it helps to keep some kind of decorum. "You want to create an open debate on an open forum that isn't censored so that people have the freedom to speak their mind," says Adelson. "But filtering is critical in order to sift through conversations and make them useful."

How to filter conversation has become a topic of intense discussion in its own right. As more and more Web surfers become familiar with message boards and comfortable firing off opinions, often anonymously, online sites are struggling to manage comments without either stifling conversation or becoming a platform for hateful speech.

The conversation over comment filtering is not new. But it was amplified recently, after a female technology blogger was threatened with sexual assault and murder in the comments section of another blog. For many online sites and publications, the ensuing blog discussions over how the hateful comments should have been handled underscored their own challenges with comments (see [BusinessWeek.com](#), 3/28/07, "[Dispatches from the Blog Battle Zone](#)").

BALANCING SPEED AND SCALE

Chief among those challenges is scale. Many online publications, including *BusinessWeek*, have human editors who scan through reader comments and approve what is posted. The goal is to ensure that, while critical comments are published, the conversation is not filled with hate speech, death threats, spam, or other attacks that can destroy reasonable debate or discourage others from participating in the discussion. The challenge with this approach is it can be time-consuming and many publications are not staffed to sift through large influxes of comments fast enough to satisfy message-board participants' desire to quickly see their posts. "To approve every comment before it goes up tends to syncopate the conversation so much that you might as well not have the conversation," says Jeff Jarvis, director of the interactive journalism program at the City University of New York's journalism graduate school.

The Washington Post ([WPO](#)) dealt with the scale issue publicly after one of its reader comment sections was flooded with

what the paper termed "hate speech," "profanity," and "personal attacks" (see BusinessWeek.com, 4/16/07, "[Web Attack](#)"). Unable to keep up with the influx of disturbing comments, the newspaper turned off the feature on the particular story that had incited the crowd. Afterward, the online publication's executive editor Jim Brady held an online discussion concerning the decision. "We need to look at how we're staffed to handle comments," said Brady. "We also need to look at the technology, specifically how much weeding out of offensive content can be automated."

However, the available filtering tools do not offer any easy solutions. Profanity filters, used by many online publications, only block or flag comments with specific unwanted words. Companies such as Neighborhood America provide language filters that let sites like ABCnews.com ([DIS](#)) search comments for specific words or phrases, or words located within a certain frequency of other words that indicate a post may be problematic, says Neighborhood America Chief Executive David Bankston. For example, the software lets site administrators search for the word "hate" in proximity to words describing an ethnic or religious group. "That way you're able to do a quick scan of what the comment is saying," Bankston says.

Not all the problematic posts are negative. For example, there are users who may want to flood ABC's site with exclamations of love or adoration for a particular on-air personality. They may be nice, says Bankston, but they don't do anything to promote conversation about a particular story. The company is working on other automated tools that will recognize images of naked bodies in video comment posts and automatically remove them.

COMMUNITY POLICING

Some sites leave much of the filtering up to the message-board users in hopes that the community will promote the appropriate level of decorum. Slashdot, an aggregation and comment site for what it calls "nerd-oriented news," has a revolving group of community moderators that are able to promote or demote comments through a scoring system. The goal, according to Slashdot's own FAQ message board, is to make the readers of the site take responsibility for what appears on it. Slashdot also gives comments higher scores when users are logged in and thus have a unique identity to which their comments can be tied.

Digg's system has a similar registration process by which users develop a reputation from their posts. The hope is that users will want to have a reputation for insightful thoughts and not, as the online community calls people who consistently post pointlessly negative comments, being a "flaming troll."

Adelson says Digg is working on additional moderation tools. "I haven't seen a perfect solution to the problem," says Adelson. "Digg has only just scratched the surface with its tools."

Perhaps the best moderation tool, says Jarvis, is joining in the conversation. He believes that if site owners and publications respond to commenters, users will consider the boards a place for adult conversation rather than a place for venomous rants.

[Holahan](#) is a writer for BusinessWeek.com in New York.

Xerox Color. It makes business sense.

Copyright 2000-2007 by The McGraw-Hill Companies Inc. All rights reserved.

The McGraw-Hill Companies