BY CHRISTINA K. PIKAS

BLOG Searching

for Competitive Intelligence, Brand Image, and Reputation Management

Do you know what is being said about you in the blogosphere? Should you care? Aren't blogs just for kids, political campaigns, and geeks, of no value in business? If you think this, you're likely missing out on an increasingly powerful information source.

Why have blogs, which are simply Web pages with reverse chronologically arranged individual posts, become so powerful? According to David Sifry of Technorati [www.sifry.com/alerts/archives/000298.html], the number of blogs has been doubling at least every 5 months. The current estimate is that there's over 10 million blogs in the blogosphere. Yet it's not the large number that creates the stir over blogs. Their power comes from the speed of information diffusion, their low barrier to access, and readers' ability to scan many blogs through XML feeds. The vast majority are not journalistic endeavors or political advertising. Blogs are as individual and unique as their writers—

and can damage your company or brand very quickly.

CONSEQUENCES OF BLOG POSTS

Take, for example, the case of Kaiser Permanente. A former employee, Elisa D. Cooper, who calls herself the "Diva of Disgruntled," posted a link to private customer health information, including information on lab tests, on her blog. Kaiser first learned about the posts from the federal Office of Civil Rights in January 2005, although she says she posted the link in July 2004 and the site to which she linked had been online since 2002. A lawsuit, filed by Kaiser, is in progress.

Another recent example is the case of the easily picked Kryptonite lock. In 1992, a bicycling enthusiast magazine, *New Cyclist*, published an article stating that circular locks similar to Kryptonite bike locks can be picked with a Bic pen. The publication of this information had very little impact and Kryptonite made no changes to the lock design. Then in September 2004, a video posted to a blog showed how to pick a Kryptonite lock in less than 30 seconds using a Bic pen. In less than 2 weeks, this video was played in all of the major media, lawsuits were brought against the company, and a movement was started to force the company to recall the "faulty locks." Even though the information was

available to bike shops and bike owners for more than 10 years, it took blogs to precipitate the crisis.

In the mainstream media where story leads can originate from company press releases, companies are in charge of the message. In the blogo sphere, where the power is in the network, the customers control the message. A good illustration of this is in *Harvard Business Review*'s realistic "Glove Girl" case study ("A Blogger in Their Midst," by Halley Suitt, vol. 81, no 9 [September 2003]: pp. 30–40). Glove Girl, who's an assistant foreman at a surgical glove manufacturing plant, blogs about the company's products on her personal blog. She achieves more audience than the official marketing campaigns and upstages the company's CEO at a major trade show. When she says something positive about a line of gloves, sales increase. When she blogs her questions about the company's ethics of doing business with a clinic with a poor record, the deal is in danger. Commentators on the case study were David Weinberger, Pamela Samuelson, Ray Ozzie, and Erin Motameni.

The blog cycle—the time between learning about something to having it on a blog feed—is very short. A blogger reads something of interest, and then posts it to his or her blog within seconds, frequently without verifying facts. Feed aggregators query the site
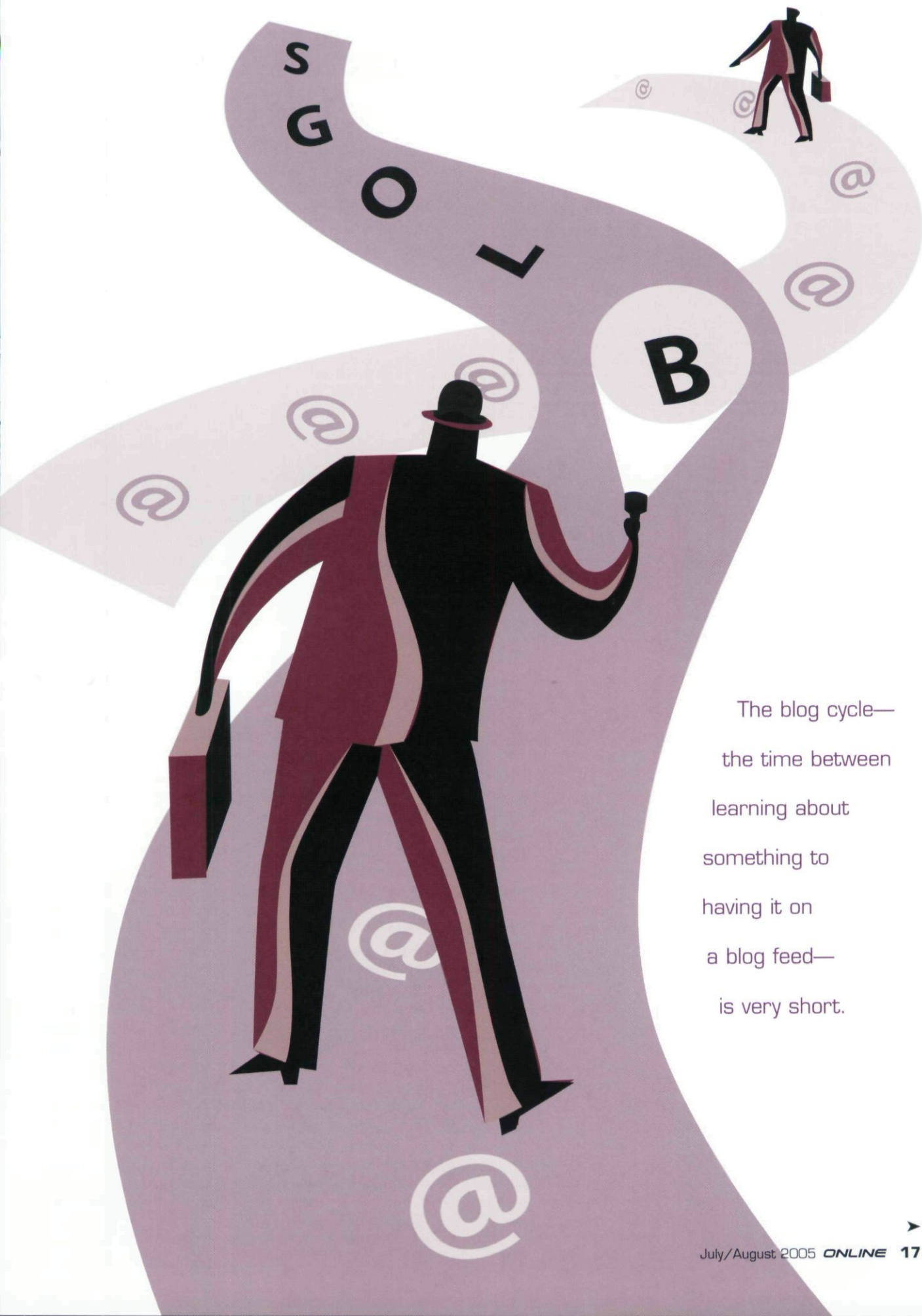
The blog cycle— the time between learning about something to having it on a blog feed— is very short.

about every hour, so readers all over the world will have almost instant desktop or mobile access to the story without searching for it. This can be great viral marketing of your product, if the post is positive. However, if the posted information is negative or even wrong, and your company is the last to know, damage control will be very difficult.

*Knowing the **structure of blogs** and blog posts in general will enable you to design your search more effectively.*

COMPREHENSIVE SERVICES VERSUS CREATING YOUR OWN

Are these examples, both real and fictionalized, enough to convince you to search across blogs for information about your company, your competitors, and your brand? Are you convinced you should create alerts so that you are up on the buzz in the blogosphere? If so, your next task is to decide how you will proceed.

As I write this, I am seeing many announcements of blog monitoring products and research from startup companies and established research labs. While this supports my point that it is important for companies to monitor the blogosphere; it makes it much more difficult to select a solution. Several companies have developed independent products that match new blog posts, search feeds, or search blogs for company, brand, or product information. Others partner with major information services to integrate this feature into existing business and news research products. For example, Moreover is adding blog information to Factiva's reputation management products. Intelliseek, BuzzMetrics, and Moreover are just a few of the companies marketing a monitoring product. In-

telliseek offers the content analysis of posts to indicate if they are negative or positive.

While providing easy solutions, the above products may not offer the precision of a custom search and may be very expensive. It is not necessary to license a commercial product to effectively search across blogs and to set up alerts for new content. Once you understand the blogosphere and how to use advanced features of general and feed or blog-specific search engines, you'll be able to get better results and will be able add new searches or tweak existing searches for better precision or recall.

THE STRUCTURE OF BLOGS

Knowing the structure of blogs and blog posts in general will enable you to design your search more effectively. In addition, knowing the specific structure of Web pages produced by the most common blogging software packages will allow for effective field searching. Although the above definition of blogs is broad enough to cover a large variety of Web sites, the lion's share of the market is held by a few hosted blogging packages such as Blogger, Type Pad, and Live Journal, according to Elise Bauer's Weblog Tools Market—Update, February 2005 [www.elise.com/web/a/weblog_tools_market_update_february_2005.php].

Blogs have a very low barrier to access. You post using e-mail, voicemail, Web forms, or a downloaded WYSIWYG program. Hosted blogs have preformatted template choices that allow varieties of layout and

color without detailed knowledge of HTML, cascading style sheets (CSS), or XML. The content can be anything from a sparse list of links to extensive collections of essays, videos, recordings, or pictures. Feeds are XML-formatted abbreviated posts or lists of headlines than can be gathered, displayed, and consumed efficiently in an aggregator.

Blogs generally have home pages with somewhat static content and a list of the most recent posts. Blogs also have individual post pages that contain a single post and may not display all of the options from the home page. Home page formats usually consist of two to three columns using tables and CSS. The largest and frequently most prominent column contains the posts that form the actual meat of the blog. (See Figure 1 on page 19.)

These individual posts are chunks of text and links arranged in reverse chronological order. Each post generally has the following fields: title, date-time stamp, body, comments, trackbacks, and permalinks. Other front-page columns may be more static or contain automatically generated content, such as headlines from recent posts, links to the blogger's personal profile, links to the archive, and a list of favorite Web sites. Archives are created immediately so the same post will occur on the main page and on the archives page. (See Figure 2 on page 19.)

URL STRUCTURES

The structure of the archives, post pages, comment pages, and permalinks can be exploited for precision searching. The table on page 20 provides examples of the format of the URLs.

GENERAL SEARCH ENGINES

Although there are several specialized blog search engines, it's wise not to overlook the general Web search engines—Google, MSN, Yahoo!, and Ask Jeeves.

I'll start with Google. Once you create your search in Google, you can visit Google Alerts [www.google.com/alerts] to paste in your successful search string and set up immediate or daily e-mail alerts. For this

Anatomy of the Main Blog Page

Title and description

Content (posts)

About the blog, profiles of the contributors

Blogroll

Sort by category

Archives by date

Source: http://stlq.info

Figure 1: Anatomy of the Main Blog Page

Anatomy of a Post

Date/time stamp

Permalink

Categories (click to find all in that category)

Citation

Title (may be hyperlinked)

# of comments – (click on permalink to see them)

# of trackbacks (click on permalink to see them)

Click to open the form to leave comments

AZoNano Launch Online Open Access Journal of Nanotechnology

Posted at 10:23 on Decem 9, 2004 in Nanotechnology and Open Access. | 0 Trackbacks | 0 Comments

:: A new, open access journal on nanotechnology is forthcoming:

AZoM.com Pty. Ltd. Sydney Australia and the UK-based Institute of Nanotechnology are pleased to announce the forthcoming launch of the Online Journal of Nanotechnology.

The Online Journal of Nanotechnology is based on a free access publishing model, coupled with what is believed to be a unique development in the field of scientific publishing – the distribution of journal revenue between the authors, peer reviewers and site operators.

The Online Journal of Nanotechnology at AZoNano.com will publish high quality articles and papers on all aspects of nanotechnology and related scientific, social and ethical issues. All the contributions will be reviewed by a world class panel of founding editors who are experts in a wide spectrum of nanotechnology science.

Via On Christina's Radar.

Post a comment to: AZoNano Launch Online Open Access Journal of Nanotechnology

Source: http://stlq.info/archives/001707.html#001707

Figure 2: Anatomy of a Post

reason and the overall success with my trial searches, I highly recommend trying this search engine first.

To use Google Alerts for monitoring brand image, try this general search string: `"product name" (inurl:YYYY/MM1 MM2 OR inurl:YYMM YYMM OR inurl:archive)`.

This works because Google indexes the "/" character and uses the ellipsis range operator. Be sure to put your product name in quotes. Select the year and the current few months or omit the ellipsis and use just the most current month. The archive term adds noise, in the form of non-

blog sites, but is necessary to be more comprehensive. Of course, you'll need to update the dates from time to time to get new results.

The above method will also pick up most Movable Type and Word-Press blogs that are hosted on the blogger's own site. The search string `"product name" (inurl:blogspot OR inurl:typepad OR inurl:livejournal OR inurl:diaryland OR inurl:blogdrive)` is more restrictive, but will pick up the hosted blogs with the majority of the market share. It is also more precise because it avoids the noise from e-com-

merce and magazine sites that may use the date in the URL.

Moving to MSN, once you narrow down your search, click on the RSS button at the bottom to subscribe to the feed for the search so new results will be sent to your aggregator. A good search string at MSN is `"product name" (site:blogspot.com OR site:typepad.com OR site:livejournal.com OR site:diaryland.com OR site:blogdrive.com)`.

MSN lacks a command that allows you to search in the URL and it doesn't index slashes and other

URL STRUCTURES

Blogger	URL	http://Name.blogspot.com (if Blogger-hosted)
	Post Page	http://Name.blogspot.com/YYYY/MM/words-of-title-hyphenated.html
	Comments	http://Name.blogspot.com/ YYYY/MM/words-of-title hyphenated.html# 000000000000000000 (18-digit comment ID) (Note: Many blogger blogs user Haloscan or other plug-in comment services that store the comments on nonblogger servers.)
	Archives	http://Name.blogspot.com/YYYY_MM_DD_Name_archive.html (The date is the first date of the archive, may be weekly or monthly archiving.)
LiveJournal	URL	http://www.livejournal.com/users/Name
	Post Page	http://www.livejournal.com/users/Name/00000.html
	Comments	http://www.livejournal.com/users/Name/00000.html?thread=11111#t11111
	Archives	http://www.livejournal.com/users/Name/YYYY/MM/DD (Month and day are optional.)
Typepad	URL	http://Name.typepad.com/Name2/ (Name2 may be similar or the same as Name; it may also be omitted.)
	Post Page	http://Name.typepad.com/Name2/YYYY/MM/words_of_title.html
	Comments	http://Name.typepad.com/Name2/YYYY/MM/words_of_title.html#comments
	Archives	http://Name.typepad.com/Name2/YYYY/MM/index.html
Diaryland	URL	http://Name.diaryland.com/index.html
	Post Page	http://Name.diaryland.com/YYMMDD_00.html
	Comments	n/a
	Archives	http://Name.diaryland.com/older.html
Blogdrive	URL	http://Name.blogdrive.com
	Post Page	http://Name.blogdrive.com/archive/0000.html (numbers refer to a post number)
	Comments	http://Name.blogdrive.com/comments?id=0000 (same post number)
	Archives	http://Name.blogdrive.com/archive/cm-0_cy-0000_m-2_d-00_y-00_o-0.html

KEY: Y – year, where YY is a two digit year (05) and YYYY is a four digit year (2005); **M** – month, where MM is a two digit month (03); **D** – day; **N** – Name is the name of the blog of account name; **Title** – indicates the title of the post.

symbols. The search string above is very narrow, but since feeds can be set up, you can easily establish quite a few alerts.

In Yahoo!, you can get alerts delivered to your cell phone or in e-mail when the new item is in the news, but not new items for the Web. Thus, you will have to manually re-run the search periodically. Here's the search string I prefer: "product name" (inurl:blogspot OR inurl:typepad OR inurl:livejournal OR inurl:diaryland OR inurl:blogdrive).

Click directly on the advance search page to input all the hostnames into the "any of these" box, then click the drop-down menu for "in the URL of the page." Yahoo! does not index symbols and does not have a range operator. Therefore, a more restrictive search is preferable in this search engine. Because there is no alert for the search, it might be useful to limit the search to the last 3 months. You may also look for the month and year: "product name" inurl:YYYY inurl:MM.

Ask Jeeves owns Teoma and uses the same search technology. I was unable to get consistently good results using either site. Alerts are not available. I do not recommend using either for this task.

BLOG-SPECIFIC

Many so-called blog search engines really only index feeds. Feeds are XML versions of the blogs that can be read using an aggregator or incorporated into portals or Web pages. Many feeds are not full text. In fact, some just include the title of the post. Additionally, many scholarly journals, mainstream media outlets, and government agencies offer feeds. Searches in these engines provide many more relevant posts higher in the results list, but return fewer results overall. Below are several of the blog specific search engines that I have found to be most useful. In each of these, no extraordinary search skills are required, because the assumption is that all of the results will come from blogs, forums, or feeds.

Bloglines [www.bloglines.com] is a Web-based feed aggregator now owned by Ask Jeeves. The search engine allows you to search all of the feeds to which Bloglines members subscribe. Once your search is complete, you can set up an alert within your account on Bloglines. For companies wanting to have the information delivered to their intranet, this may be unsatisfactory.

Feedster [www.feedster.com] has a very powerful search engine with many fields you can exploit to provide a precise search. Use the blog-only portion of the database [http://blogs.feedster.com/] to exclude news sources. Once you've tweaked the search the way you want it, you can subscribe to a feed to get new results. It is worth reviewing the advanced help page [www.feedster.com/help/advanced.php] to see what is available.

PubSub [www.pubsub.com] is not really a search engine because it matches your query against new blog posts from millions of feeds it monitors—it does not run the search against stored information from crawled or spidered sites. In this sense, it fits into the "comprehensive services" section above, but it is free for now and does not provide visualization or analysis functions. Set up your search using the box on the home page and subscribe to the feed or bookmark the digest page provides. It automatically ORs terms, so use quotes, ANDs, and parentheses. I highly recommend this source, as it monitors the largest collection of blogs.

Technorati [www.technorati.com] developed as a service to allow bloggers to track conversations on the Web. It tracks over 5 million blogs and reports when another blog links to one of your posts. More recently, Technorati has developed tags that allow bloggers to categorize their posts. See, for example, the tag "Bluetooth" [www.technorati.com/tag/Bluetooth]. Results with the same tags are imported from Flickr, Furl, and Del.icio.us—Web-based bookmark and photo managers. Create an account and conduct a simple search. You will then be given the option to "make

this a watchlist." This creates an RSS feed that will alert you when new posts match your search.

Waypath [www.waypath.com] covers fewer blogs than Technorati or PubSub, but it allows two types of RSS feeds of results: newest or best. It is also one of the few search engines that indexes the entire blog, not just the feed.

MANAGING RESULTS

I've shown you many ways to get early-warning information on the buzz in the blogosphere—almost too many. How do you go about setting up all of these alerts? What do you do with all of this information? First, get an aggregator so that you can scan incoming feeds easily. If you use Bloglines or another Web-based aggregator, your subscriptions do not have to be linked to your company (or name). In addition, you can create temporary e-mails within Bloglines to use for Google and other e-mail alerts. Even if you do use a desktop aggregator, sign up for alerts using a Web-based free e-mail account.

Second, try each of the searches above with your actual company or product name. For testing purposes, I used a fairly new mobile phone offering, the "Motorola A630." Your results may vary from mine so keep the best general search engine, PubSub, and the best blog search for you. Third, set aside time in the morning to haul in your nets. It is very tempting to check and re-check your feeds. Once the feeds are established, let them work!

Now that you know what is being blogged, it's time to turn the tables, take the initiative, and create a marketing blog for your company or organization. Your blog can counter adverse opinions in real time and in the medium bloggers understand. This may seem like quite a bit of effort, but your company and your customers will thank you for it!

Christina Pikas [cchristina.pikas@jhuapl.edu] is a librarian at The Johns Hopkins University Applied Physics Laboratory.

Comments? E-mail letters to the editor to marydee@xmission.com.