

**Free
webcast**

**Business Continuity: Stay Up
and Running Against All Odds**

Sponsored by

Strategies for today's
technology challenges

WorldCanada.com
webcast

Mobile | Feedback | Contact | Friday, June 4, 2004 Advanced Search

Subscribe!

COMPUTERWORLD*

NetworkWorld*

**CIO
CANADA**

itFocus

**CIO
CANADA** Government
Review

**Job
Universe.ca**

Knowledge Centres | Events | News | Resources | Media Resources

**What keeps YOU
up at night?**

Tech Topic

Knowledge Centres

- Enterprise Infrastructure
- Communications Infrastructure
- Information Architecture
- Integrating IT
- Departmental/End-User Computing
- Enterprise Business Applications
- Extended Enterprise
- Security
- Voice Data and IP
- IT Workplace
- Leadership
- E-Government
- Wireless/Mobile Computing

**FREE
weekly
newsletter!**

subscribe

Security
➤ Technical

A (much) clearer eye for the antispam guy

By: Alan K'necht (28 Nov 2003)

A few months ago I published a column on how to prevent spam before it starts by making a few simple changes to Web page code. The basic solution is to hide e-mail addresses from those nasty spammer utilities that crawl Web sites looking for them.

The column prompted some feedback - some of you disputed one of my recommendations, others offered additional tips to make it harder for spammer worms/bots to harvest e-mail addresses. I compiled the best of the submissions and criticisms, and here they are.

Are they worms or bots? In my column I called the utilities that harvest e-mails from Web sites "worms." Some people equate worms with a specific type of virus, and it was pointed out that these utilities work more like search engine bots (short for robots) and hence should be referenced as bots. A rose by any other name...

Use of hexadecimal values for e-mail address in code. One reader pointed out that spammers are now modifying their programs to account for this. So why, he wondered, recommend doing this at all?

This is true, but this code modification is a simple fix not designed to block all attempts - just the majority, and only for now. On my Web site I've been using this technique for a contact e-mail address that, when instituted, was spam-free and remained so for more than six months, until a publisher (not this one) inadvertently published it on their site without the use of the hexadecimal values. The e-mail address was only visible for about 36 hours before the error was rectified, but I now get about 10 to 15 spam messages a day via that address. They started appearing in my inbox about one week after the e-mail address was published. Coincidence? I think not.

Block the worms/bots before they scan your site. Just like search engine bots look for a Robot.txt file before scanning your site, it

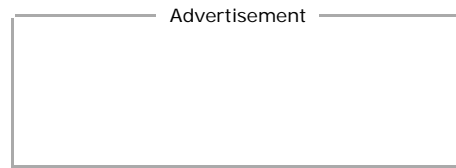
**The
Unwired
Office
starts here.**

- Longer battery life
- High performance

1 of 3

6/4/2004 10:02 AM

turns out that people believe that some of those nasty spammer utilities do the same thing. So why not disallow them access to your site by including them in the Robot.txt file?



I've yet to prove or disprove that you can effectively block them with this methodology. I did find several Web sites that reference these nasty bots, and I've compiled a list of crawlers to block, plus the code that might do the trick (www.knechtology.com/stop-spam/robots.txt).

My thoughts on this recommendation is, "it couldn't hurt," even if I think it's unlikely that all these worms/bots, whose sole purpose in life is to steal people's

e-mail addresses, would actually take the time to see if they have permission to do so. If you truly want to block these bots, you can create a ".htaccess" file that blocks them based on their user User-Agent string (see:

http://www.clockwatchers.com/robots_bad.html). This solution from a purely technical perspective will work on Apache servers. The trick is how to know which bots to block.

Use an alternative e-mail address to post to newsgroups and other places on the Web. I couldn't agree more with this recommendation. If your job requires you to post to newsgroups and discussion boards, get an extra e-mail address from your company solely for this purpose.

There is a great free service that provides throwaway e-mail addresses for just this purpose. Be sure to checkout www.sneakemail.com. The beauty of this product is that when used correctly, you can actually figure out where they got your e-mail address and complain to the source. You can also use services like Hotmail and Yahoo Mail.

Use Cascading Style Sheets (CSS) and a graphic. (Provided by John A. MacDonald of Innisfil, Ont.) This was one of the more intriguing suggestions, and my favourite (even if I could only get it to work with IE). While I was aware that you could use the URL function in CSS to insert background images behind an element, it never occurred to me to use it to hide an e-mail address.

Here's how it works. First you create a graphic of your e-mail address, as you want it to appear on your Web page. Insert the graphic into the page. Make the image clickable by placing an empty "ahref" tag around it (e.g. ``) and assign a class called "mail" to the ahref tag (e.g. ``). Now in your style sheet create that class as a subset of the "A:Active" function called mail and insert the value "background: URL('mailto:you@your.com')" in the class.

The appearance to the user is transparent and the e-mail address is completely hidden from the worm/bot. Now, when the user clicks on the graphic of the e-mail address in your Web page, their

e-mail program simply pops up. This lets you promote your e-mail address on your site and allows the user to contact you with their personal e-mail program. Worried about accessibility? Add the appropriate alt tag to the graphic, but put in "at" instead of the "@" symbol. Note: you can see an example of this at www.knechtology.com/stop-spam/.

Thanks to everyone for comments and suggestions. If we all start implementing these ideas, perhaps we'll be able to slow down those worms/bots. Sure, that might be wishful thinking, but it's a start. If you have any more tips, send them in.

K'necht is a regular speaker and president of K'nechtology Inc., a technology strategy, search engine optimization and Web development company. www.knechtology.com.



page 1 of 1

©2004 ITworldcanada.com All rights reserved.

[Copyright Information](#)

[Privacy Policy](#)