Bill would force Web sites to delete personal info

By Declan McCullagh

Story last modified Mon Feb 13 15:30:32 PST 2006

A bill just announced in Congress would require every Web site operator to delete information about visitors, including e-mail addresses, if the data is no longer required for a "legitimate" business purpose.

The proposal, introduced Wednesday by Rep. Ed Markey, seeks to import European-style privacy regulations by imposing a broad data-deletion requirement. It would apply to every U.S. Web site, even ones run by individuals, bloggers or nonprofit groups and charities.

Markey said the measure would help stop identity theft. "This warehoused personal information about consumers' Internet use should not be needlessly stored to await compromise by data thieves or fraudsters, or disclosure through judicial fishing expeditions," the Massachusetts Democrat said in a statement.

Also, Markey said, the bill would help address the issue of search engines storing data about their customers' search terms, a subject that received attention when the Department of Justice subpoenaed Google, Yahoo and other sites for such information.

It's not clear that Markey's proposal, called the Eliminate Warehousing of Consumer Internet Data Act of 2006 (click here for PDF), would have much effect on attorneys seeking search terms through a subpoena. It defines personal information as including name, home address, e-mail address, telephone number, and so on—but it doesn't explicitly include search terms or Internet addresses.

NetCoalition, a lobbying group for Internet businesses, was skeptical of Markey's proposal. The group represents companies including Google, Yahoo and News.com publisher CNET Networks.

"You're putting the federal government in the position of deciding for U.S. companies what a 'legitimate' business purpose is," said Markham Erickson, policy director for NetCoalition. "It's a blunt way to address a more nuanced problem."

Most Web sites keep logs that record which Internet addresses visited their Web sites. In some cases, that can identify individuals--such as rr.cs.cmu.edu, which refers to Raj Reddy, a computer science professor at Carnegie Mellon University. Others are as ambiguous as cache-mtc-ac06.proxy.aol.com, which might act as a proxy for thousands of America Online customers.

Because Markey's bill defines personal information in a vague manner--"information that allows a living person to be identified individually"--it's not clear whether an Internet address must be deleted by default. Violations can be punished as a "deceptive" business practice by the Federal Trade Commission, even if the Web site operator is an individual instead of a business. 

"This is a product of Google failing to get ahead of this problem," said Jim Harper, an analyst at the free-market Cato Institute. "They've known that people are concerned about their privacy policy, which allows them to keep data forever."

In response to a survey conducted last week by CNET News.com, Google confirmed that it ties search terms to Internet addresses. (John Battelle and Adam Fields received a similar response.)

Harper suggested that Google and other search engines find ways to aggregate visitor records without attaching personal identification. He added, though, that Markey's bill is overly broad and "clearly a hamhanded attempt at regulation."

Copyright ©1995-2006 CNET Networks, Inc. All rights reserved.