Networking: Not-so-secret documents

February 6 2006

Last fall a controversy erupted when the details of the assassination of former Lebanese Prime Minister Rafik Hariri were revealed in a United Nations report -- after a cunning reader spotted a "track changes" mistake in the layout of the document. That political controversy is one of the latest tempests to emerge over "metadata," or data about data, contained in Microsoft Word and Adobe PDF documents, easily accessible by millions of readers over public networks like the Internet, experts tell United Press International's Networking.

By clicking on the "track changes" feature in Word, readers can see who wrote a particular document, when it was written, what edits were made and comments made by editors and redactors -- something government officials, working with official secrets, or confidential information, most definitely don't want released for review in the court of public opinion. By deleting text blocks -- used to blacken out information in PDF files -- readers can see what was originally written there.

There is also a danger that computer-savvy terrorists could use the "track changes" feature of the word-processing software and other techniques to access data the U.S. government doesn't want them to see -- such as negotiating options discussed in earlier drafts of government documents, troop-deployment schedules or other top-secret information.

The National Security Agency, the government's electronic eavesdropping agency, worried about these, and other, digital document threats, recently issued guidance to federal agencies about how to properly "sanitize" Word and PDF documents about to be sent out over

networks. It is up to each federal agency, however, to implement the suggestions, experts said.

The government isn't the only one concerned about metadata problems -- companies are too.

"The specific concern that lawyers have had -- and that I believe NSA is concerned about -- is text remains with the document when it is saved as a file and delivered to another party," Paul Dalton, an attorney with the Dallas-based law firm of Cowles & Thompson, P.C., told Networking. "Microsoft Word, in particular, stores information that has been deleted from a document within the document file itself. That's how the 'undo' and 'redo' features -- which we all find helpful -- are able to bring back several levels of prior text as one edits a document."

Another vulnerability is the "comments" feature in Word, Dalton said. "Those comments normally are hidden from view until the user turns them on using the 'show' feature under 'track changes.' If one creates a document and sends it to someone else for review, that person adds his thoughts using the 'comments' feature, and the document goes out without those comments having been affirmatively removed, a subsequent reader would be able to see all of the initial reviewer's recorded observations."

Finding that hidden text could be very embarrassing to a lawyer who represents the facts and the law one way to one group of people, then writes another, completely contradictory view in a document draft, and then, lastly, changes what is written there for public consumption in the final draft of the document.

Computer experts have known about metadata problems for years, but the problem is just now coming to public attention, especially as communicating over the Internet and private networks has become the

norm.

"The action by the NSA, however, starts the ball rolling on another issue -- standards of care for the people and enterprises that are delivering the electronic document," Dan Venglarik, an attorney with the law firm of Davis Munck Butrus, P.C., told Networking. "But with the U.S. government now having defined specific procedures for ensuring that confidential information is not inadvertently 'leaked,' a definite threshold is apparently set for negligence purposes. Damages on such negligence claims are likely to be difficult to prove in most cases, but I would not be surprised to see some cases starting to be brought over the next few years."

These problems with metadata are more than a mere "formatting glitch," said Joe Fantuzzi, chief executive officer of Workshare Technology Inc., a developer of document-management technology with offices in London and San Francisco. "There are serious security concerns with the major document formats that businesses and the government use every day," said Fantuzzi. "And if these problems can happen at the White House and large companies like Merck, they can happen anywhere."