(U//FOUO) OPSEC Bulletin: Converting Sanitized Text from Word to PDF? -- Caution!

FROM: [redacted] E-Space Analysis Center (S1E)
and the OPSEC Working Group
Run Date: 04/03/2006

(U) You've just finished using Microsoft Word to sanitize a text for release and you convert it to the Adobe PDF format. When you look at the PDF file, you see that material that you edited out of the Word file has mysteriously reappeared! How did this happen?

(U//FOUO) A report issued recently by the Information Assurance Directorate (I333) - " Redacting with Confidence: How to Safely Publish Sanitized Reports Converted From Word to PDF " -- highlights this problem. The report notes, " *As numerous people have learned to their chagrin, merely converting an MS Word document to PDF does not remove all metadata automatically.* " It cites three examples of **sanitization snags** :

1. **Redaction of Text and Diagrams** - Covering text, charts, tables, or diagrams with black rectangles, or highlighting text in black, is a common and effective means of redaction for hardcopy printed materials. It is not effective, in general, for computer documents distributed across computer networks (i.e. in "softcopy" format). The most common mistake is covering text with black.
2. **Redaction of Images** - Covering up parts of an image with separate graphics such as black rectangles, or making images "unreadable" by reducing their size, has also been used for redaction of hardcopy printed materials. It is generally not effective for computer documents distributed in softcopy form.
3. **Metadata and Document Properties** - In addition to the visible content of a document, most office tools, such as MS Word, contain substantial hidden information about the document. This information is often as sensitive as the original document, and its presence in downgraded or sanitized documents has historically led to compromise.

(U) How to avoid this problem? If your document is all text and it would be easy to reformat it in PDF, you could just copy the text from Word into a type of file that will not carry over the formatting information (e.g. an email message or a .txt file), and THEN copy it into a PDF file (and reformat it as necessary).

(U) If the above tactic is not feasible (perhaps because the file contains graphics, or is extremely lengthy and reformatting it in PDF would be too time-consuming), there are instructions in the IAD report on how to convert the information from Word to PDF safely. Again, here's the link:

Redacting With Confidence

(U) We are today dependent on information technology that encompasses not-so-obvious security gaps that could possibly represent a sizable security risk. The tools we use to conduct our day-to-day business may have important potential OPSEC consequences for those not familiar with the capabilities of modern software. Until we have information tools that warn employees of unauthorized information disclosure, OPSEC will remain the responsibility of each employee. So, before you transmit (E-Mail, FTP, etc.) or copy (via any media) that file, check to make sure that the information you provide is authorized for the intended recipient.

"(U//FOUO) SIDtoday articles may not be republished or reposted outside NSANet without the consent of S0121 (DL sid_comms)."