

White House Accidentally Exposes Data in PDF File

"In the latest error, the White House posted a copy of President Bush's "Plan for Victory in Iraq," the heart of his speech last week at the Naval Academy. But the Adobe portable document format file on the Web site also contained the hidden name of the original author of the document: Peter Feaver, a Duke University political science professor who joined the National Security Council staff last June as a special adviser." Patience Wait, *GCN Magazine*.

http://www.gcn.com/vol1_no1/daily-updates/37688-1.htm

We used Document Detective to take an even closer look at the President's victory plan. We discovered that the original document was a Word document, and that Tracked Changes had been enabled through the Ad Hoc Review process. Mr. David Sherzer from the Executive Office of the President (EOP) has the distinction of being the last person to email the document in its original Word format using Outlook before it was converted into a PDF document.

Beginning with Office XP, Microsoft enhanced the Tracked Changes feature with a new feature called an Ad Hoc Review. When the Ad Hoc Review is enabled, special identifiers are embedded into the Custom Document Properties section of the Meta data that allow the Microsoft application to track the Review's process. This information includes the name, email address and subject of the email used to send the document for review. A unique Review Cycle Identifier (RCID) is also added so that the file can be identified even if the reviewer changes the file name. If the reviewer sends it to another reviewer, then a second RCID is added. The review cycle can be extended to any number of reviewers, but only the last two RCIDs are included in the document. The Ad Hoc Review Cycle must be reversed to recover the other RCIDs. The following information was recovered from the President's victory plan:

```
pdfx: _EmailSubject>Iraq Strategy Final Draft!  
pdfx: _AuthorEmailDisplayName>Sherzer, David  
pdfx: _AuthorEmail='David_Sherzer@who.eop.gov'  
pdfx: _AdHocReviewCycleID>- 1523295987  
pdfx: _PreviousAdHocReviewCycleID>-1826856961  
pdfx:Company>EOP
```

The real danger of the Ad Hoc Review is that Microsoft automatically enables it without warning the user! Any time a Word, PowerPoint or Excel document is sent via email using Outlook 2002 (Office XP), the Ad Hoc Review is enabled by default. This is the probable cause of the recent rash of Tracked Changes incidents reported in the press.

We also discovered that the document was prepared on a classified computer with access to a classified network. This is a dangerous security practice that is discouraged or prohibited by most Government organizations. The flag images that appear at the top of all the pages include a Uniform Resource Identifier (URI) that is shown below:

<http://www.cia.sgov.gov/cwfb/cfactbook/flags/iz-flag.html>

Sgov.gov is a SECRET network used by non-DOD Agencies (like the Department of Homeland Security (DHS)) and the Intelligence Community for sharing National Security Information. This domain has access to the SECRET Internet Protocol Routed Network (SIPRNet). The URI was embedded inadvertently when the image was inserted into the document. For this URI to be embedded in the document, it had to be on a computer that had access to the sgov.gov network.

The sgov.gov URI can be found using a text or binary editor, but there is an easier way. Open the document in the Acrobat Reader and press the tab key. This action will take you to the first header

containing the Iraqi flag and highlight the flag with a small box. Hover the mouse over the highlighted image, and the URI will appear in the tool tip balloon.

The President's victory plan for Iraq clearly demonstrates that the conventional strategy of using a PDF document to "sanitize" documents does not work. Documents should be sanitized using a tool like Document Detective before the document is ported to a PDF document. The document should be checked again after it has been ported to ensure the porting process did not add undesirable information to the new PDF document.

Democratic National Committee Exposed as Author

An unsigned Word document critical of Supreme Court nominee, judge Samuel A. Alito, Jr., being circulated as independent was actually written by the Democratic National Committee. Meta data within the document revealed its origins. While this was not a surprising revelation, it is yet another reminder that hidden data in electronic documents is common.

<http://www.nytimes.com/2005/11/07/business/07link.html?ex=1135486800&en=918ea797afbb98b1&ei=5070>

Microsoft Marketing Memo Written on a MAC

This information is nearly nine months old, but it just came to our attention. A marketing memo touting the superiority of Microsoft Office for creating professional looking documents that can be shared seamlessly with others was actually a PDF document created using QuarkExpress on a Macintosh. This information is recorded in the Meta data that appears near the end of the document. The Meta data is not compressed, so it can be viewed with any text editor that can load the file. John Lederer quipped in an ABA discussion group that Microsoft had used the MAC produced PDF instead of using Microsoft software and file formats, because they needed a professional looking document that could be seamlessly exchanged with others.

<http://mail.abanet.org/scripts/wa.exe?A2=ind0403&L=lawtech&D=1&O=D&F=P&P=33500>

Document Detective PDF Parser Developments

SRS Technologies released Document Detective version 1.1 with our first generation PDF analyzer on 28 November 2005. The capabilities of this new analyzer were discussed in our 31 October 2005 newsletter. Since that time, we have made some adjustments that allow us to recover deleted pages and text from a PDF document. The Adobe PDF Editor lets users insert or delete text, images, pages, etc., but deleting information does not remove it from the document. Each time the document is saved, a new Page Tree is inserted that references all of the objects currently in use. "Deleted" objects are still in the file, but they are not referenced by the new Page Tree. Document Detective recovers all of the objects from the file, and any objects that are not used to create the current document pages are listed in the Objects folder for the reviewer's consideration. If the compression algorithm is known, then the objects in the Objects folder are decompressed, and decompressed objects that are recognized as encoded text are decoded. As a result, the deleted text and other text objects that are not yet processed by Document Detective (such as Annotations) will be exposed for review.

While testing the new modifications, we noticed something unusual about text edited with the Adobe editor. We could actually track the individual keystrokes being typed! We could see spelling errors being corrected by typing the backspace followed by the corrected text. We are still looking into this unexpected result, and we will report our findings in a future newsletter.

It is now clear that Adobe PDF documents are not safe documents! You should be aware that porting a document to PDF does not sanitize the document. Some hidden data from the original document will port to the new PDF document, and there will be digital fingerprints that could reveal the origins of the

document. Editing PDF documents creates additional security issues and should be avoided unless absolutely necessary.

SRS Technologies is committed to continuing research and product development. While no process or procedure is perfect, we can assure you that Document Detective is the best available solution for reviewing and sanitizing electronic documents, especially when National Security Information is at stake.

Please feel free to forward this message to anyone you believe would benefit from this information.

This newsletter is dedicated to raising awareness about the Desktop Publishing Threat to Information Security and to disseminating information that will help mitigate these risks. Our newsletter is published aperiodically when there is news to report. We will not clutter your inbox with idle chatter when there is nothing significant or useful to say. This newsletter is available free of charge. Subscriptions are at the discretion of SRS Technologies.

Suggestions and comments are always welcome. If you do not wish to receive this newsletter, please e-mail us and we will gladly remove you from the list. All EDS related correspondence should be sent to rhackett@stg.srs.com

For a list of articles and information regarding EDS, please visit our web page:
<http://www.stg.srs.com/eds/>

For information about Document Detective electronic document review and sanitization toolkit, please visit our web page:
<http://www.stg.srs.com/eds/docdet/>

To subscribe or unsubscribe from this newsletter, send email to:
rhackett@stg.srs.com

SRS Technologies
Systems Technology Group
500 Discovery Drive
Huntsville, AL 35806
Copyright 2006 SRS Technologies