



IMAGE &amp; DATA MANAGER

## The double-edged sword

### Portable document format

**P**DF can be described as the backbone to a complete document management strategy. It is searchable, stable, secure and can be redacted. With the PDF/A standard now an internationally accepted under ISO 19005, it will also be viewable in the future. Yet in Australia, we have not reached the broad level of understanding the U.S. has of PDF. IDM looks at why enterprise-wide policy on documents leaving the corporate firewall has eluded many Australian organisations.

#### Redaction

Every boss dreads the day that trade secrets or financial data are leaked. Most attention in recent times have focused on the dangers presented by mobile storage devices and laptops, however a more benign, but no less damaging threat, lies in information inadvertently leaked via documents sent beyond the company firewall.

Redaction is one function that PDF vendors are beginning to offer to help manage electronic documents sent outside the company. Adobe offers this as an add-on for \$150 while Nuance this month signed a deal with Appligent to offer a plug-in to PDF Converter Professional v4.2 which Nuance's Mark Austin says is available for \$195.

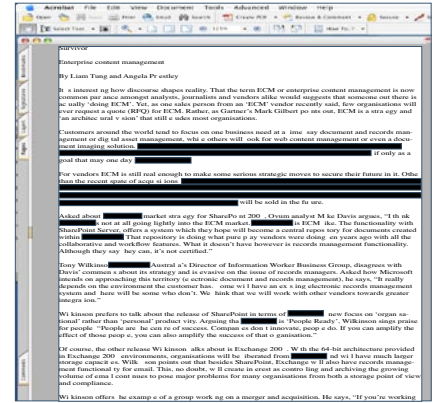
David Woolstencroft, managing director of Australian software developer DocsCorp, points out this has been a standard feature of pdfDocs Desktop for a number of years. Woolstencroft says the latest version, pdfDocs Desktop 1.9, can also scan from paper for redaction.

Prior to the rise of PDF-based redaction tools, you could simply blacken or literally cut text out from a paper document, then photocopy or scan the redacted version for public release. It is a laborious but easily understood task. The problem in the digital world is that people can apply similar processes (or habits) under a different reality. It is easy to fall into the "Seeing is believing" mentality, yet when this is applied to electronic documents – email, PDF, Word, Notes etc – it will likely land you in trouble.

In 2005 the Pentagon posted a document on its website in response to allegations that US troops in Iraq had shot Italian secret agent Nicola Calipari. The PDF, as far as the Pentagon was concerned, had been redacted yet within hours an inquisitive Italian journalist, aided by a German 'hacker', had uncovered the name 'Specialist Mario Lozano', further inflaming an already tense situation.

A similar mishap occurred in Australia. In late 2005 Westpac's chief executive officer David Morgan had little time to bask in the glory Westpac's record \$2.8 billion end of year profit result; a staff member had accidentally released profit results to analysts before lodging them with the Australian Stock Exchange. Consequently Westpac was forced to stop trading for two days. The document had been converted to PDF and emailed under the belief that certain elements of the document had been redacted. Similar to the Pentagon's error, the text was only obscured.

Both these instances are the results of users only 'masking' the text rather than



redacting it. As Woolstencroft points out, "Masking is not redacting. Masking text with opaque objects, as some have discovered to their embarrassment in the US military and government, is not a foolproof method of redaction. Saving the PDF document back into Word for example, displays the redacted text."

These examples raise questions about process and user's comprehension of the relationship between PDF and electronic documents. So a question the person whose head is on the chopping block should ask is: Do my staff understand the underlying technology enough to ensure that I will never face this situation?

This can only be resolved through correct training and a deeper understanding of the benefits and risks associated with its use.

But these above instances are only the ones that make headlines. All government agencies have staff that face freedom of information requests. Under FOI laws, anyone can request documents and the FOI officer responsible for redaction needs to ensure the document complies with relevant exemptions. It is a serious problem if a trade secret or third party's name is released in that document, but the person who actually redacts the document faces a different question: how long will it take me to complete redacting this 250 page file? Multiply this by ten and you have your average workload for many FOI officers. The benefits are lower labour costs involved in complying with FOI exemptions and, assuming the original document is in electronic format, the need for paper is completely eliminated.

**KNAPPCOMM**

 Knapp Communications. PO Box 576, Strawberry Hills 2012.  
 Ph: +61 2 93182644. Fax: +61 2 93104608. info@knapp.com.au

 THIS ARTICLE WAS REPRINTED FROM  
 IMAGE & DATA MANAGER,

JANUARY / FEBRUARY 2006, PP26-27.

© Copyright Knapp Communications. All rights reserved.

CONDITIONS OF USE This document is for use only by the purchaser and cannot be altered in any way. It is not to be sold to or used by any other third parties, vendors, distributors or retail outlets for any marketing purpose and can only be used for marketing by the purchaser. Any infringement of these conditions will be considered a violation of the publisher's copyright control.

**IDM**  
 IMAGE & DATA MANAGER

As one FOI officer pointed out, paper-based redaction often requires two rounds of photocopying because text is still viewable through black-ink markers.

### Controlling metadata

The other major concern is metadata and it is here where Mr Woolstencroft argues pdfDocs offers a major advantage. That is, providing an affordable means to automate processes for documents being sent beyond the corporate firewall.

Lurking beneath the visible content of any file is metadata, says Mr Woolstencroft. He reminds us that this “consists of who wrote the document, when it was edited as well as the changes that were made to the document. Metadata is dangerous for all these reasons. When you convert Word documents to PDF, this information is not carried over to the PDF file. In this respect, PDF is a safer format to distribute outside the company firewall.”

So, besides watching every email that leaves the firewall, what exactly should you do to ensure this never happens?

Austin suggests looking at your organisation's unique needs. “What you should do is analyse the requirements of your situation and implement appropriate procedures for yourself and your organization (or get competent advice). Users should only use PDF creation tools that specifically support redaction. If you don't have such a tool, the work-around is likely to be along these lines: Save the document in an image format such as TIFF that removes all text from the document.” After this TIFFs can be converted into an image PDF using a PDF tool if it is required. “Be aware that metadata may also be stored in the document and that also needs to be removed if it contains sensitive information.”

The approach Woolstencroft suggests is different. He argues that pdfDocs has been created to allow organisations to easily apply enterprise-wide policy on documents leaving its boundaries. He says, “pdfDocs Desktop ensures that MS Office attachments are automatically converted to PDF attachments once the Send button is clicked. The rules

associated with email integration can be set to convert all documents emailed to PDF, or only ones that are going outside the company firewall.”

He warns though that “Redaction is a manual process that must be done by the document author. Redaction, metadata removal and a rules-based email policy company-wide that converts all outgoing MS email attachments to PDF provides a great degree of security to companies large and small.”

So have Australian organisations taken an enterprise-wide approach to documents leaving the company? Woolstencroft says, overall no. The reason? In one word, Adobe. “When large enterprises want to standardise the way documents leave the company firewall, they simply can't afford to with Adobe because it's simply too expensive.”

Due to its global focus on the enterprise market and partnerships with document management vendors such as Hummingbird, Interwoven, Tower and Worldox, Woolstencroft says DocsCorp continues to develop software targeting process efficiencies designed to make policy easier. Version 1.9 of pdfDocs, he says, maintains this focus.

### Compliance and collaboration

The market for PDF in the US is very different compared with Australia. Mr Woolstencroft says the US is far more accepting of PDF as an essential tool.

Woolstencroft argues that the take up in Australia for PDF has been led by more conservative markets such as accounting where there are tighter regulations and users face demanding workloads. These conditions demand process efficiencies in converting documents into PDF format.

He says, “For many corporate organisations in Australia, PDF is seen as a nice thing to have but really it hasn't made it through to the extent of the Americans. The U.S. is a more mature market that is looking for ways to squeeze out extra productivity. Their knowledge of PDF and electronic document management systems is more mature so they have looked at the product range and incremental

changes to increase productivity.”

Admittedly Sarbanes-Oxley and tighter, nationwide regulations have helped drive the acceptance of PDF in the US however Woolstencroft also points to a stronger culture of collaboration. With a higher frequency of inter and intra organisation document transmission, they demand both efficiency and a stable document format. The U.S. has also made PDF an acceptable format to supply court documents in; to help lawyers comply with court requirements for e-filing, Mr Woolstencroft points to the ‘split feature’ included in version 1.9. This automatically splits PDF files based on file size or page numbers. Again, DocsCorp's focus on process efficiencies at the enterprise level has been designed to offer the benefits of PDF coupled with greater control under demanding circumstances.

This focus appears to be working for DocsCorp on the international scene. Its recent wins include Wendys and the Federal Reserve in New York, as well as wins in the UK and major Scandinavian legal firms. So much so that DocsCorp has recently opened its first office in Pittsburgh and will soon open others in Chicago and New York. There, DocsCorp is able to play a critical role in helping Worldox, Interwoven, Hummingbird and Tower deliver a complete document management solution.

Austin agrees that the U.S. is a more sophisticated market. “North America leads the way in sophisticated use of the PDF standard. However, redaction is increasingly a requirement for the legal sector in Australia, particularly with consulting firms and sensitive government areas. Redaction has been increasingly requested as the use of PDF has grown.”

It will take time for Australia to be on par with the US in terms of PDF. But as we gear up for a world that demands and fosters greater collaboration (eh hem, Microsoft Office and SharePoint 2007), it's time we took a leaf from the Americans and ask ourselves how we protect information leaving the firewall if for no other reason than to prepare ourselves for the future. **IDM**

**KNAPP COMM**

Knapp Communications. PO Box 576, Strawberry Hills 2012.  
Ph: +61 2 93182644. Fax: +61 2 93104608. info@knapp.com.au

THIS ARTICLE WAS REPRINTED FROM  
IMAGE & DATA MANAGER,

JANUARY / FEBRUARY 2006, PP26-27.

© Copyright Knapp Communications. All rights reserved.

CONDITIONS OF USE This document is for use only by the purchaser and cannot be altered in any way. It is not to be sold to or used by any other third parties, vendors, distributors or retail outlets for any marketing purpose and can only be used for marketing by the purchaser. Any infringement of these conditions will be considered a violation of the publisher's copyright control.

**IDM**  
IMAGE & DATA MANAGER