DOCUMENT SECURITY

ABOUT THE AUTHOR

Ronald L. Mendell holds a Master of Science degree in Network Security from Capitol College in Laurel, Maryland. He also holds the Certified Information Systems Security Professional (CISSP) designation. He has also held the Certified Legal Investigator (CLI) designation from the National Association of Legal Investigators (NALI). A member of the Information Systems Security Association (ISSA), he is a Distinguished Visiting Lecturer in Network and Computer Security at Our Lady of the Lake University in San Antonio, Texas. A writer specializing in investigative and security topics, he has numerous published articles in magazines such as *Security Management* and *The ISSA Journal* with subjects ranging from business intelligence to financial investigations to computer security. This is his fourth book for Charles C Thomas Publisher, Ltd. He works for a high-tech company in Austin, Texas.

DOCUMENT SECURITY

Protecting Physical and Electronic Content

By

RONALD L. MENDELL, MS, CISSP, CLI

Master of Science in Network Security
Certified Information Systems Secutity Professional
Certified Legal Investigator
Member of the International Systems Security Association (ISSA)
Member of High Technology Crime Investigation Association (HTCIA)

CHARLES C THOMAS • PUBLISHER, LTD.
Springfield • Illinois • U.S.A.

Published and Distributed Throughout the World by

CHARLES C THOMAS • PUBLISHER, LTD.
2600 South First Street
Springfield, Illinois 62704

ISBN 978-0-398-07766-2 (hard)
ISBN 978-0-398-07767-9 (paper)

Library of Congress Catalog Card Number: 2007015249

Printed in the United States of America
UB-R-3

Library of Congress Cataloging-in-Publication Data

Mendell, Ronald L.
Document security : protecting physical and electronic content / by Ronald L. Mendell.
p. cm.
ISBN 978-0-398-07766-2 (hard) -- ISBN 978-0-398-07767-9 (pbk.)
1. Computer security. 2. Computer networks--Security measures. I. Title.

QA76.9.A25M457 2007
005.8--dc22

2007015249

PREFACE

Several electronic layers exist in most documents, a fact overlooked by many writers. Probing these sublayers often reveals information not intended for release by the author. Documents in electronic formats create a "palimpsest" that even semiskilled investigators can probe for sensitive data.

Palimpsest seems like an exotic word. But literally, it means "scraped again" from the Greek word roots. In ancient and medieval Europe, writers often scraped off previous writing on a manuscript and wrote new text. (Writing media were in short supply and were expensive.) With modern forensic techniques like ultraviolet light and photography researchers uncover the original layer of writing.

Using computer forensic techniques, twenty-first century sleuths discover text and data in electronic documents thought erased by previous users. Modern electronic media are inherently palimpsestuous. Secrets become visible through metadata in documents, slack space in files, magnetic remanence, and other thorny ironies of information retention. They disclose information often, under the radar, by unintentionally making sensitive information Web-facing or not encrypting data on a laptop, which results in information leakage.

Overconfidence that one's sensitive data is not leaking through to the outside world will vex security professionals in the twenty-first century. Immense security resources go to prevent deliberate network intrusion. However, content security is not always on the forefront of security thinking. More information leaks out of organizations unintentionally than corporate America would like to think about. Many of the most recent headline-grabbers about security breaches involve documents or files leaked by a stolen laptop or by "misplaced" computer tapes or by being inadvertently Web-facing. The text identifies common pitfalls in document security and suggests remedies to prevent future headlines.

INTRODUCTION

The "hacker" culture dominated network security throughout the 1980s and 1990s. As the exploits of teenagers cracking into the systems of multibillion dollar corporations grew, basic countermeasures evolved to deal with the onslaught. As the twenty-first century arrived, the criminal sector caught on to the treasures lying in the data on those systems. While "hackers" have not disappeared, the dangerous attacks are now less thrill-motivated and more geared toward seizing valuable data.

Financially motivated crime continues to grow in cyberspace. The target is files or documents. Content, whether it be credit card numbers, social security numbers, banking information, customer lists, or trade secrets, has become "king." Some of the most notable headlines involve organizations losing databases, misplacing files or documents containing customer data, or having laptops stolen with, of course, confidential data on them.

Organized criminal rings target financial data online through a variety of schemes ranging from phishing to planting malicious code, such as Trojans, on PCs to simply researching public records available on the Web. Spies obtain proprietary data through finding Web-facing documents via search engines, and social engineering continues to trump the best of network security technology. Kevin Mitnick and Robert Schifreen acknowledge in their respective books, *The Art of Intrusion* and *Defeating the Hacker*, that social engineering often is the shortest and easiest route to most secrets.

In twenty-first century America, individuals and organizations leak information on a regular basis. In some cases, they hemorrhage data, albeit unintentionally. Protecting networks is essential, but due attention needs to go to protecting content, even when it is not residing just in electronic form on a network.

Information leakage or compromise happens in the following ways:

1. Web-facing documents contain sensitive or confidential data. Employees, however, place the documents on an "internal server," thinking the information will remain visible only within the internal network. Unfortunately, the information becomes visible to the external world through Internet access.
2. Documents undergo multiple drafts and then get sent to recipients in electronic form. Savvy readers can learn about the history of the document and even view redacted sections by accessing the metadata within the document.
3. Documents on laptops and PDAs containing sensitive data have no encryption protection, or they lack robust encryption protection. When the laptops and PDAs are lost or stolen, the critical data has little protection.
4. Storage media for documents in electronic format do not have proper markings as to content and sensitivity. Tracking procedures do not exist for the media. No encryption is in effect for the data. Such media are easily misplaced, lost, mislabeled, or stolen.
5. Documents, whether in paper, physical, or electronic format are not disposed of in a secure manner.
6. Reuse of electronic media occurs without following recommended secure procedures. Persons with a minimal understanding of computer forensics can read sensitive information remaining on the media.
7. Digital devices record all activity on the machine. Computer forensic examination recovers much of what the uninformed user thought he or she had deleted.
8. Web pages contain details about the hiring of technical staff, recent network infrastructure enhancements, and details about the enterprise's business organization. All of this available information aids corporate spies and hackers.
9. Disinformation on fraudulent Web sites compromise legitimate businesses' logos, branding, and services.
10. Credentials from business organizations can be easy to forge or to fake. These vulnerabilities permit fraud in gaining employment, in obtaining physical entry to the facilities, and in impersonating the business in the marketplace.

In other words, paying attention to documents and their content covers considerable security territory. Most of the leakage of sensitive information is not intentional. Workers and managers do not mean for it to happen. Often, the compromise of data arises from someone working extra hard. They take sensitive files home and before anyone realizes the problem the data becomes compromised. It is lost, stolen, or accidentally placed in the trash.

Thinking to help others, employees place information on the Web. When it is available online, information becomes easy to disseminate and to update. These advantages improve internal communication within an organization, but they also facilitate hacking and information theft against the organization.

The text strives to alert an audience of managers, security professionals, and workers who come in regular contact with sensitive information. Document security is not an accident. At any point in the life cycle of a document if it faces exposure to unauthorized eyes, compromise and loss of confidentiality occurs.

Recognition of how sensitive documents can violate the principle of confidentiality is the primary focus. Continuous protection requires understanding all of the possible avenues for compromise. Those avenues include the following:

A. Not understanding the information conveyed in metadata.
B. Not employing robust encryption protection.
C. Inadequate monitoring of business channels and subsequent filtering to reduce information leakage.
D. Inadequate erasure of magnetic media to reduce remanence.

Chapter 1 discusses metadata in documents. The most common metadata Microsoft Office documents are in the document properties section. The statistical information available there can reveal how long it took to create and to revise the document. In addition, previous revisions of the document may be discoverable. Paying attention to this issue can reduce unintentional release of sensitive information.

In Chapter 2 the text explores Web-facing documents and how search engines like Google® can uncover sensitive data in those documents. This is a widespread problem, and it requires constant attention by security to reduce or eliminate the exposure.

Business channels range from e-mail to instant messaging to FTP transfers. Chapter 3 discusses how filtering these channels is feasible with modern technology. However, the telephone and events like trade shows and professional meetings also provide business channels that are difficult to filter.

Chapter 4 covers the theft of digital devices such as personal data assistants (PDAs), laptops, and cellular telephones. These devices all contain documentary information. The chapter discusses the use of global tracking technologies and encryption to protect vital information from this growing problem.

Erasing most computer media does not completely remove the information. Special procedures are necessary to completely remove sensitive data. Chapter 5 discusses this issue and explains methods for disposal and reuse procedures.

In Chapter 6, paper and physical documents, such as information written on whiteboards or printed on boxes, pose unique control, disposal, and storage challenges. These documents bring the physical security force into the information security effort, if the organization uses the force properly. Protecting paper and physical documents forms the core of any document security program. Carelessness here is symptomatic overall of a weak information security effort.

Forensics involving computer-based documents looks at digital fragments on hard drives and on other computer media. These fragments tell a story about what a user thought was deleted or written over on the computer. Chapter 7 examines the whole issue of “slack space” on a computer and what security can do to make users aware that computers are the ultimate recording machines.

Chapter 8 continues the discussion by describing anti-forensics. These techniques minimize what forensic examination can uncover. Nothing is foolproof, but awareness goes a long way to preventing inadvertent passing of sensitive data on a data storage device.

Being deceived or fooled by documents is an important issue for security. Chapter 9 deals with the evaluation of online information. Bogus sites can imitate legitimate ones, and other Web sites can pass on disinformation to facilitate phishing and other scams. Learning to evaluate the validity and reliability of online information should be a part of the security training for all employees.

Chapter 10 discusses document forgeries. The increasing sophistication of desktop publishing programs, scanners, and printers means security has to be able to detect forged credentials and vital documents as a part of protecting an organization. Bogus documents necessary for securing employment continue to proliferate. In addition, academic and business records are also the subject of growing forgery trends.

The basic principles of information security as to documents require understanding the vulnerabilities that information faces. Upon creation, an electronic document may leave unintentional clues as to its content. Even if the main document remains secure, metadata about its contents may be elsewhere on the PC or PDA. Mirrored images may reside in swap files or in backup storage.

In storage, an electronic document may face surreptitious copying or alteration. If not properly classified, confidential electronic documents may encounter unauthorized eyes. Paper documents, due to them being commonplace in work areas, tend to be ignored as a security red flag. Those individuals, however, with a need to know, albeit not a necessarily authorized need to know, will haunt accumulation centers for documents to skim for information. Physical documents written on chalkboards and whiteboards often convey sensitive information in a completely innocuous way. Unless procedures exist to erase this information timely, unwanted eyes may get to study it.

Reusing electronic media has its own special dangers. A disk, a hard drive, a USB drive, or a backup tape that contained confidential data may end up being recycled for nonsensitive use. The remanence of sensitive information compromises the data to unauthorized users. Unless stringent procedures guard against sloppy reuse expect proprietary and confidential data to go walking out the door.

Lastly, destruction of confidential documents requires careful planning and thought, whether those documents are paper-based, physical, or on electronic media. It is a difficult argument to make that someone stole your trade secrets when that person was able to recover them from your unlocked dumpster.

CONTENTS

Page

DOCUMENT SECURITY

Chapter 1

METADATA

The Preface introduced the term, “palimpsest,” to describe the texture of electronic documents. Much like an ancient or a medieval parchment, a hidden layer exists below the surface text. With proper forensic techniques this substrate becomes visible. Paintings sometimes have a layer of a previous drawing or painting underneath what our eye perceives. Building on these analogies, we understand that electronic documents often have an unintentional subtext, which, if ignored, may result in the leakage of sensitive information.

In the BBC Web-based article of August 18, 2003, “The Hidden Dangers of Documents,” Mark Ward offers several insights into this unique vulnerability. First, documents with numerous revisions, especially if there are multiple collaborators, are prone to information leakage via metadata not being removed after the document’s drafting. (Metadata is information about the document itself: the authors, the number of revisions, the time required to produce the document, and so on. But most important, it includes text, tables, and graphics, the authors thought they obscured or deleted.) People do not realize that many word processing systems like MS Word® automatically record this production data and statistics. They fail to recognize that using the command to hide text or illustrations fails to prevent inquiring eyes from discovering the information later. Also, common techniques like whitening text or blackening a graphic or table often fall short in protecting sensitive data. (Numerous business and consumer software products, such as MS Office®, which include Excel® and PowerPoint®, possess this vulnerability.)

Second, the problem is widespread. Mark Ward cites a study by computer researcher, Simon Byers, where Byers gathered 100,000 Word documents from various Web sites. There was not a single document that did not contain some kind of hidden information. With this shocking evidence, the conclusion that metadata results in the significant leakage of sensitive, confidential, or embarrassing information in both government or business is an information security nightmare that rears its head every day.

Finally, Ward discusses several incidents of metadata telling more than the authors intended. In the United Kingdom, the publicized Iraq “dodgy dossier” unintentionally contained the names of civil servants who worked on it. In America, during the period of the Washington sniper attacks, the *Washington Post* published a letter sent to the police that included confidential names and addresses. Ward notes a case where an employment contract received by an applicant contained previous revisions. The applicant used that sensitive information to negotiate a better deal.

IMPLICATIONS

Why bother about metadata? If all that business and technical writers ever did was print out what they wrote into hard copy and distribute their work product on paper, metadata would not be an issue. Electronic documents allow information to pass rapidly across great distances, and they facilitate twenty-first century commerce. Storing electronic documents uses little physical space compared to paper, and these documents permit searching for the phrase or section heading on the tip of your tongue. In other words, electronic formats for information will continue to stay on the forefront of business and governmental communication.

Awareness of what lies in the sublayers of electronic documents is an important security concern for now and the future. Having an electronic document say more than the author intended is not difficult. Vigilance against these information leaks requires user education, and that education process involves recognizing the main avenues for metadata telling too much.

Begin educating users by explaining that all electronic documents possess properties. Those properties include statistics about the document: editing time, the number of pages, the number of paragraphs, file dates, and how many revisions. While at face value, these numbers appear innocuous. Imagine, however, if a writer bills eight hours to a client for a document where the metadata indicates total editing time was only two hours. True, the writer took into account time to research and to plan the project, but the metadata raises doubts in the client’s mind. Knowing the number of revisions may give clues about the difficulty and complexity in the document's composition. Again, claims of an arduous drafting process may be questioned if the statistics suggest a less difficult composition effort.

Other general properties provide the names of authors, collaborators, and author's comments about the document. Custom properties include the document number, the group creating the document, the language used, the editor, and other facts about the text or file. Routing slips containing email addresses of reviewers or collaborators, when using the "File Send" function, also act as another "metadata trap." Document authors often forget that these internal properties exist as metadata behind or below the visible, overt data. While at face value, little harm results in most cases if a third party sees this data, yet in certain documents, one may not want outsiders to know all the collaborators on a project, or who reviewed the document prior to publication.

Most problems resulting from metadata information leakage arise when the user or author tries to hide portions of the document. Hiding equates to security in many writers' minds. But, security through obscurity often fails in practice. Common methods for hiding are:

1. Suppressing portions of text.
2. Hiding comments appended to a text, a spreadsheet, or a slide in a presentation.
3. Suppressing headers and footers or footnotes.
4. Whitening text on a white background.
5. Making text very small (usually on Web pages).
6. Hiding slides in a presentation.
7. Suppressing cells, data rows, and columns in a spreadsheet.
8. Suppressing embedded objects such as graphics or photographs in a document.
9. Suppressing hyperlinks or using text as an alias for the URL.
10. Redacting sensitive portions of a document by blackening or otherwise obscuring the area.

A majority of the items on the list (except for items 4, 5, and 10) occur during the drafting of the document and quickly get forgotten as being a hidden part of the final draft. If an author uses the "Track Changes" feature during the writing process, the history of changes to the document remain as a sublayer in the final draft. Many desktop suites like MS Office make the suppression of portions or a section in a document just a matter of a few keystrokes. Turning on the "Reveal Formatting" feature is one way of uncovering such efforts at obscurity. (Table 1.1 summarizes MS Office's common metadata weak points.)

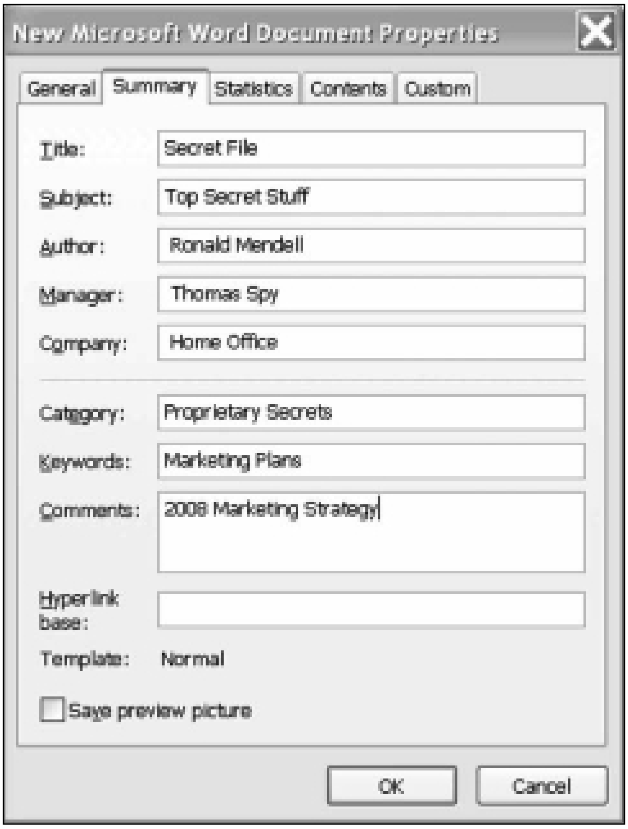
Metadata on Web documents is a very large problem. As Simon Byers's research indicates, a vast number of documents end up facing the Web in their original application's format. True, some metadata found in the hypertext markup language (HTML) enhances the value of the Web

Table 1-1: Some Metadata Types in Microsoft Office

Type	Description	Product
Comments	This element appears in document properties, in presentations, in text documents, and in spreadsheets.	MS Office
Custom Properties	Document number, group, language, editor, etc.	MS Office
Data Rows and Columns	These elements can be hidden in a spreadsheet.	MS Excel
Document Properties	General characteristics, summary, statistics, contents.	MS Office
Document Statistics	Editing time, number of: pages, paragraphs, lines, etc.	MS Office
Embedded Objects	Suppressed spreadsheets, graphics, etc.	MS Office
Fast Saves	Changes to the file appended to the document's end.	MS Word
File Dates	See Statistics.	MS Office
Footnotes	Suppressed in text.	MS Word
Headers and Footers	Suppressed in text.	MS Word
Hidden Cells	Suppressed cells in a spreadsheet.	MS Excel
Hidden Slides	Suppressed and forgotten in a final presentation draft	MS PowerPoint
Hidden Text	Hidden and often forgotten	MS Word primarily
Hyperlinks	Text used as hyperlink.	MS Office
Previous Versions	See Document Properties	MS Office
Routing Slips	File Send allows routing of documents to different email addresses	MS Office
Small Text	Very small text used on Web pages	MS Office
Track Changes	Done during drafting process, often forgotten about in final presentation copy.	MS Word
White Text	White font on white space to hide text.	MS Office

site. Such "tags" enable search engines to locate the pages more easily and authors use techniques such as very small text or whitened text hopefully to aid search engines while being less obvious. Unfortunately, search engines like Google permit keyword searches in specific formats like .doc, .xls, .ppt, and many more. For example, the inquiry '"marketing plan" filetype:doc' yields all MS Word documents containing the phrase "marketing plan." (See Chapter 2 for details about Google hacking.) Downloading the resulting document or documents as an MS Word file allows for the internal examination for any metadata not sanitized by the author. In addition, tools exist on the Internet, which permit the capture of an entire Web site. Then, one can examine the HTML source code looking for clues to sensitive data embedded in the Web documents.

Speaking about metadata in documents, however useful, does not replace seeing a few examples. "Properties," as seen in Figure 1.1, has



Title: Secret File

Subject: Top Secret Stuff

Author: Ronald Mendell

Manager: Thomas Spy

Company: Home Office

Category: Proprietary Secrets

Keywords: Marketing Plans

Comments: 2008 Marketing Strategy

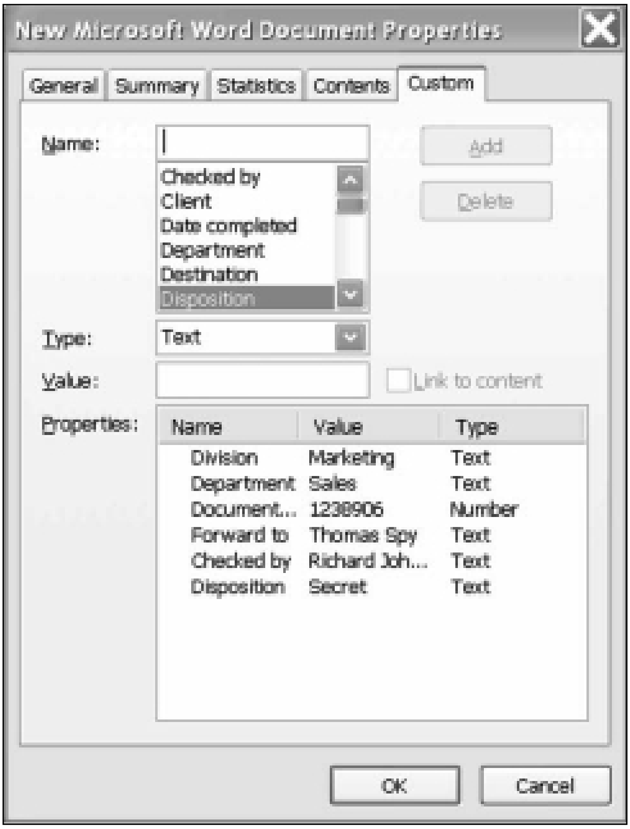
Hyperlink base:

Template: Normal

□ Save preview picture

Figure 1.1: Common Properties

information tabs for “General,” “Summary,” “Statistics,” “Contents,” and a gateway tab to “Custom” properties. The General section tab includes type of document, location on the computer, size of file, MS-DOS name for file, dates and times for created, modified, and accessed, and the file attributes (whether the archive bit is active). Summary’s tab includes title, subject matter, author, manager, company, category, keywords, comments, main hyperlink, and the template used. Created, modified, and accessed dates and times, “last saved by” (by which user or author), revision number, total editing time, and the number of pages, paragraphs, and characters are all under the Statistics tab. The Contents tab has a sectional outline of the document. As indicated in Figure 1.2, Custom properties allow the author or user to select from a list of additional properties ranging from “Checked by” (who reviewed the document for accuracy) to “Typist.” For every property selected, the



Name:

Type: Text

Value:

□ Link to content

Properties:

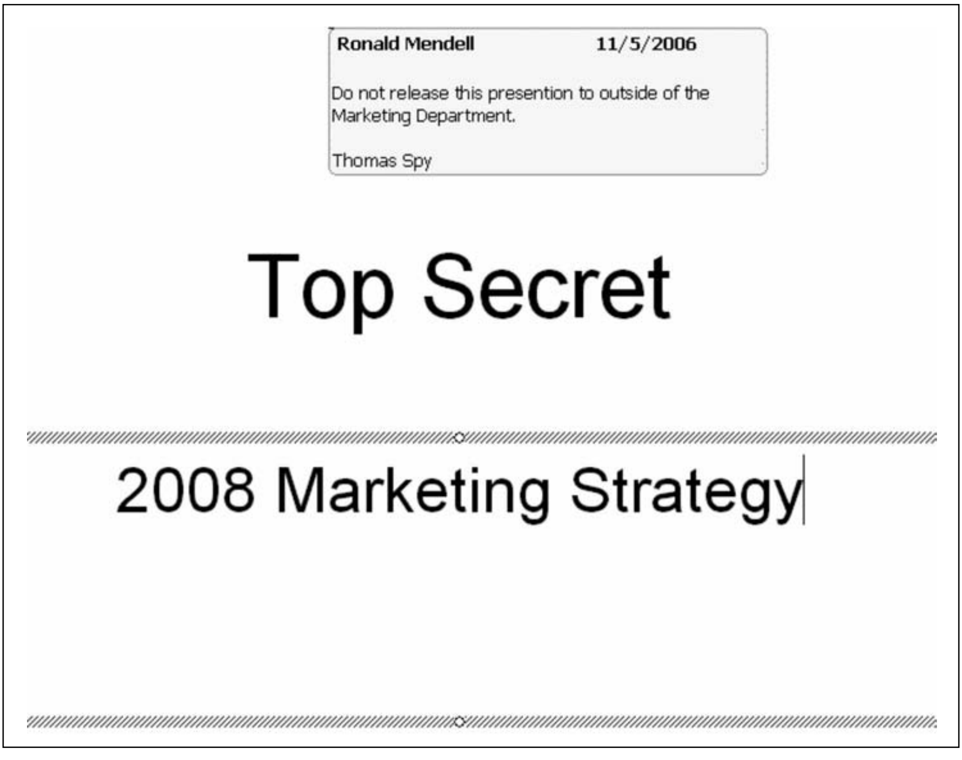
Name	Value	Type
Division	Marketing	Text
Department	Sales	Text
Document...	1238906	Number
Forward to	Thomas Spy	Text
Checked by	Richard Joh...	Text
Disposition	Secret	Text

Figure 1.2: Custom Properties

author adds a value for the field; for example, “Typist” could be “Mary Jones.”

The metadata issues go beyond MS Word. In Figure 1.3, the comments on a PowerPoint slide become apparent. Hiding a cell in an Excel spreadsheet is visible in Figure 1.4. These problems emerge when authors publish a document but forget that such metadata exists, or they think simple hiding methods suffice to cover up facts about the document they do not wish to be made public.

Obviously, a solution to this information leakage challenge takes two forms. First, the original document can undergo a sanitizing process to remove all the undesirable metadata. This approach requires thoroughness and patience, along with careful proofreading once the process is complete. The alternative method involves placing the original content, which the author wants outsiders to see, into another, secondary document that does not permit the transference of the metadata. Either methodology results in a metadata “safe” document, provided that the



Ronald Mendell 11/5/2006

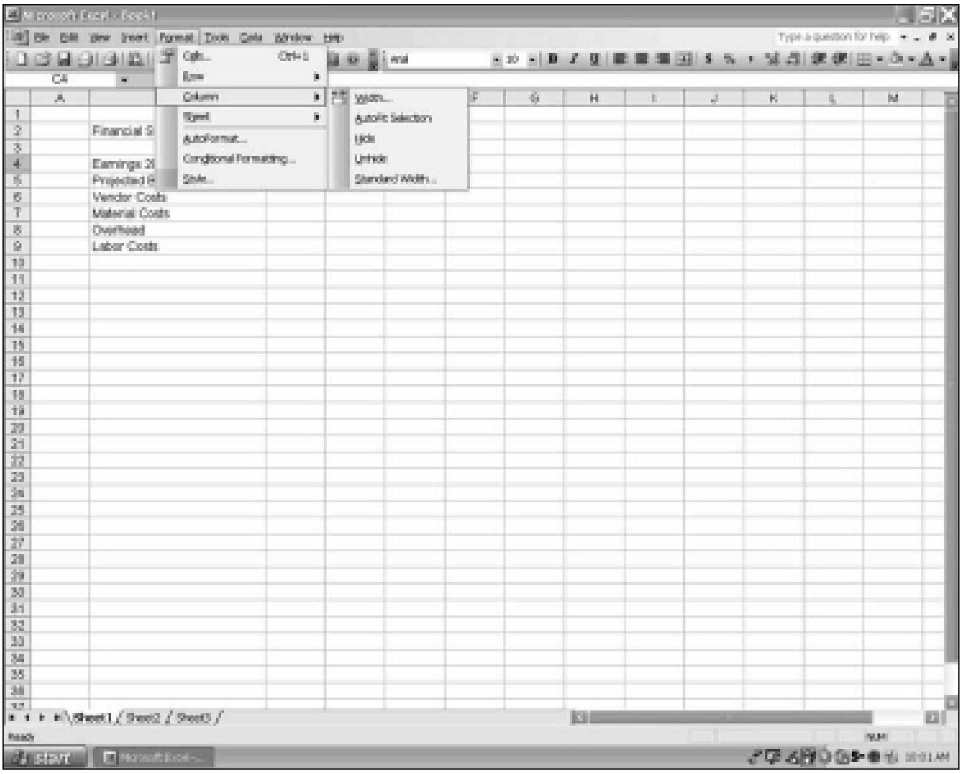
Do not release this presention to outside of the Marketing Department.

Thomas Spy

Top Secret

2008 Marketing Strategy

Figure 1.3: Comment on a PowerPoint Slide



Financial S
Earnings 2(
Projected E
Vendor Costs
Material Costs
Overhead
Labor Costs

Figure 1.4: Hiding a Cell in Excel

author follows all the correct procedures. In the next section, we will examine countermeasures for cleansing and transferring content.

METADATA COUNTERMEASURES

Table 1.2 summarizes both effective and ineffective countermeasures for dealing with metadata. Covering text or diagrams or reducing images usually does not work against a savvy reader. Sanitizing a document, however, through a series of steps, we will look at shortly. The four other effective approaches are as follows:

1. Use the Microsoft add-in program, "Remove Hidden Data" (RHD).
2. Use the MS Office document's drop-down menu "Tools/Options/Security/Privacy Options" to alert the author or user to metadata problems in the document.

Table 1-2: Controlling Metadata

Technique	Description	Effectiveness
Covering text or diagrams	Blackening or whitening with a rectangle sensitive data or text in a document.	Usually not effective. Removing covering in the copy is not difficult.
Reducing images	Reduce the image to point where it is not visible.	Usually not effective. Reveal markup commands in the application can uncover such images.
Sanitizing a document before copying	See the steps in Table 1-3.	Quite effective.
Use "Remove Hidden Data" add-in program	Available from Microsoft. http://support.microsoft.com/kb/834427	Effective in most cases.
Use Tools/Options/ Security	Check the "Privacy Options" to warn about metadata.	Effective in alerting author or user to problems.
PDF Utility	Appligent has redaction utilities for PDF documents. http://www.appligent.com/products/product_families/redaction.php	Effective with PDF documents.
Antiword and Catdoc for Linux and Unix users	See these Antiword links: http://en.wikipedia.org/wiki/Antiword http://gnuwin.epfl.ch/apps/antiword/en/index.html For Catdoc: http://www.45.free.net/~vitus/software/catdoc/	Renders MS Office applications used in Linux or Unix environments into simple text files without metadata.

3. Employ Appligent's PDF utility to ensure unwanted metadata does not pass through to the published document from the PDF original.
4. Use Antiword or Catdoc for MS Word files on Linux and Unix machines.

Locate RHD's description on Microsoft's Knowledge Base (Reference 834427 at http://support.microsoft.com/kb/834427). This program works on individual files or on multiple files. Collaboration features like Track Changes, Comments, and Send for Review will not work after the user or author applies RHD. So, use RHD only after the drafting process is complete. Microsoft states that RHD can remove:

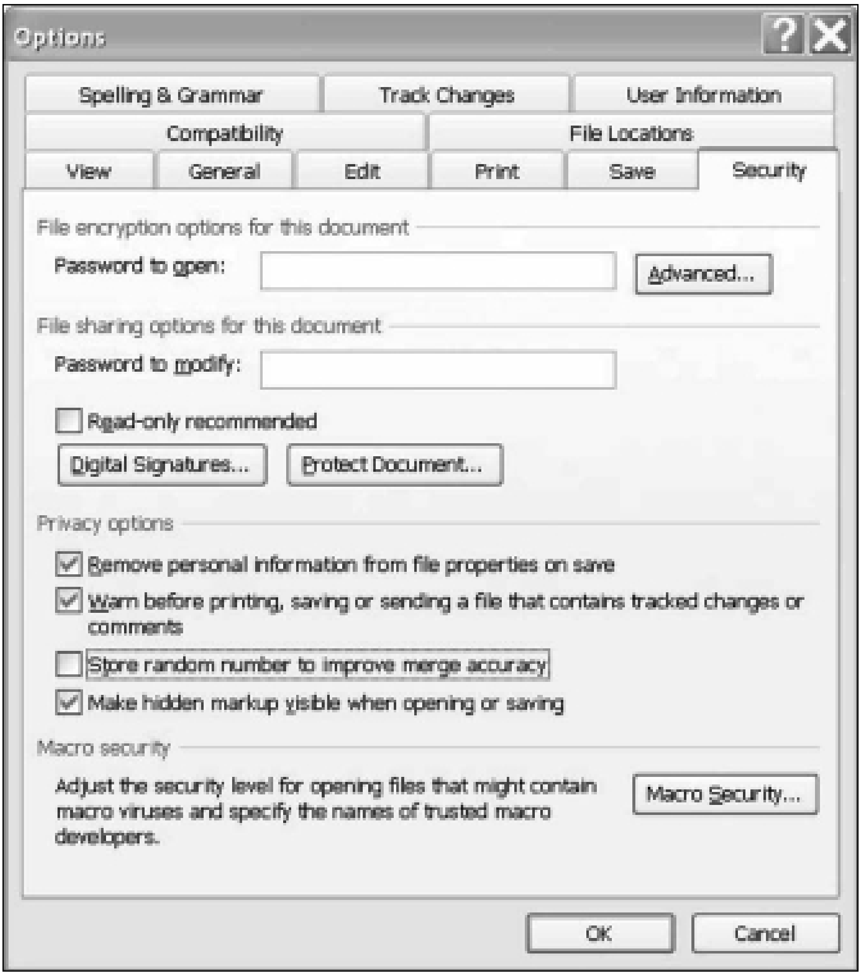
- Comments
- Previous Authors and Editors
- User Name
- Personal Summary Information
- Revisions
- Deleted Text
- Visual Basic Macros
- Merge ID Numbers (See check box "C" below.)
- Routing Slips
- E-mail Headers
- Scenario Comments
- Unique Identifiers

Use the drop-down menu for "Tools" in the document's toolbar. Select "Options" and click on the tab "Security." (See Figure 1.5.) Under "Privacy Options" the following check boxes are available:

[A.] "Remove personal information from file properties on save"
[B.] "Warn before printing, saving or sending a file that contains tracked changes or comments"
[C.] "Store random number to improve merge accuracy"
[D.] "Make hidden markup visible when opening or saving"

These security options act as a first line of defense against metadata passing unnoticed into a published document. Items A, B, and D are self-explanatory and need to be checked so that they will be active. Item C, if unchecked, will not store GUIDs (Generated Unique Identifiers [numbers]) when doing document merges. GUIDs, although useful in temporarily tracking merge documents, can identify the computer used to create the document if left stored on the machine. If you wish your computer to remain anonymous in the published document, uncheck this feature.

Appligent's family of PDF redaction tools is quite effective in removing metadata from PDF files. (PDF files are a type of universal document format that permits the documents being read on any computer that has an Adobe® PDF document reader on the machine.) Redax 4.0 automatically removes any document metadata and even marks up sensitive visible data like social security numbers, zip codes, and telephone numbers. Unlike conventional blackening or whitening of sensitive text,



File encryption options for this document

Password to open:

File sharing options for this document

Password to modify:

□ Read-only recommended

Privacy options

☑ Remove personal information from file properties on save

☑ Warn before printing, saving or sending a file that contains tracked changes or comments

□ Store random number to improve merge accuracy

☑ Make hidden markup visible when opening or saving

Macro security

Adjust the security level for opening files that might contain macro viruses and specify the names of trusted macro developers.

Figure 1.5: Tools/Options/Security

Redax's process is inexpugnable. Someone cannot just change the covering color to see the underlying text. So, the tool is excellent for redacting documents pursuant to Freedom of Information Act (FOIA) or Open Records requests. Appligent has an excellent white paper on their site, "The Case for Content Security," at http://www.appligent.com/docs/tech/contentSecurity.pdf if one desires more background on the redaction process.

Antiword and Catdoc offer some relief to Linux and Unix users. If they use Microsoft Office applications, these tools render documents into text free of metadata. Antiword is a free MS Word reader that has versions for Linux, RISC OS, Free BSD, Be OS, Mac OS X, and various flavors of Unix systems. Catdoc is an MS Word reader that extracts out text from the formatted MS Word document. Its cousins, which are xls2csv and catppt, create the same capability for Excel and PowerPoint documents respectively.

Sanitizing a document through a series of manual steps rounds out the discussion of countermeasures. These steps are from a National Security Agency (NSA) publication, "Redacting with Confidence: How to Safely Publish Sanitized Reports Converted from Word to PDF." Protecting documents for dissemination from inadvertent metadata requires vigilance. Nothing can replace your own visual examining of the final public document. Careful review of a lengthy document may require several sets of eyes, and while initially this process may seem unduly tedious, remember that you and your team are being digital sleuths. Try thinking of the digital document as something to attack. View it as a spy would do. (See "Being a Metadata Sleuth" at the end of the chapter.)

Please refer to Table 1.3 during this commentary. The first step is to create a new copy of the document. Then make sure the "Track Changes" feature is turned off. (You do not want to add any more metadata to the document.) Review the document and delete sensitive data or content. Replace deleted items such as graphics, tables, paragraphs, and text boxes with rectangles containing meaningless data like 1's and 0's. (If pursuant to a FOIA or Open Records request, then do this procedure to show the items and areas redacted.) DO NOT simply cover the area by using a dark color or by whitening the text. DELETE all the text or graphics, and then replace the missing area with the meaningless data.

Review the redacted copy again for any possible oversights. Have other authorized persons double-check your work. Then, select all the document contents and paste them into a virgin, blank MS Word document. This step is very important. It minimizes the amount of metadata in the MS Word document that you plan to convert into a PDF document. Review the document again. Then, ensure that Adobe PDF conversion settings are correct by having unchecked the options: "Convert Document Information" and "Attach Source File." If the document passes all the inspections and these PDF conversion settings guarantee that metadata will not pass through, then convert the MS Word document to PDF format. Finally, do an inspection of the PDF file to make sure no undesired data is either viewable or searchable. Run some test searches within the PDF document using key words or phrases you

Table 1.3: Sanitizing an MS Word Document

No.	Instructions	Check
1	Create new copy of document.	
2	Turn off "Track Changes" in copy.	
3	Review copy and delete sensitive content.	
4	Replace deleted items such as graphics, tables, paragraphs, and text boxes with rectangles containing meaningless data. (If necessary to show items and areas redacted)	
5	Review redacted copy for errors and omissions. Then, select all the document contents and paste them into a virgin, blank MS Word document.	
6	Review the document again. Ensure that Adobe PDF conversion settings are correct by having unchecked options: "Convert Document Information" and "Attach Source File."	
7	Convert document to PDF format	
8	Review PDF document for any errors or omissions regarding undesired metadata.	
	Source: "Redacting with Confidence: How to Safely Publish Sanitized Reports Converted from Word to PDF", National Security Agency. http://www.fas.org/sgp/othergov/dod/nsa-redact.pdf	

redacted to make sure you have a clean public document for dissemination. Also inspect the "Properties" of the PDF file.

MICROSOFT'S ONLINE HELP WITH MS OFFICE METADATA

Microsoft's Knowledge Base (KB), its online encyclopedia of help and advisories for users, has several articles regarding eliminating metadata from various MS Office applications. (Go to http://support.microsoft.com/ and enter into the search box the article number desired.) Articles numbered 237361 and 290945 cover issues with MS Word. These articles begin with a general overview of the metadata items that can reside in an MS Word document, which we enumerated earlier in this chapter. What is particularly useful about the KB articles is that they explain how to remove each individual category of metadata. Users can go down a hyperlinked list and choose the particular category they wish to remove. If one is only interested in removing Personal Summary information, for example, the associated hyperlink quickly takes the user to the relevant portion of the article. This organization of the article permits quick resolution of issues when a user has only a certain key metadata element to remove.

Article 223789, "How to Minimize Metadata in Microsoft Excel Workbooks," again gives an overview of the possible metadata items or categories within an Excel document:

> The following are some examples of metadata that may be stored in your workbooks:
>
> - Your name
> - Your initials
> - Your company or organization name
> - The name of your computer
> - The name of the network server or hard disk where you saved the workbook
> - Other file properties and summary information
> - Non-visible portions of embedded OLE objects
> - Document revisions
> - Hidden text or cells
> - Personalized views
> - Comments

As with the MS Word articles, this article on Excel provides a hyperlinked list for each category of metadata to permit easy searching and resolution of the problem. Please note the unique issues identified for Excel such as hidden cells and personalized views. Make sure your document production staff understands that regardless of the data's format metadata is always an issue. True, MS Word receives a lot of coverage regarding metadata security, but those that work with spreadsheets need to keep in mind that large scale information leakage can occur here too. Excel spreadsheets deserve as much attention in purging metadata, prior to dissemination, as other MS Office applications.

Articles numbered 314797 and 314800 deal with removing metadata from PowerPoint documents. Again, Microsoft follows the same format as with other MS Office applications. Some of the hyperlinks for PowerPoint the Knowledge Base quotes include:

> How to Delete Your User Name from Your Programs
> How to Delete Personal Summary Information

How to Delete Personal Summary Information When You Are Connected to a Network
How to Delete Comments in a Presentation
How to Delete Information from Headers and Footers
How to Disable Fast Saves
How to Delete Hyperlinks from a Presentation
How to Delete Routing Slip Information from a Presentation
How to Delete Your Name from Visual Basic Code
How to Delete Visual Basic References to Other Files
How to Delete Network or Hard Disk Information from a Presentation
Embedded Objects in Presentations May Contain Metadata

Again, the user must understand that every electronic document transmitted to others has the potential for information leakage through metadata. Since PowerPoint presentations get e-mailed or sent via file transfer protocol (FTP) all over the world for meetings, conferences, seminars, and the like, special vigilance is necessary. Because PowerPoint has exceptional visual capabilities, one forgets it may contain hidden text that should not pass to outsiders.

Before the discussion moves on to digital sleuthing, an important question arises: what about other applications outside of the MS Office suite? How does one find information about addressing metadata issues in WordPerfect® and other suites? Again, the Web search engine is the security professional's best friend. A quick check under "WordPerfect Metadata" in Google at the time of the writing of this chapter produced numerous online references, including a PDF file from Corel (http://www.corel.com/content/pdf/wpo12/Minimizing_Metadata_In_WordPerfect12.pdf, "Minimizing Metadata in WordPerfect12"). Any application with a significant distribution will have something on the Web about contending with metadata. If online resources prove unsatisfactory, please contact the respective manufacturer through their Web page for assistance.

BEING A METADATA SLEUTH

Becoming a digital detective is one of the themes of this book. Looking below the surface appearance of an electronic document is something that an adversary will do with great care. Security professionals need to adopt the same attitude when checking documents for metadata. The first step in this process involves learning the vulnerabilities of the application that created the document. If an investigator finds a document on the Web in its original composing format such as MS Word or PowerPoint or WordPerfect, immediate alarms should go off. As we have seen earlier in the discussion, those formats are fine for document creation but not for publishing. Those formats usually carry unintended or forgotten metadata, or poorly redacted text and graphics. One establishes an evaluation list for examining the document by visiting the respective manufacturer's metadata Web site.

The general principles of sleuthing an electronic document follow from traditional observation skills. Go beyond what the document is trying to say. Understand what it is also trying not to say. Redaction is the ellipsis of sensitive information. What makes portions of a document sensitive varies from document to document. Perceive the document's theme or mission and then try to understand what the author would try to hide. Sensitive material falls into the following general categories:

1. Who created or collaborated on the document?
2. Who reviewed or approved the content?
3. The timeline of the document's creation or editing. How many times was it revised? How long did the editing process take?
4. Personal information such as telephone numbers, social security numbers, addresses, the names of persons guaranteed anonymity, and the like.
5. Legally sensitive information required by law to be kept confidential such as medical information or student records or employee information.
6. Author's comments regarding the text. Editorial comments.
7. Proprietary data or trade secrets.
8. Classified information or data that could help reveal classified information.
9. E-mail addresses or universal resource locator (URL) links (Web pages) that the author does not wish outsiders to know as being related to the content of the document.
10. Revision marks and information from "Tracked Changes."
11. Templates and old file versions.
12. Headers and footers that are hidden, and other hidden text.
13. Visual Basic® references to other files and embedded objects.

14. Errors or omissions within the document that give clues to sensitive information. (For example, deleting one personal identifier for an individual but forgetting another identifier in hidden text or existing as a caption for an illustration.)

When sleuthing a document, start with sections that are obviously redacted. If the author has darkened an area, change the covering color to a lighter one. You may be surprised to find that the underlying text or image becomes visible. Turn on the "Reveal Codes" or "Reveal Formatting" feature to see if any text has undergone whitening.

Activate the "Track Changes" or "Markup" feature to reveal any comments or tracked changes in the document. The sleuth can use this feature in conjunction with "Reveal Formatting" to discover embedded objects, hidden text, hidden headers and footers, and revision marks.

Most documents have easily viewed "Properties" by clicking on the "File" tab and then clicking on "Properties." You can also view "Custom" properties as an internal tab within "Properties." Again remarkably, many writers and editors fail to remove sensitive information from this collection point for metadata. The history of a document often lies here: revisions, edit time, and the identity of authors and collaborators.

Remember the things in documents beyond text that people hide: slides in presentations, cells in tables or spreadsheets, rows and columns in spreadsheets, charts, and illustrations. In Excel documents, a few simple menu commands reveal most secrets. Drop down the "View" menu and click on "Comments" to see all the hidden comments in the spreadsheet. The "View" menu also reveals "Headers and Footers" by clicking on the same. To uncover hidden rows or columns, use the "Format" drop-down menu and choose either "Row" or "Column" and click on "Unhide." For addition tips on locating hidden items in Excel, use the drop-down menu "Help" and search with the word "hiding."

In PowerPoint presentations, the drop-down menu "Slide Show" has a "Hide Slide" feature. To see a list of hidden slides, right click on any slide in a slide show and click on "Go to Slide." In the list of slides that appears all hidden slides will be identified. Show hidden comments or changes by using the "View" drop-down menu and click on "Markup." For addition tips on locating hidden items in PowerPoint, use the drop-down menu "Help" and search with the word "hiding."

As far as Web pages go, viewing the source code (HTML) in a browser is usually just a matter of selecting the "View" drop-down menu and clicking on "Source." If you want to examine an entire Web site, purchasing a Web site capturing program like Web Site Downloader will do the trick. This program, for example, permits various types of filtering when doing the capture onto your hard drive or onto a CD or DVD disk. Filtering allows selecting particular files or pages to capture if you do not wish to download the entire site. Either a full or partial capture permits later detailed analysis of the contents for sensitive metadata.

If you want to examine documents outside of their native application, using a HEX (hexadecimal) editor will prove effective. A good one is WinHex. Depending upon the version purchased, this tool can offer a disk editor, a RAM editor, the ability to view up to twenty different data types, and the ability to analyze and to compare files. The viewer in WinHex allows an investigator to see text in ASCII format (basic alphanumeric characters) while also seeing the corresponding hexadecimal code. When you want to see the actual data in a document at the lowest level, a HEX editor is an excellent tool. (See the Web site for WinHex at http://www.x-ways.net/winhex/.)

These basic techniques, if used consistently, will uncover most of the metadata that slips through into published electronic documents. Knowing what to look for is the first step in ensuring that your documents do not say more that what you want them to say. (See Table 1-4 for a summary of the sleuthing techniques.)

Table 1.4: Sleuthing for Metadata

Issue	Techniques	Comments
Uncovering colored text, graphics, or diagrams	User or author employs blackening or whitening with a rectangle to cover sensitive data or text in a document. • Change the covering color to a lighter one. • Turn on the “Reveal Codes” or “Reveal Formatting” feature to see if any text has undergone whitening.	Removing covering in the copy is not difficult.
Discovering reduced images	User or author reduces the image to point where it is not visible. • Revealing markup or formatting commands in the application can uncover such images.	Images usually are not difficult to find by revealing the formatting.
Discovering “Tracked Changes” and hidden formatting	User or author turns off the display of “Tracked Changes” and formatting codes in the final copy. • Activate the “Track Changes” or “Markup” feature. • Turn on “Reveal Formatting” to discover embedded objects, hidden text, hidden headers and footers, and revision marks.	Many authors forget that this metadata passes on into the published electronic document.
Uncovering the document’s properties.	User or author forgets or overlooks what is in “Properties.” • Click on the “File” tab and then click on “Properties.” • View “Custom” properties as an internal tab within “Properties.”	This can be a goldmine on the document’s history.
Revealing hidden information in Excel and PowerPoint documents	People hide slides in presentations, cells in tables or spreadsheets, rows and columns in spreadsheets, and charts and illustrations in both presentations and spreadsheets. • Use the drop-down menu “Help” and search with the word “hiding” to find all the methods for showing hidden data.	Assume that any spreadsheet or presentation has something suppressed.
Mining Web sites for documents and information	Web sites often reveal far more than the designers intended. • View the source code from the browser. • Download the entire Web site with capture software: http://www.web-site-downloader.com/entire/	Web sites can be a rich source of intelligence about a company or organization.
Examining documents	Examine documents outside of their native application by using a HEX editor.	Low-level viewer

Chapter 2

WEB-FACING DOCUMENTS

Web applications continue to grow in focus by the information security community. Port 80, which permits HTTP (hypertext transfer protocol) connections, is open on the perimeter of most networks that depend upon the Internet for commerce and for information flows. Hackers and crackers exploit this opening to leverage attacks against the network as a whole. Professional security testers using tools like WebInspect™ and AppScan® probe Web applications looking for holes in the defenses. Web Application Security, however, is not the subject of this chapter.

Instead, we will concentrate on documents, not applications. Very few of the sophisticated computer skills used by top-tier hackers are necessary to discover sensitive information when one focuses on finding documents. All that is required is knowing where and how to find such documents on the Internet. The techniques are simple. How these documents come to be exposed to the Web is the main issue this chapter explores.

When someone searches for information leaks via Web-facing documents, two different strategies present themselves. First, the researcher can focus on a particular Web site and try to glean as much information from that site as possible. Usually, this approach lends itself best when the researcher has a clear target. Gathering business intelligence on a competitor works well with this tactic, or, if someone is planning a broader attack on a specific target, this approach helps to build a comprehensive picture of the target's "information footprint."

The online researcher may not care about a specific target. Rather, the information category itself becomes the object of inquiry. If someone is looking for marketable personal data like credit card numbers, proprietary data such as company financials, or lists of customers or sales leads, any information that can be sold in cyberspace, this second approach makes sense. If one, for example, sells mailing lists, conducting searches for that pattern of information should produce sufficient "loot" to stock the database that ends up being sold to others.

All information has a certain pattern in its organization and content. A financial balance sheet of a business may appear in a word processor document or in a spreadsheet. Regardless of the application, however, the content of the information will contain certain words, phrases, symbols, formatting, and punctuation. The same principle applies to a police report, a medical record, a driver's license record, or to any of a myriad of documents used in commerce and in daily life. If one knows how to search for the pattern and the common formats where it is found, finding all sorts of information is not difficult, and sensitive data leaks through to the outside world through Web-facing documents by one of two means.

The first way for information leakage is the stand-alone sensitive document. Somehow, someone placed a document in a vulnerable place on a network where it faces the Web. The document *by itself* reveals the sensitive data an information predator is after. No other resources are necessary for the sensitive information to be compromised.

A more insidious threat is the posting of multiple documents that individually do not have sensitive data. When taken in *aggregation*, however, they build a picture regarding sensitive information. Building a dossier about an individual from multiple Internet sources is a common example of the aggregation technique, and it is difficult to protect against. We will visit the concept more as we go along.

The main treasures that farmers of the Web for sensitive documents seek include:

- Proprietary Data (Trade secrets, Research and Development data, Internal documents, and Production processes)
- Financial Data
- Marketable Personal Data (Personal identifiers)
- Marketing Plans
- Customer Lists
- Supplier and Vendor Information
- ITSEC Information (Network configurations)
- Databases

Proprietary data may exist in concentrated form in a single document. When such a document finds its way into a Web-facing portion of a network, then a profound breach of security has occurred. More often, though, proprietary data is diffuse. It leaks out in small portions here and there. A published paper in a professional journal that tells a bit too much, an employment ad detailing the skills needed for a technical job, a posting in a newsgroup asking for technical advice, and biographical article about a key researcher in the company, these documents all become cumulative in the story they tell. Each alone speaks softly, but together, they form a chorus providing deep insight. Aggregating these pieces of information creates knowledge about an organization's proprietary operations. Broad search engine techniques like "Google hacking" aid in the aggregation process. Developing appropriate search patterns requires knowledge of the industry or business and the associated terminology.

Financial data often gathers into concentrated form in balance sheets, financial reports, and forecasts. These documents do end up facing the Web usually through users' error. Business intelligence researchers that find them definitely have hit a gold mine. Such data can be also diffuse: found in business articles, in news accounts, in presentations before professional groups, and in filings with regulatory agencies. In searching for this information, the method can be either a Web site download or a broad Web search engine query. Aggregation works quite well when sources are varied and multiple. Patterns to look for in a search include financial terms, financial document headings, certain financial ratios, and dollar amounts.

Marketable personal data occurs in concentrated form and also tends to be scattered across multiple sources like resumes, public records, membership information for groups and associations, news accounts, and in personal postings like individual Web sites and "blogs." (A blog is an online form of personal journalism, an upscale diary for the public to read and comment upon.) Unfortunately for those concerned with privacy, many of these sources are available online, and so data aggregation is not difficult. Data patterns include names, addresses, dates of birth, social security numbers, telephone numbers, credit card numbers, and so on. These patterns are simple to search on the Web, and sometimes, handlers of sensitive documents post them in the wrong places leaving concentrated personal data exposed.

Marketing plans generally tend to be a stand-alone document. Like any business information in the twenty-first century, however, contents may leak out in bits and pieces in the variety of sources previously discussed.

In fact, Open Source Intelligence (OSI) offers the business intelligence analyst, investigative reporter, or private investigator a powerful, legal way to discover sensitive information on individuals, businesses, and organizations. OSI is the art and science of gathering diverse source into a coherent intelligence picture. (For more about OSI, see Ronald L. Mendell, "Intelligence Gathering for ITSEC Professionals," *The ISSA Journal*, December 2005.) Broad Web search engine queries are an effective way of doing OSI for marketing plans or data. Web site downloads also can uncover these documents. Typical search patterns include marketing terms, marketing jargon peculiar to the targeted enterprise, and document headings unique to a marketing plan or forecast.

Customer lists usually are stand-alone documents. Typical search patterns for them include names, addresses, and contact information. If the ownership of the list is not critical (not a specific target's list), then a broad Web search engine query can locate them across the Internet. If a specific target's customer list is sought after, then a Web site download from the target's Web-facing servers is in order. Aggregation from diverse sources is also possible if trying to build a list for a given target. This aggregating technique uses multiple sources like news accounts, public records, transaction data, and published reports.

Very similar in content to customer lists are lists of suppliers and vendors. This data can be aggregated from diverse business sources and public records as with customer lists. Again, a broad Web search engine query can locate either stand-alone documents or bits and pieces of information from diverse sources pertaining to a given target.

Information security (ITSEC) information contains data about the configuration of directories on network servers, FTP servers, and Web servers. Knowledge of the directory structure on Web-facing servers forms the basis for pattern searching. (See the "Google Hacking" section below for details.)

Databases contain a wide variety of sensitive data including the categories just discussed. Both broad Web search engine query methods and Web site downloads can facilitate access to databases. Search patterns depend upon the content of the database. Knowledge of the subject area is especially important in crafting queries.

Table 2.1 summarizes these primary targets of those researchers and analysts that mine information from the Web. Search engine techniques, which the text discusses in the next two sections, enable aggregation from a broad range of identified sources. The techniques also help identify

Table 2.1: Targeting Web-facing Documents

Targets	Stand-alone or Aggregate	Patterns of Data	Search Type
Proprietary Data (Trade secrets, Research and Development data, Production processes)	Aggregation quite effective	Business or industry terminology, technical terminology	Broad search useful in locating documents or in aggregating data
Financial Data	Usually stand-alone documents	Financial terms, certain number patterns and ratios	Either a broad Web search engine query or a Web site download can work.
Marketable Personal Data	Stand-alone lists and databases, also data aggregation can be quite effective	Names, addresses, DOBs, SSNs, telephone numbers, credit card numbers	Broad Web search engine queries
Marketing Plans	Generally, a stand-alone document, however, aggregation from diverse sources can work	Marketing concepts, language, and headings on marketing topics: "Expansion Plans, Strategy, Rollout Plans"	Broad Web search engine query can look for a specific target's documents; a Web site download also can be effective.
Customer Lists	Stand-alone document	Names, addresses, telephone numbers, customer number, other contact information	Broad Web search engine query effective
Supplier/Vendor Lists	Similar to Customer Lists, stand-alone document	Names, addresses, telephone numbers, vendor number, other contact information	Broad Web search engine query effective
IT Security Information (Configuration of Network Servers)	Stand-alone directories on servers	Directory tree structure (See "Google Hacking" section)	Broad Web search engine query effective
Databases	Usually stand-alone on a server	Depends upon the content of the database	Either a broad Web search engine query or a Web site download can work.

stand-alone documents on Web sites, which are rich in desired information and content. The appropriate Web site download software permits the downloading of these data goldmines.

How do these sensitive documents end up facing the Web? The answer lies in the lack of controls. Controls arise from a security policy put in place to address the problem. If an organization does not have the means to identify its information assets, then it cannot protect them. As difficult as the task is, the enumeration of sensitive information assets forms the basis of protecting them from theft, unauthorized copying, and compromise. The fluid nature of information commerce in the twenty-first century makes the challenge especially difficult.

Yet, clearly defined classification procedures for sensitive documents need to be in place even in the most fluid of environments. A lack of document classification procedures will result, otherwise, in serious information leakage problems. The dividing of the information and knowledge base of the organization into security zones minimizes the danger of leakage. Based upon the principle of "least privilege," security zones recognize that someone cannot create a sensitive document unless he or she has access to sensitive information. Divide the information in an organization into Public, Internal, Sensitive, Confidential, and High Security zones. Obviously, the creation of a document classified Confidential in the Public zone simply will not do. (Assigning relative values to the different classifications depends upon the organization's needs. See Tables 2.2a and 2.2b for suggestions regarding classification.)

Each employee or authorized user has assigned privileges to the zones needed for their job (least privilege). They must label a document according to the zone where the document creation occurred. Once labeled, a document may not pass to a zone of lesser security. Physical and information security zones mirror each other so equal protection extends to physical and electronic documents. Your IT network, for example, needs the same zone segregation as used with manual, physical systems like paper documents. Many documents will be classified as Public or Internal. Internal documents pose minimal harm to the organization if compromised. Security measures discourage their disclosure, but if Internal documents get compromised the result is not a disaster. Your security efforts in the first two zones strive to prevent unauthorized access to the inner and more sensitive zones.

Any documents intended to be Web-facing must reside in the Public zone. They must also be created in that zone. Documents in any of the higher zones cannot be Web-facing and servers directly accessible from the Web cannot be in these higher zones. Authors and users may be ignorant of what servers face the Web, so robust compartmentalization of

Table 2.2a: Security Zones for Documents (Lower levels)

Level	Impact if Compromised	Document Types	Protection Methods
Public	None	• Documents intended for public release • Documents or databases intentionally placed on the Web server for public access • Technical or business papers for public consumption • Speeches at public meetings.	Physical: (Control access from public areas into internal areas of the organization) IT Security: (Use a DMZ with IDS monitoring to isolate the Web segment from the rest of the network)
Internal	Some embarrassment	• Internal memos and e-mails of a non-sensitive nature • Intranet postings • Employee announcements and updates • Internal newsletter	Physical: (Controlled access with badges, security officers, CCTV, intrusion detection systems) IT Security: (Isolate the segment for Internal documents with internal firewalls and IDS monitoring)
Sensitive	Significant harm (Civil lawsuit)	• Clients' or customers' proprietary data • HIPAA or GLB protected information • Employee or student records	Robust physical and IT security. Internal card access to internal areas and IP address separation of sensitive network areas along with internal firewalls and IDS monitoring.

the Network into security zones is essential. Security auditing software checks electronic documents' labels to ensure they are in their respective zones. Physical audits and controls prevent paper documents of high sensitivity from leaving their respective security zones. Only documents classified as Public can make their way into mobile devices like laptops, PDAs, and computerized cellular telephones. If documents of higher classification must go mobile in these devices, stringent encryption and security measures must be in place. (The text will discuss more about these issues in Chapter 4.) It is extremely unwise, however, to permit Confidential and High Security documents from ever going "mobile," regardless of the security protections on the device. Any such movement must be considered very carefully by management.

Table 2.2b: Security Zones for Documents (Higher levels)

Level	Impact if Compromised	Document Types	Protection Methods
Confidential	Grave harm financially	• Mergers and Acquisitions information • Marketing plans • Trade secrets and other critical intellectual property • Customer lists and databases • Supplier/Vendor lists • Internal auditing papers	Robust physical and IT security. Internal card access to confidential areas and IP address separation of confidential network areas along with internal firewalls and IDS monitoring.
High Security	Criminal prosecution or serious endangerment of human life	Documents pertaining to national security or to ongoing criminal investigations	Compliance with DITSCAP/DIACAP security.
Note: References on DITSCAP/DIACAP (DoD Information Technology Security Certification & Accreditation Process/ DoD Information Assurance Certification and Accreditation Process) include: http://iase.disa.mil/ditscap/index.html, http://www.blackbirdtech.com/ditscap.pdf, and http://www.nswc.navy.mil/ISSEC/Form/DITSCAP/DITSCAP_intro.html.			

Gray areas do exist in a security zone scheme. Disaster and emergency plans generally require dissemination throughout the organization to be effective. Management still has to be careful about what information could become public. Enterprise-wide communications like newsletters, company news bulletins, and announcements pose another challenge. Could details about operations, facilities, events, and employee serve as leverage in an information attack on sensitive resources? Extranets used by customers and business partners, while valuable in e-commerce, can provide a portal for information leaks. They require close monitoring and supervision.Information links, tools, and Web sites used by customers, suppliers, and vendors can also provide avenues for sensitive information leakage. Constant vigilance on the content of documents passing through and into these portals is absolutely necessary. Chapter 3 discusses technology for monitoring these business channels.

GOOGLE HACKING

"Google hacking" describes the art of using the search engine of one of the world's largest databases to find highly specific information.

Google searching enables the location of specific Web sites that can address a particular research concern. If appropriate to the needs of the researcher, a Web site download can occur. Then, the researcher does detailed analysis on the site's content, but Google has another capability. The search engine locates diverse Web resources to address a specific query, so Google has immense powers of data aggregation.

When seeking a particular class of information such as credit card numbers, Social Security numbers (SSN), dates of birth (DOB), Excel spreadsheets with confidential financial data, or Montana vehicle license plates, if the information faces the Web, Google can find it. It can cast a broad net for information about a specific target; for example, ask Google for all information about John Quincy Doe from Any Town, Pennsylvania that has been posted to the Web from March 16, 1998 to the present. That type of search is quite feasible. In other words, the fantastic data aggregation capabilities of Google allow for finding a generic class of information or for locating very specific information about a given person or entity. Whether one wants to compile lists of marketable personal information for identity theft or build a dossier on John Q. Doe specifically, Google is an excellent tool to use.

Our aim is not to criticize Google for ethical reasons. After all, Google is a neutral technology, which can be called upon to serve good or evil, much like the telephone permits summoning emergency medical help or facilitates planning a bank robbery. By examining Google hacking techniques, the text seeks to make the reader aware of the dangers of placing documents in view of the Web without proper forethought. Before reviewing those techniques, the author needs to establish a few ethical ground rules.

The text looks at Google hacking methods from a general perspective. Details about finding specific marketable personal identifiers like credit card numbers or social Security numbers are omitted. The results of searches on specific persons or organizations are also omitted. Stressing the concept of the danger facing sensitive documents in the Web's plain view, the text deliberately skips over teaching specific techniques beyond this concept level. Understandably, the aim is not to create a treatise on how to commit identity theft or corporate espionage.

Google provides two basic approaches to finding Web-based information: command line searches and advanced searches. The advanced search is available by clicking on the "Advanced Search" link on the main Google page at http://www.google.com. That link takes the user to a data entry screen. The main sections on that screen are as follows:

- Find
- Language
- File Format
- Date (Updated)
- Numeric Range
- Occurrences
- Usage Rights
- Safe Search Filter
- Page-specific Searches
- Topic-specific Searches

The Advanced Search permits the user to craft a rather complex query with ease. For the input screen handles all the search variables in a structured way to permit the user to concentrate on what he or she is asking in the query without having to worry about the query's syntax.

In the Find section the user chooses from the search term matching: "All the Words," "The Exact Phrase," "At Least One of the Words," or "Without the Words." The last category means that search excludes from the result the words in the search term. For example, I want to see all Excel spreadsheets on the Web, but not those containing the word "bankruptcy." On the other hand, if one wants to see only Web pages with the precise phrase "Lady Macbeth" then the option for the exact phrase is the way to go.

The Language parameter merely specifies Web pages written in a certain natural language like English, French, or Russian. The usual default is English of course. File Format, however, specifies the type of document based upon the three character document name suffix like ".doc" for word processing documents. Other common suffixes include ".pdf" for Adobe PDF files, ".ppt" for MS PowerPoint files, ".xls" for Excel spreadsheets, and ".rtf" for rich text files produced from word processing applications. There are many more suffixes that are searchable beyond those listed in the advanced search. (One finds those documents through the command line search, which we'll discuss shortly.) Using File Format in a search creates a high level of precision, especially when combined with other search parameters. As stated earlier, we can narrow down our search to all Excel spreadsheets that do not contain the word "bankruptcy," or we can limit the search to only those containing that word. The combinations are endless.

In the Date section the user may specify locating Web pages only if they have been updated since a certain date. Numeric Range specifies some kind of numerical scale. For example, I only want to see digital cameras priced from $350 to $500. This search, in other words, returns Web pages containing numbers between the two specified numbers.

The Occurrences section states where the Google engine will look for the search terms. Choosing from the menu, the user may select:

- "Anywhere on the Page"
- "In the Title"
- "In the Text"
- "In the URL"
- "In the Links"

These options convey a feel for the structure or the "geography" of Web pages. A user or researcher may be not only highly precise in the exact language of the query, but he or she may define also in where to look for the information. These combined search features give Google immense search capabilities. (So, those who post documents to the Web need to be careful about every element in their document. Careful researchers will not miss anything, no matter how obscure the author feels it is.)

The sections Usage Rights and Safe Search filter are of just passing interest to our discussion. For the former allows the user to select pages that carry certain licensing rights, and the latter feature allows searching with certain security filters in place to prevent objectionable material.

The Page-Specific section allows the user to find pages similar to the page they have located or find pages that link to the located page. These features become quite valuable when trying to determine the reputation of a Web page in the Internet community. The number of links to page supply a gauge to evaluate how authoritative or infamous it is to the Internet community.

The final section, Topic-Specific Searches, has the following links:

- Books
- Code Search
- Google Scholar™
- Google News and Archive
- Apple Computers
- BSD UNIX
- Linux
- Microsoft
- US Government
- Universities

Most of these links are self-explanatory: they facilitate searching within a given topic. Code Search, however, is worthy of note to our discussion. This link brings the user to a page that permits searching using regular expressions, a new, very powerful feature. Regular expressions represent strings of numbers or characters in a generalized way. For example, a telephone number or a Social Security number has a generalized pattern. Accordingly, by creating this pattern using a regular expression, a user can find all Web pages that have specific numbers or character sets that conform to the pattern. Google's page has a link to explain how to craft regular expressions, the specifics of which are beyond the scope of this text. It is important to remember, though, that this capability further enhances Google's immense search capabilities. Let the unwary beware.

Command line searching occurs on Google's main page. It requires a bit of knowledge of Google's searching syntax, but no advanced computer knowledge such as programming or scripting is necessary. Using the command line offers greater flexibility in searching, and once a user becomes familiar with it, command line evolves into the preferred method. A number of search line commands exist; many of them mirror the features in Advanced Search. The most significant ones for the purpose of this discussion are:

- Allintext
- Allintitle
- Allinurl
- Daterange
- Filetype
- Intext
- Intitle
- Inurl
- Link
- Numrange
- Site

"Allintext" specifies to look for all of the search terms in the Web page's text. "Intext" looks for any of the search terms in the Web page's text. Examples of these searches include: allintext: "John Q. Public", which searches for the exact phrase in the text, and *intext: Administrator login,* which searches for either *Administrator* or *login* in the text.

The search pairs, "allintitle" and "intitle" along with "allinurl" and "inurl," perform the same search functions for terms in the Web page's title and URL respectively. "Intitle" is particularly useful in finding directory listings on servers; for example, a common search is: *intitle: Index.of "parent directory".* The URL search terms locate specific words in the address for a Web page, for example, the search, *allinurl: cgi-bin password,* identifies a page that may contain password information.

"Daterange" specifies searching for Web pages published between two dates. These dates are Julian dates, expressed as follows: *daterange: 2452122-2452234.* Julian dates provide a standardized method for expressing civil date formats like "June 6, 1944" or "6 June 1944" or "06/06/44." (For the details on calculating Julian dates from civil dates, see http://www.numerical-recipes.com/julian.html.)

"Filetype" is a powerful way to search for information. As indicated prior in the discussion, a search can specify an Excel file, a word processing file, a PDF file, or a PowerPoint file. The format for this search is *filetype: xls.* Actually, these common file types, such as Excel, are from among hundreds available on Web pages around the world. (For an extensive listing of file type extensions, see http://filext.com/index.php.)

"Link" permits identifying those pages that link to the specified Web page. This feature on the command line permits discovering additional information on a topic or on a targeted subject. "Site" allows including a particular domain name, such as *cnn.com* for the Cable News Network, in a search or identifying a specific type of site, such as *.gov* for a government site, as a part of the search term.

Finally, "numrange" specifies a search range between two numbers as with the Advanced Search featured described in prior paragraphs. An example in the command line would be: *DVD player $250..350*, which means for the search to locate DVD players in that price range.

Basic operators also function in search statements in the command line. The common ones are the exclusion sign (-), the inclusion sign (+), the exact phrase notation (double quotes around the phrase), and the wildcards for a single character (.) and for any word (*). *Orange – bowl* returns any page with "orange" in it but not those with "orange bowl." The inclusion sign forces the inclusion of something common as with *Star Wars Episode +1*, which returns all references to "Star Wars Episode One." *M.nd* returns "mind" and "mend." *Orange** returns "orange" combined with any other word. (For a complete explanation of all Google search terms and operators, see http//:code.google.com/apis/soapsearch/reference.html.)

Once the reader understands the syntax of Google's search terms, then the hacking part of the process becomes fairly clear. Google hacking involves combining the various search methods into a focused attack strategy. In developing these strategies, a Google hacker considers Web location (site), text location (in the body, in the title, or in the URL?), file type, and keywords. Other search factors also include time (date range) and number range, and if a researcher becomes familiar with regular expressions, as discussed in the Advanced Search, then pattern becomes a search factor too. Numerical patterns like telephone numbers, Social Security numbers, shipping tracking numbers, and so on are fairly easy to search for. In fact, Google has a page to facilitate searching for these common number patterns at http://www.google.com/help/features.html.

Textual pattern searches are also possible. Of course, the use of regular expressions enables looking for patterns of characters in addition to numbers. but the keyword component of a command line search also has the ability to pick up on common phrasing used on Web sites. Such patterns often come into play when trying to map out the structure and the configuration of a network. While these search methods are a little off the topic of content/document security, they do reveal, however, that both network infrastructure and content on the network are vulnerable to Google hacking. Consider the following searches:

1. "Microsoft-IIS/5.0 server at"
2. Intitle: "Welcome to Windows 2000 Internet Services" IIS 5.0

Both of them key into phrases used on the Web server pages using specific operating systems. When trying to compromise a network, search terms like these offer valuable reconnaissance data.

As far as content goes, picture using keyword phrases such as "checking account," "confidential," "for internal distribution only," or "executive salary." Combining these keywords with other search terms produces highly focused inquiries such as the following:

1. "Not for External Distribution" confidential site:gov
2. "Executive Salary" filetype:xls
3. "For Internal Use Only" filetype:pdf site:edu
4. "John Q. Public" salary filetype:xls
5. "Sally Q. Public" resume "c.v." filetype:doc

The combinations are endless. Google affords the ability to do a very broad search for specific information on an individual or an organization, or it permits a dragnet for particular information patterns. The technology facilitates both broad data aggregation and highly focused site searches. (See Table 2.3 for an example of how a Google can expand by adding search terms.)

OTHER SEARCH ENGINES

The discussion has concentrated on Google so far as a means of creating awareness regarding the vulnerabilities of Web-facing documents. Other search engines, however, also offer capabilities for data mining the Web. Dogpile® is one of note. It is a multisearch engine in that it searches Google, Yahoo!®, MSN Search, and Ask.com™ at the same time. So, Dogpile supplies very broad search capability.

Table 2.3: Expanding a Google Search

Location	Text location	File Type	Keywords	Range	Pattern
			"John Jones, Jr."		
	intext		"John Jones, Jr."		
	allintext	filetype:xls	"John Jones, Jr." salary		
site:com	allintext	filetype:xls	"John Jones, Jr." salary		
site:com	allintext	filetype:xls	"John Jones, Jr." salary	$50000..$150000	
site:com	allintext	filetype:xls	"John Jones, Jr." salary	$50000..$150000	Regular expression for Social Security number
Note: The search starts with just the target's name, but it increases in specifying particular data. Finally, it asks for a document, an Excel spreadsheet, which contains the subject's salary information and any data resembling the pattern for a Social Security number.					

Having an Advanced Web Search page, Dogpile like Google permits qualifying a search by "All of these words," "The exact phrase," "Any of these words," or "None of these words." Dogpile's Advanced Web Search also allows selection of language, date range, and domain (.gov, .edu, .com, and so on). Selecting the Results Display option permits sorting output by relevance to the search term (the most relevant first) or by search engine. Like with Google's Advanced Search, Dogpile also permits search filtering to screen out explicit adult content, and both Dogpile and Google permit searching for images not just text.

Ask.com offers numerous search tools in a menu format allowing a researcher to choose from Web, images, news, maps, Encyclopedia, blogs and feeds, and the like. Blogs, by the way, are a growing information source on the Web. Do not underestimate the amount of information that lies in this twenty-first century mode of communication. Someone may be acting as a watchdog of your organization and documenting a great deal on a regular basis. Be on the lookout. The search engine provides a preview binocular feature for certain sites, enabling the user to see the Web page before actually clicking on the link. Somewhere on the Web, someone has written about or otherwise documented your organization or you as an individual. Ask.com provides a convenient way to discover your "footprint" on the Web.

Alta Vista™ at http://www.altavista.com has an Advanced Search very similar to Dogpile's. It includes filtering by file type, domain, host (site), link, title, and URL, but, in addition, it offers searching by Boolean expressions, which is a fancy term for including AND, NOT, and OR in searches. These reserved words add a precision to doing textual searches, for example:

1. "Peanut butter" AND jelly returns pages containing both.
2. Peanut NOT butter returns any pages with references to "peanut" but none that contain "peanut butter."
3. Peanut OR walnut returns pages containing either term.

The usefulness of this search strategy comes into play when a name is fairly common or used in multiple contexts. For example, Francis Bacon is a famous Elizabethan author, but there is also a modern painter by that name. If one searches for the author, the search syntax can be "Francis Bacon" NOT painter.

The resources for locating Web-facing documents and information are immense, regardless of the researcher's preference for search tools. Never consider security through obscurity as an option against this reality. If sensitive data faces the Web by being accessible to the Internet, then someone will find it. The next and final section of this chapter suggests countermeasures for dealing with the challenge.

COUNTERMEASURES

Developing effective countermeasures against unintentionally Web-facing documents requires a good document security policy coupled with the establishment of security zones for a network. The information in Tables 2.2a and 2.2b summarizes both approaches. The cornerstone of the security policy is accurate classification of documents by their sensitivity level. Mandatory labeling of documents provides a method of tracking the location of sensitive information.

Well-defined security zones in the network ensure appropriate depositories for documents based upon their security level. Internal firewalls segregate the network into separate security zones. At higher levels of security, IP address segregation further reinforces the integrity of the security zones, and IDS (Intrusion Detection System) monitoring looks for unauthorized traffic between the zones. Deploying software to do internal checking for sensitive documents being located in incorrect zones is another recommended practice. (Of course, this is a good reason why documents need a security label.) These measures combined offer layered security to deter sensitive documents from becoming Web-facing.

In a network with a well-defined security perimeter, these countermeasures can be quite effective. Yet, the twenty-first century increasingly yields an information environment involving mobile computing and information flows that facilitate commerce but diminish the effectiveness of a traditional security perimeter. A counterbalance to these trends is a strong intelligence gathering campaign to learn what is available on the Web about one's own client or organization. The savvy information security professional constantly scours cyberspace to see what data is floating about regarding the entity he or she seeks to protect. Obviously, despite these efforts some documentation gets created outside the organization. Information given out at trade shows, professional meetings, and in published articles comes to mind. Policies on external publishing, on public speaking, and on public postings in blogs and in newsgroups are also necessary.

Other information flows such as e-mail, instant messaging (IM), FTP, postings on other Web sites, faxes, and printings to remote printers create a different form of documentation, which requires effective security countermeasures. Chapter 3 addresses the challenges of these information channels. Also, mobile digital devices such as PDAs, laptops, and digital cellular telephones pose unique risks to the security of the documents that they store. Chapter 4 deals with protecting these devices.

Remember to think always in terms of the following security classifications:

- Public
- Internal
- Sensitive
- Confidential
- High Security

Regardless of where your documents are physically located, regardless of the server or digital device that stores them, ask yourself if the level of protection matches the sensitivity of the document. Staying in that frame of mind will minimize information leakage from your organization.

Chapter 3

INFORMATION LEAKAGE IN BUSINESS CHANNELS

What is a document? The traditional explanation would probably include something like "an organized body of information containing text and, if appropriate, graphics to communicate ideas." A published document then is traditionally something stand-alone and fixed, static in nature. From the Latin word, *documentum*, which is literally "a lesson," our word "document" ultimately, derives from the Latin verb *docére*, which is in English the verb "to teach." A document is then "the means to teach."

In the twenty-first century, our traditional concept of a document yields to the dynamic reality of fluid communications running across computer networks and cyberspace. Text and graphics cut and pasted, instant messaging (IM) connecting scattered individuals, Web postings documenting thoughts and transactions, e-mail crossing offices and continents, peer-to-peer networks humming with activity, and FTP servers receiving and sending all manner of files, these images reflect the fluidity of modern e-commerce.

No longer can security professionals focus on just the content in traditional documents, whether they are paper or electronic. Instead, professionals must recognize that all the other data in the stream of electronic commerce constitutes a new documentation. While at times ephemeral, this data is the "means to teach" beyond what its authors intended. Tapping into one of these data streams may provide invaluable insight to a business intelligence analyst or to other individuals looking for sensitive information. Again, the reader must understand that gathering intelligence is almost always putting together the fragmentary and the piecemeal. All intelligence gatherers quickly become documentalists, specializing in the nuances of a wide range of communication types and document forms. Expect them to pick up on the details that help build a larger picture of an organization and its operations.

No one piece of information alone tells the whole story, but when aggregated, the diverse pieces reveal a broad picture. Such aggregation occurs across a broad band of information streams. These sources include:

1. Web postings (Blogs, newsgroups, electronic bulletin boards, online chat, discussion areas on sites, and so on).
2. E-mail (both internal and external).
3. Instant messaging (IM).
4. FTP servers (files deposited and retrieved).
5. Peer to Peer networking.
6. Faxes.
7. Networked printers (unsecured).

Some of these data streams are open source, such as Web postings and certain anonymous FTP sites, which permit guest postings and retrievals. Anyone knowing where to look can legally find the information. The others, while originating as an internal source, can become rapidly compromised by outsiders or exploited by unauthorized internal users.

For example, a user called Bob sends an e-mail containing sensitive information to Carol, who is authorized to receive the information. She thinks it has some great ideas in it, so she copies and pastes portions of the e-mail into a presentation she's making before a local professional group. Unfortunately, the thought of information leakage isn't setting off warning bells to her. After all, she is not divulging the entire document. Regardless of whom the users are, this activity goes on regularly, without any malice on the users' part. Cutting, copying, and pasting are universal activities in an age of constant information flow.

Instant messaging offers rapid convenience in communication. Immediate online conferences are possible between coworkers. The technology facilitates the transfer of files and documents between individuals. Unfortunately, most of this activity occurs in plaintext (unencrypted). Usually IM creates a transcript of conversations, which can be copied and pasted into other documents. Creating sensitive documentation "on the fly" makes IM a treacherous communication medium. Even if a company restricts IM access only to internal communications, information leakage is still a serious problem. Bob, a project manager, has a lengthy IM session with Carol about some recent software development issues. Later, Carl, an IT administrator with access rights to individual users' machines, remotely does a routine update to Bob's PC. During the process, Carl opens up Bob's IM program and sees the conversation transcript with Carol. He remembers his buddy Jasper is working on a problem like that at another company, so Carl copies the transcript and e-mails it to Jasper to help his friend out. Carl isn't a spy, and he isn't making a penny out of his action. He has the pure motive of just assisting a friend with a little bit of useful information.

John needs to send Richard a specifications document on a new digital camera his team is developing. The document, however, is too large to e-mail. In the alternative, John calls Richard and tells him that in the next ten minutes the document will be on an FTP server used by the company for public document downloads to customers and vendors. The server permits anonymous logins without an account or password. Even though the "spec" document has sensitive information, John thinks it will be all right since the document will be on the server for just a short time. After placing the copy of the document on the FTP server, John receives an urgent call to come down to the production line to consult on a critical issue. Meanwhile, Richard downloads the copy successfully from the FTP site, but he does not delete the document on the FTP server. Unfortunately, after several hours John returns to his office, but he forgets about the document due to pressing distractions. Two weeks later, when Richard mentions to him how useful the document was to his portion of the project, John remembers to remove it from the FTP server.

Peer-to-peer (P2P) networks operate without central administration and establish direct trust relationships between the users. Many security concerns arise from these networks regarding the authentication and authorization of users. A wise practice is not to permit sensitive documents to enter this networking environment due to the lack of formal access controls. The challenge for enterprises is to prevent the migration of sensitive documentation into P2P settings since authorized users may belong to P2P networks in addition to the enterprise's network.

Fax machines and printers, unless in properly secured rooms, may have sensitive documents lying unattended for extended time periods. Unauthorized individuals can read and copy these documents with no one the wiser. When sending a fax containing sensitive information, the sender has no assurance that the receiver's security is adequate. Some fax machines have a memory feature to permit a printout of the history of sent faxes and to allow the temporary storage of received or transmitted documents. The memory of some printers is accessible. Recent printers have slots for memory cards storing digital images and documents. Users may forget to retrieve their memory card after printing. In other words, fax machine and printers are venues for high latency of data, whether that data is in physical, paper documents or in electronic format. These machines become magnets for the inquisitive, the nosy, and those disposed to corporate espionage.

CONTROLLING BUSINESS CHANNELS

What concerns us is the lifeblood of modern commerce: the free flow of information. Without this flow twenty-first century business and government cannot function. As security professionals we face the dilemma of trying to impose reasonable constraints upon business channels without seriously hampering the commercial interaction of ideas. This task is not easy, and the previous chapters emphasized external attacks, which in many ways are less difficult in establishing countermeasures. Internal threats, however, are common with business channels. Those with access to sensitive information compromise such assets with relative ease, and they may commit these compromises through accidents and errors not just malice.

Education helps to a certain degree. Training staff on the security pitfalls of business channels may relieve some of the problems. Yet, the press of business, the need to meet deadlines and emergencies, and the compelling requirements of information sharing in the Digital Age all work against education as a sole remedy. Technology must assist in protecting business channels against:

- Cutting, copying, and pasting information from a document of a higher security level to one of a lower classification.
- Sending documents with a sensitive classification to a channel of lower security classification.
- Placing sensitive documents in a security zone of lower classification (See Chapter 2 and Tables 2.2a and 2.2b about security zones.)

- Creating sensitive information "on the fly" (IM, e-mails, and Web postings) without appropriate controls to detect it and to prevent its unauthorized dissemination.
- Allowing the high latency of sensitive information in and around fax machines and printers and in portable storage media such as memory cards, memory sticks, USB drives, DVDs, CDs, and removable hard drives without proper safeguards.

This huge security mission requires addressing these issues in three steps. First, an organization must have a method for tagging and fingerprinting sensitive content in electronic documents based upon the security policy. If sensitive content cannot be identified, then countermeasures will not work. Second, the business channel security system must have the capability to monitor outbound and internal traffic on the protected network. If a user tries to copy and paste sensitive information, that action must be subject to detection and inspection. The same goes for a user trying to place a document in an incorrect security zone, and third, the defensive system must enforce the security policy. Mere monitoring is insufficient. Immediate response may require one or more of the following: (a) blocking the action, (b) quarantining the sensitive data fragments or the whole document, (c) encrypting the information, and/or (d) notifying the security staff.

In addressing the three steps the enabling software must integrate into the organization's information infrastructure. In complex environments the components include databases, financial systems, file severs, mainframes, networking infrastructure, and network storage. Again, security zones become vital to the success of the security policy imposed upon the enterprise's network. Internal and external access to higher level zones requires strict access control and monitoring.

The key to the access control and monitoring is a centralized administration of content security. Content security enforces the integrity and confidentiality of individual documents, and it maintains the boundaries of the security zones by looking for out-of-bounds documents. Policies protect the integrity of sensitive documents by preventing unauthorized writing to and writing from them. Confidentiality is enforced by preventing the unauthorized transmission of a document into an unauthorized channel and by detecting documents that have a sensitivity label greater than the zone where they reside.

Central administration of content security occurs logically between the data sources such as databases, file servers, and mainframes and the business channels. In the central administration segment on the network, the processes of fingerprinting or tagging sensitive information, of identification and subsequent enforcement, and of security policy application occur. This technology along with the enforcement of security zones deters a low-level asset like a laptop from being able to access a high-level asset such as a sensitive database. It also greatly reduces the possibility of data fragments of high sensitivity being placed into a document of low sensitivity or being passed into unsecured channels. Paul Williams, the Chief Technology Officer of Grayhat Research Corporation in Houston, Texas, commented in his course on "Advanced Information Security" that "one does not want a $750.00 laptop to be able to compromise a multi-million dollar database." This thought needs to remain on the forefront of all of our thinking about content security.

Content security administration resembles IDS and IPS technologies. Intrusion detection and protection systems come in different "flavors." The simplest method of detection used by them is signature recognition. A particular attack has a definite signature or pattern that the IDS or IPS technology recognizes. In content security the software imposes a fingerprint upon sensitive information according to the security policy. This fingerprint serves as the identifying factor. In addition, the security software is able to detect sensitive information using the following methodologies:

1. Keywords
2. Patterns of words or phrases
3. By referring to lexicons or dictionaries
4. Using regular expressions (REGEX), for example, the system could look for Social Security numbers based upon a regular expression
5. Templates for specific security requirements or models like HIPAA or Sarbanes-Oxley

Content security imposed internally to protect business channels is a modified form of mandatory access control. Sensitive objects either have security sensitivity labels or acquire them through the detection process, and subjects cannot access objects if the access violates the security policy. If a user creates sensitive information, for example, on-the-fly, then the software examines the data using one of the methodologies above. When the monitor detects a possible violation (the user wants to e-mail the information outside of the company) then it blocks that action, and the e-mail is quarantined for examination by security. Obviously, data already fingerprinted would be immediately recognized by the system and similar defensive measures would occur.

Does this granular control smack of George Orwell's *1984* and its ultimate totalitarian figure, Big Brother? Are we heading toward a digital society where every transaction and thought in text undergoes scrutiny? Any thinking person must reflect in a sobering way on the possible consequences. While this text is not a treatise on political or social theory, it can offer some suggestions on making these controls acceptable to users.

First, have a clear, concise, and understandable policy on document and content security. Then, communicate that policy to all employees throughout the year. Stress that keeping sensitive information secure protects everyone's jobs.

Second, explain to everyone why certain information and documents will end up quarantined. Monitoring business channels has become a security necessity, but that issue does not mean security will become oppressive. Rather, the monitoring software and system is a safety net for information, not a tool for an inquisition.

Third, recognize that people make mistakes. Conduct investigations in a fair, courteous, and friendly manner. Avoid being overbearing in interviews associated with blocked or quarantined documents. Aim at educating users rather than chiding them. This approach does not mean that those who deliberately try to circumvent information security safeguards will not face discipline, but disciplinary action is always the fallback position not the starting point.

Fourth, constantly fine-tune the content security policies and make sure they stay current with the enterprise's needs. Recognize that some information previously thought sensitive may no longer be sensitive. Adjust your policies accordingly, and refrain from overenforcement. Too much security interferes with the work of the enterprise and chafes employees on a daily basis. Always seek an appropriate balance between security and information flow. In order to do this balancing act, consulting with data owners needs to be a regular part of the security policy development process.

Finally, have policies and auditing procedures to ensure security personnel and managers do not abuse the content security system for personal gain. Under no circumstances should personal information or communications be acted upon unless those transmissions violate the security policy or the acceptable use policy of the organization. The arbitrary surveillance of certain employees is taboo. No one's specific communications should undergo intense monitoring unless the facts point to the need for a legitimate investigation. These restrictions are matters of simple fairness and respect for employees' privacy. Abuses in these areas create an atmosphere that raises serious questions of civil liberty.

WHAT DO INFORMATION THIEVES WANT?

Information thieves seek marketable data on individuals, sensitive business information, and information about network infrastructure that can lead to exploits against the network. Personal identifiers such as Social Security numbers (SSN), credit card numbers, bank account numbers, and insurance policy numbers are among the common targets. These consumer identifiers, while prime targets, offer an advantage to a content security system in that they have a data structure, which is fairly easy to recognize. Regular expressions usually provide a searchable pattern to detect the unauthorized movement or copying of these data types. As far as network access or infrastructure information goes, logins, passwords, internal IP addresses, and machine names also have recognizable patterns subject to detection by regular expression and keyword searches.

Fingerprinting or tagging confidential records such as medical documents and financial statements, spreadsheets, or documents found on the network enables a reasonable degree of tracking. In addition keyword, pattern recognition, and lexicon-based searches facilitate detecting and locating these records. Source code, proprietary recipes, engineering designs, patent information, marketing plans, and documents pertaining to trade secrets are all tracked using fingerprinting and textual searches.

The tagging or fingerprinting of sensitive documents and the pattern recognition of sensitive information create an inner layer of security. These measures deter unauthorized entry of sensitive documents and information into various business channels. Myopic vision results, however, unless the security professional realizes that these measures are but one level of security within an overall plan. Defense in depth resolves into two bottom-line issues: preventing outsiders from gaining access to the inner sanctum and preventing insiders from exploiting their access rights to the inner sanctum.

Outsider attacks on the inner sanctum of sensitive documents have several manifestations:

1. *Exploiting a trust relationship.* The attacker feigns being a user, a process, or a server that the sensitive object trusts.
2. *Obtaining an authentication credential.* A malefactor steals or compromises a login and password, a token card, or an access card to authenticate herself to the object.
3. *Usurping a trusted access channel.* A vendor, a supplier, a customer, or a mobile employee has trusted access to the network through a VPN or a portal on an extranet. The attacker finds a way to hack into the channel. Always assume that third parties holding trusted access will not be as careful as you are about security. (The classic discussion of this attack scenario is Carolyn P. Meinel's October 1998 article in the *Scientific American*, "How Hackers Break In . . . and How They Are Caught.")
4. *Social engineering their way in.* An attacker fools employees or third parties with access rights into either granting him access or in sending him the sensitive information.
5. *Researching for the sensitive information.* People do get careless and place sensitive documents on the network with a view to the Web. As depicted in Chapter Two, Google hacking can then take over.
6. *Hacking their way in using technical methods.* Buffer overflows, spoofing attacks, and compromising the Web server all come into play.

Developing a counterpoint for each of these methods requires a layered approach. It is not simply a matter of saying, "we have content security software in place, and therefore we are secure." Obviously, a network additionally requires perimeter security measures like a screening router, a firewall, and IDS and IPS (intrusion detection and protection) monitoring to detect and to block external technical attacks. Defenses against Google hacking, covered in Chapter 2, include security zones and the proper placement of documents within those zones. (Content security monitoring also can help locate misplaced sensitive documents.) Education against social engineering attacks is a constant matter for employee training. Content security measures help to a degree in this area because they monitor the traffic in sensitive information.

Any trusted channel or relationship requires extra monitoring by an IDS or IPS and by the content security system. Special attention needs to be paid to the traffic moving in and out on that channel. With regard to authentication, single factor authentication is very bad for accessing any sensitive object, whether it is a file or a database. Use two-factor or multi-factor authentication for access to sensitive materials on the network. (Factors come in three forms: what you know, a password; what you have, an access card or token; and, who you are, a fingerprint. Using more than one factor results in two- or multi-factor authentication.) Build into contracts with third parties the right to audit their security procedures to ensure security compliance. And, never grant anyone *carte blanche* access. They should only have the amount of access necessary to get their job or mission accomplished.

Insider attacks on sensitive documents require cunning, but these attacks have the advantage of already being within the security perimeter. They include:

A. *Cutting, copying, and pasting* sensitive information into documents or messages of a lesser security level.
B. *Placing sensitive documents* onto portable media like a USB drive.
C. *Paraphrasing sensitive information.* Attackers try to alter the pattern of the language but still convey the ideas.
D. *Printing out sensitive files.* Then, they carry them out as paper documents.
`E. *Trying to strip off the security sensitivity label on the document.* By removing the electronic tagging or fingerprinting, they elude detection of the document.

Content security monitoring measures deter “Item A” actions. With regard to transferring sensitive documents to portable media, content security software should create an audit trail for transfer transactions. In addition, security software is available to prevent the unauthorized copying of files from a local machine onto an external drive or medium for an added layer of protection.

“Item C,” at first, seems a clever “dodge.” Yet, if the content security monitoring software has enough depth in resources to look for sensitive information, in that it uses multiple methodologies, the insider will still face a high possibility of detection. Printing is a business channel like any other, so monitoring should detect it. Electronic fingerprinting of sensitive documents should be robust enough to resist tampering. In any event, even if a label gets removed, then the other detection methodologies should still recognize it as a sensitive document.

OTHER CHALLENGES

Monitoring business channels aims at detecting textual information of a sensitive nature in digital format. It does nothing with regard to analog sound waves. In other words, conventional telephone communications are not a business channel content monitoring can address. True, security personnel can randomly monitor conventional telephone conversations, and technology can play a role in monitoring Voice Over IP telephony (VOIP) since that information is digital. But the question becomes, is it worth it? Earlier in the text, we discussed the Orwellian image of Big Brother. Verbal communications are a fundamental human freedom, and except in rare instances of an authorized, legitimate investigation, our courts and the people balk at any overzealous monitoring of those communications. Any verbal communication, whether on the telephone, at a bar with friends or chatting up a member of the opposite sex, or in giving a lecture runs the risk of information leakage. Education is the first line of defense when it covers protecting against corporate espionage and social engineering. Investigations come into play in matters regarding serious breaches of information security. Although the bandwidth of verbal content is enormous, it remains an area for reasonable restraint in everyday security monitoring.

Obfuscation is another challenge for information security professionals. Insiders can try to encrypt sensitive data to elude detection. They can employ steganography to hide sensitive files within innocuous image or sound files. CAPTCHA (Completely Automated Public Turing Test to Tell Humans and Computers Apart) techniques are another method for obfuscation. These techniques distort text so that "bots" and other computers systems cannot read the text. Humans can still read the information because our pattern recognition skills are much broader in scope than a machine's. We can deal with some ambiguity. A software program can render a sensitive document into CAPTCHA text to avoid detection by the content monitoring system. (For more information on CAPTCHA, see these Web sites: http://www.captcha.biz/ and http://en.wikipedia.org/wiki/CAPTCHA.)

Another obfuscation method is to convert the text into a foreign language. Even worse, the insider uses software to translate the sensitive document into a language that uses a different alphabet or writing scheme such as Greek, Russian, Arabic, Korean, Japanese, or Chinese.

A less exotic method requires simply taking a screenshot of sensitive data on the screen and saving the information as a graphics file. Since it is an image and not text, detection of the sensitive information that it contains will be more difficult. However, depending upon the content monitoring system used, statistical analysis may be able to detect certain color layouts or mappings. In any event, insiders will always be trying new techniques to get around any content safeguards. Help is always available out there for those seeking secrecy or privacy. The August 2006 issue of *Wired* magazine, for example, has a section, "Foil a Snooping Boss." This section has suggestions on cloaking Web searching, encrypting e-mail, encrypting IM conversations, and on other cloaking methods.

Another challenge to document or content security is out of boundary creation of sensitive information. An employee memorizes sensitive information or becomes familiar with it in the course of their work. The individual then takes the information out of the organization in their head and later renders it to paper or into electronic form outside of the organization's security perimeter. In the age of digital photography, capable of high resolution images using small, compact cameras, insiders can take photographs of sensitive information and documents. The cameras and their associated memory cards are easy to transport off-site with little chance of detection.

RISK MANAGEMENT

Evaluating methods for deterring the compromise of sensitive documents in business channels requires taking a broad view. The foundation of any effort is a sound security policy, for it is nearly impossible to enforce content security without the guiding vision that such a policy offers. Such a policy must consider the following:

- What documents are sensitive, and which security zones will protect them?
- What level of monitoring will be applied to the business channels?
- Will content security monitoring be able to employ a variety of pattern recognition methodologies? How will monitoring handle documents or messages in foreign languages or scripts? How will image files be examined?

- How will documents or communications encrypted by the user be handled? How will steganography tools be detected on the network? Will monitoring include scanning for steganographic embedded data in files?
- Will monitoring be able to detect files with CAPTCHA text or other distorted text schemes?
- What will be the policy on the use of digital cameras? How will their use be controlled?
- What level of education or training will employees undergo regarding social engineering and corporate espionage?
- Will presentations before professional and trade groups require clearance by security prior to the event? What about an employee writing an article or a book, will that require clearance too?
- Will customer, supplier, and vendor lists be "trapped" with some bogus listings that contain an address, e-mail address, and telephone that you control? (If you get contract there from a company soliciting business, then you know the list has been compromised.)

Protecting business channels is a complex endeavor for information security. In assuming that responsibility, real world constraints require security accepting the fact that some information will leak. Preparing for the various modes of attack and developing measures to detect and to block those attacks via business channels is essential. Equally vital is to limit the exposure of highly confidential or sensitive documents only to those with a real need to know. This enforcement of *least privilege*, along with the establishment of security zones, will maximize the odds that when leaks occur they will be of low-level information assets.

Table 3.1: Business Channels

Channel	Characteristics	Countermeasures
E-mail	Prone to copying, cutting, pasting of sensitive information. May use various obfuscation techniques.	REGEX, Pattern recognition, Document fingerprinting
Web Postings	Divulging sensitive information in newsgroups, chat rooms, filling out forms on Web pages.	REGEX, Pattern recognition
IM	Creating sensitive information on-the-fly a problem. Obfuscation: encryption.	REGEX, Pattern recognition
Printers	Physical security of the printer. Printer must be trusted to receive sensitive information.	Blocking sensitive documents, Document fingerprinting
FAX	Physical security issues. Unauthorized channel for sensitive documents.	Blocking sensitive documents, Document fingerprinting
FTP	Improper placement of sensitive documents.	Blocking sensitive documents, Document fingerprinting
Peer-to-Peer (P2P)	No centralized security administration. Sensitive documents should not pass to this channel.	Blocking sensitive documents, Document fingerprinting, REGEX, Pattern recognition
Verbal Communications	• Analog telephone is out-of-band. • Includes public conversations and discussions. • VOIP	Education of employees, Monitoring of VOIP is possible
External Publications	• Technical papers • Articles • Documents filed with regulatory agencies	Out-of-band: Education and security pre-screening
Digital Images	• Screen captures • Digital photography of sensitive documents	• Monitoring of image files • Control of cameras

Chapter 4

DIGITAL DEVICE THEFT

In an increasingly mobile and digital society, employees and information workers process data in a wide range of environments. Coffee shops, libraries, airports, lobbies, shopping malls, and other numerous locations welcome the mobile computing professional. PDAs and laptops carry large amounts of sensitive data and documents, often without adequate safeguards to prevent their theft and the outright compromise of their information contents. Thieves find mobile computing devices as juicy targets. They are reasonably easy to sell for quick cash. More sophisticated thieves look beyond the physical asset and see the value of the information contained on the device. Information security professionals need to guard against both types of thieves.

Every conceivable organization has lost valuable, sensitive data and documents via mobile computer theft: hospitals, universities, consulting firms, banks, government agencies, health insurers, and even charitable organizations like the YMCA. Defending against this growing threat requires a threefold effort. First, educating users is vital. They must understand the dangers of transporting sensitive information in electronic files and documents on their laptop or PDA or other mobile digital device. Second, appropriate countermeasures need to be in place to guard against physical theft. If physical theft does occur, technical means need to be in place to aid in the retrieval of the equipment. Finally, the sensitive data on the device needs to be protected via encryption. Additionally, if the information is a customer list or a supplier or vendor list, it needs to be "trapped" with bogus listings where security controls the addresses, e-mail addresses, and telephone numbers for those listings. If contacted via any of those points, security will know the list has been compromised and who purchased it.

Savvy in the field of information security does not guarantee safety unless these countermeasures undergo strict implementation. Stephen Manning in his article, “The Biggest Threat to Computer Security? Carelessness,” (*Austin American-Statesman*, June 19, 2006) points out that the Department of Veteran Affairs lost data on 26.5 million veterans via a laptop theft. He also identified Ernst and Young, a national accounting firm offering information security consulting to major corporations, who encrypted 30,000 of its own laptops for use by staff, still had its client, Hotels.com, face a compromise of 243,000 customer records due to a stolen laptop.

The first rule of thumb is that in many cases the extensive data files do not have to be on the mobile computer in the first place. In pre-digital days, information workers stuffed their briefcases with every document they could find on their desk before heading home so they could “get things done” that evening or over the weekend. In reality, most people did not even look at most of the paper they toted home. The Digital Age has the same practice going on. Workers have whole universes of data on their mobile computers, much of that data, however, needs referencing at best occasionally while working. Getting around this “packrat” mentality is difficult, but it needs to be a part of mobile users’ training.

Users also need to understand where they are prone to physical theft. Any place where a user leaves the laptop unattended runs a risk of theft, and perhaps, many people will think that is just common sense. “Common sense” often yields to practical needs. For example, during a night on the town with business associates, a user probably would leave his or her laptop at the hotel. One wouldn’t normally take a laptop on a romantic date. Leaving a laptop in a work area while everyone goes to lunch together is another common practice. Does a user take the laptop to the restroom in a restaurant or in a café? How does one keep an eye on the laptop when going through security at a building or in an airport? All of these issues are valid considerations. We’ll explore possible solutions in the next few paragraphs.

Hotel rooms are terrible places to leave any sensitive information. The guest has little control over who enters the room when he or she is not there. Staff members of the hotel usually have passkeys to access the room. During cleaning by staff, doors are left open and virtually anyone can go into the room. If the room does not have a safe, place your laptop in the hotel’s safe or lockbox. Never dispose of sensitive papers or computer media in the hotel room’s trash. If visiting a client or customer that can be trusted, dispose of such materials in one of their shredder collection boxes. If you are not comfortable with that action, bring your sensitive, data trash back with you for disposal within your own security perimeter.

Executives traveling, who require heavy information support by bringing a sizeable number of sensitive documents, whether in electronic or paper form, should have a security professional accompany him or her. The security professional can watch over the materials and equipment when the executive is not present. While this step is an added expense, it adds to the executive's convenience by not having to constantly pack up and move sensitive materials.

When working on a project with a group, and the group decides to leave for lunch, have the laptops locked in a secure room. Ongoing projects that require collaboration over a period of days should have a designated secure room to prevent the theft of digital devices and sensitive documents used as work-in-progress. If a room is not available, have your laptop fitted with a locking bracket so it may be secured with a locking cable to a desk or wall. Never assume that leaving a laptop or PDA unattended at your place of work means that it is safe. Much of the digital device theft that occurs happens in the workplace. In the absence of a locked room or using a locking cable, place the laptop inside your desk and lock the desk before you leave the area. Whenever the device is out of your sight, consider it prone to theft unless properly secured.

In a restaurant or café, use the buddy system. Have someone you know and trust watch your device while you leave the table. If you are alone, leave it locked in the trunk of your vehicle if temperatures permit doing so. (As an added protection, have a locking bracket in the trunk of your vehicle and attach the device with a locking cable to it.) If leaving the device in the trunk is not practical and you must leave your table for a few minutes, then take the laptop with you.

In an airport, use the buddy system if possible. Have a coworker, relative, or friend watch the device when passing through security. Having a laptop with a removable hard drive and/or a removable battery reduces the usefulness of the device for a thief. Carry those removable items separately. Be aware of who is around you in an airport, a mall, or other public area. Your laptop or PDA is highly vulnerable in these settings. Be on guard against a sudden theft. Never leave the device unattended. Do not use the device in public at an airport or a mall unless you really need to get something urgent done. Those places are magnets for thieves on the prowl. (Using a laptop or PDA in a wireless café is a bit of a different setting as many people go there for that purpose, but you still need to be watchful and on guard.) Do not carry it in a bag that advertises: "I've got a laptop." Carry it in another more ambiguous piece of luggage.

Public workplaces such as copy centers or business centers in hotels appear relatively safe. Do not assume that they are any safer than any other public area. Unattended laptops, PDAs, digital media, and sensitive documents can disappear in an eye blink while you are asking a clerk at the counter a question. Again, keep these items with you at all times. While these places serve as offices away from home, they are not your office. Strangers are nearby; keep that in the forefront of your mind.

Before moving on from the topic of mobile theft, be sure to remember to exercise restraint in the amount of information that you carry with you when traveling. Have only the physical documents that you actually need. On your digital device, have only the files that you actually need. On your cellular telephone, keep the amount of information stored on the device to a minimum. If you are a person that is constantly leaving your cellular telephone back at the restaurant, keeping the information "footprint" on the device small is a wise precaution. What you do not have with you, you cannot have lost or stolen.

TECHNICAL DEFENSES

What are the main technical concerns with protecting your PDA or laptop from theft or information compromise? First, the user or owner should have countermeasures in place to minimize the impact of a theft, should one occur. These measures form an inner line of defense if the physical steps previously discussed fail. They, of course, are no substitute for the physical security measures and vigilance required of the user or owner. Second, technical measures must in place to prevent the interception of the PDA's or laptop's communications, whether over the Internet or via a wireless connection. Both areas of technical countermeasures are essential to maintaining a high level of security.

Laura Taylor in "PDA Security 101" (*InternetJournal*, April 7, 2003) makes several important observations about technical security for Laptops and PDAs. The first important concept is that of a VPN, a Virtual Private Network. With a VPN, a user can transmit over the Internet to communicate securely with a server or network back at the business or organization. A VPN uses a suite of protocols and encryption technology to protect the communication path between the PDA or laptop and the intended network. She advises wrapping a PDA or laptop in a VPN when communicating back with the organizational network. The digital device and the network should authenticate using X.509 certificates, which are a technical way of saying that the two communication points should recognize each other via a standard digital certificate, a cryptographic shaking of hands. The process is very similar to making a purchase or connecting to one's bank using SSL over the Web, creating a secure HTTP connection.

VPNs permit the accessing of a database or file remotely in a reasonably secure manner. They can reduce the risk of lugging around extensive files or databases containing sensitive information on a PDA or laptop. In other words, VPNs bolster physical security when traveling on the road, because they reduce the information "footprint" or "load" on your mobile device. They also protect your communications back to your base operations. Expect, however, a performance hit on processing information when using a VPN because of the added processing overhead they create due to the added protocols and encryption.

PDAs are essentially always on. The reason requires them to be frequently cradled to link to a PC so their battery can receive a recharge. Since they are always on and if they are wireless capable, they require shielding when in transit. Use a shielding bag to avoid leaking wireless transmissions. Access points (APs) to wireless networks are quite prolific these days, and your device could end up communicating with an AP without your knowledge. Mobile Cloak is a manufacturer of shielding bags. Look them up via a Web search engine for specifics on implementing this protective technology.

Encryption is the next technical defense available to protect the data on PDAs and laptops if stolen. Choices in this area include PGP (Pretty Good Privacy), Windows EFS (Encrypting File System), and X Tool Data Protector. (See the Bibliography, "Useful Web Sites.") Other encryption solutions are available through various vendors on the Web, but always seek vendors that supply a publicly-known encryption algorithm. Proprietary encryption products, where the algorithm is secret, and therefore untested, may not be robust in their protection.

Encryption renders plaintext, ordinary readable text, into ciphertext, which cannot be read without a decryption key. The encryption process “scrambles” the plaintext in two possible ways” substitution or transposition. In a substitution cipher the plaintext letters or characters are replaced by other characters, letters, or symbols. A transposition cipher rearranges the plaintext into an unrecognizable form. Robust encryption involves usually the combination of both techniques, and most important, the encryption process has multiple rounds of substitution and transposition.

Since computers are incredibly fast compared to humans, simple substitution and transposition schemes will not stand up to computer analysis. A robust algorithm or procedure is necessary to ensure the encryption resists computer attacks for the foreseeable future. When purchasing and deploying encryption protection for documents, one should stay clear of proprietary encryption. A proprietary encryption method keeps its algorithm (procedure) secret. While this secrecy sounds like it increases security, in many cases such security does not guarantee robust protection. A secret encryption method may not be mathematically sound. Without public testing of the algorithm by experts, its robustness to attack cannot be established. Publicly tested algorithms like Triple-DES, AES, and RSA establish their resistance to cryptological attacks through the cryptology community’s rigorous testing.

This short overview of cryptography hopefully establishes the importance of using publicly tested, robust encryption tools. To a forensics expert, just because a file is encrypted does not mean the file’s protection cannot be broken. The algorithm may be weak. The password or key may not be strong, or the party using the encryption may be sloppy and keep the password or key in an easily accessible place (like on the hard drive of the same machine). Encryption, if weak, is a small hurdle, but not a barrier to a document’s contents.

Make sure that your users understand how to use encryption and make it a default process for files on their device. Keeping the key or password on the device or on a sticky note on the device is an absolute “no-no.” Education of the users is critical to ensure appropriate cryptographic safeguards are followed. Each manufacturer outlines the necessary steps to guarantee their encryption remains robust. Covering these guidelines is a necessary part of user training. Properly encrypted data, if on a stolen digital device or otherwise compromised, will not do a thief much good, which is what you want. If you can’t protect the device, at least you can protect the data.

Make sure that you establish security policies for PDAs and laptops. These policies should cover the following issues:

1. Who may use mobile digital devices in the field or while traveling?
2. What data is permissible on the devices?
3. What data should be accessed by a VPN instead of being on the device itself?
4. What type of encryption must be used?
5. How will the user ensure that encryption of sensitive files occurs?
6. What physical safeguards against theft must the user employ?
7. Will the devices have anti-theft tracking on them?
8. What backup procedures will be used for the devices?
9. Will shielding be used for PDAs?
10. Are personal PDAs and laptops permitted for holding the organization's data?
11. Under what circumstances will wireless use be permitted?
12. Will mobile devices be equipped for electronic shredding of documents?
13. What protection against viruses and mobile malware should be provided for the devices?
14. What training in mobile computing security should users receive?

Being able to trace a stolen PDA or laptop is an essential security step in providing layered security. Computer tracking software permits obtaining a geographic fix on a stolen a machine anytime the machine connects to the Internet or to a telephone line. A central monitoring station gets an alert as to where the machine is located. This alert takes place unknown to the person in possession of the machine or device. The party who stole the device will not know the tracking software is on the device, and he or she will not be able to remove it even if they find out.

The central monitoring center, in addition to receiving an alert regarding the location of the stolen device, has an added capability. Remotely deleting or erasing files and data is also possible. A thief may have physical control over the PDA of laptop, but denying him or her access to your sensitive files is quite possible through the remote deletion and erasure capability. This feature has its greatest power when the files are also encrypted. Encryption prevents a thief from copying or transferring sensitive data prior to hooking the laptop or PDA to a telephone line or accessing the Internet.

Several vendors provide the online tracking service. The Stealth product is well known and cited by Laura Taylor in her article. The Web site for the Stealth product is http://www.computersecurity.com/stealth/computer_tracker.htm. In addition to the online detection service, the same vendor also provides a tool for bit wiping or electronically shredding data on your laptop or PDA. The Web page for the vendor's Data Protector is http://www.computersecurity.com/stealth/data_protector.htm. (Taylor's article also lists other mobile computing security vendors covering areas of protection ranging from VPNs to encryption.)

Another important factor to consider for mobile device security is data backup. One of the reasons that workers lug around large amounts of data is convenience. Having a database on your local machine speeds up the work process since this requires checking the database on a constant basis. The added convenience of large capacity, but small USB and removable drives also makes the process much easier. Carrying one's sensitive data on removable drives as a backup, however, creates additional exposure. Drives are easily lost or stolen. They are also prone to rapid copying.

A solution to the problem created by this added exposure is online data vaulting. A user in the field can establish a secure link to an online vault and send the data there for backup storage. This practice eliminates the need to store the data on tapes or backup media, such as removable drives in the field. Coupled with the use of a VPN to access sensitive information, online vaulting offers protection at both ends of the security equation. The data remains secure when it is accessed behind a secure perimeter via a VPN. It remains secure when in storage in the online vault.

In addition to these technical defenses, the owner should also have basic identification about their PDA or laptop. If a theft does occur, having the necessary identifying data aids in reporting the incident to the police. The identifying information includes the following:

1. Manufacturer
2. Model
3. Year/Date purchased
4. Serial Number
5. Value
6. Other identifying marks

Michelle Johnston Sellicito's article, "Securing Your Laptop" (informit.com, June 4, 2004), makes several suggestions regarding physical security that echo the previous discussion. She observes that laptops and other mobile digital devices are easily left behind. A laptop is a theft-prone commodity, and its data makes it even more attractive. Sollicito's main points include:

1. Keep mobile devices in a locked room or area when not in use.
2. Carry them in a bag or case not specifically designed for laptops.
3. Do not put the devices in areas where you do not have direct control.
4. Do not leave the device in the car.
5. Where feasible, secure the device by cable lock.
6. Use a theft prevention plate that has a bar-coded unique identifier. (See http://www.computersecurity.com/stop.)
7. Consider using alarms on the laptop to deter physical theft.
8. Use an online tracker to establish the device's location in the event of theft.

Theft is often a crime of opportunity. The more that a user disregards this fact the greater the chances are of becoming a victim of theft. Retail stores constantly battle this theft as a fundamental part of doing business. Multiple layers of defense involves physically locking down the devices, limiting the files a potential buyer can access, and using the devices in a restricted interface such as being able to execute only certain programs. These steps ensure that users cannot exploit the device if they steal the computer. Table 4.1 summarizes applying these protection fundamentals to mobile devices in various settings.

To summarize, the most vulnerable places for the theft of digital devices are, in order of vulnerability:

1. Airports, malls, and around public telephones and information kiosks in retail areas.
2. Security checkpoints.
3. Office work areas.
4. Copy centers and business centers.
5. Restaurants, bars, and cafes.
6. Hotel rooms.
7. At schools, universities, and colleges.
8. At trade shows and speaking engagements.
9. In libraries.
10. At business meetings.

Table 4-1: Minimizing Mobile Device Theft

Approach	Methods
Physical Lockdown	• Lockdown cable • Anti-theft plate
Segregation	• Locked room • Secured area • Guarding of device
Rendering Data Unreadable	• Encryption • Electronic shredding
Limiting the Information Footprint	• Use a VPN for remote access • Use online vaulting • Limit the files available on the device
Tracking the Device	• Online anti-theft tracking • Anti-theft plate
Controlling Communications Leakage	• Use a VPN • Place in a shielding bag • Limit transmitting sensitive information on a WiFi

Remember that digital device theft is one of the primary attack methods for compromising sensitive information. If an adversary cannot technically hack into your network or system, he or she will try to steal any mobile information that you have circulating outside of your security perimeter.

Chapter 5

MAGNETIC, ELECTRONIC, AND OPTICAL PERSISTENCE

Even with the best of network security, and the taking of extensive security measures such as security zones, serious information leakage still occurs unless measures for the secure destruction and/or reuse of computer media are in place. If one wants to find out about an organization or business from an insider's perspective, all he or she needs is a computer or computer media containing sensitive information. Many times, the information researcher does not have to breach the target's security perimeter to gain access. The sensitive data ends up as a machine or storage device tossed in the trash, donated to a charity, or sold away a bargain prices. Focusing only on "something to get rid of," the person in charge of disposal may be ignorant of the steps necessary to prevent sensitive data from leaking to the outside world in the disposal process.

Reuse of computer media is another serious concern in developing a document security program. If sensitive documents are on storage device, it should not be reused by another party, unless that data undergoes a secure removal process. Many users think that deleting a file in an application removes it from the storage medium. Unfortunately, the erasure process is not that simple. Ordinary deletion of a file simply makes that file no longer visible to the application and to the operating system. Most operating systems simply change the initial character in the file's name to render it "invisible." Even when the operating system writes over some of the sectors containing the original file, portions of the original file may remain in "slack space." Knowing how to recover "deleted" files and data in "slack space" are what earn computer forensics experts their salaries, as we shall see in Chapter 7. Meanwhile, accept as a maxim that if sensitive data remains on a digital device or medium, then someone can find it. Incorporating policies and procedures for data object reuse and secure disposal should be a part of any document security program.

A recent news account indicated that examining "old" computers and discarded computer media is a big business in Nigeria. Computers discarded, trashed, or given away are ending up overseas. Recovering sensitive data from these computers is becoming a boon business in Third World countries like Nigeria, a country noted for being the source of many online scams. The information serves as the basis for identity theft and corporate espionage. When your client places an "old" computer or associated media into the stream of commerce by relinquishing all physical security over the machine and its contents, then do not be surprised in whose hands the machine and its data end up residing.

The mindset of many users is that once a computer floppy hits the trash that it becomes unusable. This wishful thinking does not protect the data, and while this storage technology is on a major wane, plenty of floppies containing valuable and sensitive files are still out there. When they get replaced by USB drives or memory sticks, in the trash they go. The business intelligence analyst needs to know only a little bit about how to clean the floppies to get them usable again. *Secrets of a Super Hacker*, published in 1994, when floppies were "King," describes in detail how to retrieve them from the trash, how to clean them, and even how to place them in a new jacket. Expect any competent intelligence specialist to have at his or her disposal computers equipped with drives and interfaces to handle almost any kind of computer media from 5.25" floppies to SD cards. Also expect them to have enough knowledge of computer forensics to glean information from almost any "trashed" computer storage medium.

Residual data on magnetic, optical, and electronic storage media pose a grave security risk to every organization. Proper disposal and reuse of these media are not optional, nice things to do when you get a chance. Security in this area is mandatory to protect sensitive documents from unauthorized disclosure. In fact, at the federal level FIPS-199 and 44 U.S.C. 3542 establish the duty to protect sensitive and confidential information by appropriate means.

The National Institute of Standards and Technology's (NIST) "Special Publication 800-88" regarding sanitizing computer media identifies two forms of media: hard copy and soft copy. Hard copies cover paper, microforms, printer ribbons, and platens used or produced by computer systems. The common way to destroy these items is shredding or burning. (Chapter 6 discusses these media at length.) Soft copy media contain magnetic, optical, or electronic representations of data. They pose unique reuse and destruction issues. The protocols for each depend upon the characteristics of the storage medium, which later portions of this chapter discuss.

The soft copy media that the chapter focuses on are as follows:

- Computing machines, devices, and network equipment
- Floppies
- Hard drives
- USB drives
- Memory cards and sticks
- Zip disks
- Magnetic tapes
- Optical disks
- Memory
- Magnetic cards

According to "NIST 800-88," the sanitization methods fall into the following categories:

1. *Disposal* – where the media are discarded in a prescribed manner but undergo no sanitizing prior to disposal. (Recycling paper or other media not containing any sensitive information, for example.)
2. *Clearing* – data is removed from the storage device, usually in a manner prescribed by the manufacturer, so that it is no longer recoverable by the device's keyboard or user interface. This method can include overwriting the data on the medium. (Resetting or clearing the contents of a cellular telephone is an example of this method.)
3. *Purging* – removing data so that it cannot be recovered in a laboratory. (Degaussing magnetic media is an example of purging.)
4. *Destroying* – renders the storage medium physically incapable of being a source for the data. (Shredding, burning, pulverizing, and melting are common forms of media destruction.)

One point to make clear before moving on with the discussion is an issue about magnets. A common notion suggests that placing a magnetic storage medium in proximity to a magnet or an electromagnet automatically erases the data. The author has experimented with exposing floppy disks to various handheld magnets and small electromagnets and has concluded that casual contact between the two does not erase the data. Waving a refrigerator magnet a few times over a floppy probably will have no effect on the data. This is not to suggest that exposing one's hard drive or magnetic storage devices to magnets is a recommended practice. It is not. However, to be assured of robust purging of magnetic media, use a degaussing device, not a kitchen magnet.

Obviously, in evaluating the steps to reuse or to destroy computer media containing sensitive information, the level of the data's sensitivity becomes an important factor. Data of low sensitivity may require a single method of sanitization. Moderately sensitive data may be a candidate for employing two methods. Highly sensitive data may require three methods. If, for example, a PDA contains highly sensitive materials and disposal is necessary, security may clear the device as per the manufacturer's instructions, then overwrite or purge the data prior to submitting the device for physical destruction. This methodology is layered defense at its utmost, but it may be warranted when dealing with high-level sensitive data.

HANDLING THE SANITIZING OF DIFFERENT MEDIA

Paper and microforms require specific destruction procedures. They must be secured in a locked area prior to removal for destruction. If an independent firm does the shredding and/or the burning, the firm must possess appropriate bonding and must be obligated contractually to follow an agreed upon procedure as to the destruction process. Following a recognized standard, such as NIST 800-88, ensures that all parties understand what needs to be done. The NIST standard recommends for paper using a cross-cut shredder that produces pieces 1 × 5 millimeters in size. In the alternative, pulverizing the paper with a machine conforming to the National Security Agency (NSA) pulverizing standard is also acceptable. Microforms should be burned and reduced to white ashes.

Handheld devices include cellular telephones and PDAs. The first step in sanitizing these devices for reuse is to manually delete all sensitive information. This means clearing out all the databases on the machine. While this step can be tedious, it does ensure that no obvious data is available to the next user. Then, contacting the manufacturer is necessary to learn how to do a hard reset of the device back to its factory settings. With today's Web available, most manufacturers have online the user manuals for their electronic products. If you cannot locate the printed user manual, then consult the one online to obtain the details on the hard reset procedure. Destruction of the device includes the options of pulverizing, shredding, crushing or disintegrating, or incineration.

Networking devices such as routers require a manufacturer's hard reset to factory default settings to erase all rules and routing tables. Again, the Web serves as resource for locating user manuals and manufacturer's instructions. The destruction of these devices involves the same options as for handheld devices, which shall be known as the "standard options" for the remainder of this chapter. One may question the level of sensitivity for screening rules and routing tables. Any information on how your network functions can be very valuable to an attacker trying to penetrate the network's perimeter. For a "document" is anything relied upon for information or understanding. A screening router can be a "document' of great import in knowledgeable hands, so treat it with the sensitivity that it deserves prior to reuse or when placing it in unknown hands.

Equipment, such as copy machines and fax machines, has the same reuse and destruction guidelines as for networking devices. These pieces of essential office machinery are often forgotten when it comes to monitoring and to safeguarding for information security. They frequently have large mounts of memory, which may reveal a great deal about your organization's functioning and operations. Again, they like other electronic devices or equipment contain "documents" in the broadest definition of the term. Make sure that they are properly reset before your organization donates or wholesales them. When in doubt, have them destroyed by a bonded vendor.

Magnetic disks like floppies can be overwritten with a program that places 1's and 0's over the existing data. As indicated prior, degaussing with an approved NSA degaussing device ensures correct erasure of the medium. The standard options for destruction are adequate for floppies.

Overwriting ATA hard drives is feasible using similar software. Degaussing, however, is not an option for current hard drives if you plan reuse. Using a degaussing device renders the drive permanently unusable. For purging the contents, NIST 800-88 guide recommends using the "Secure Erase" software available from the University of California at San Diego (UCSD). The standard options for destruction apply to ATA hard drives.

The same sanitizing principles apply to USB removable media that use hard drives. These include thumb drives, pen drives, flash drives, and memory sticks. Purging using the UCSD software is recommended over degaussing if you plan on reuse. The standard options for destruction apply to these drives.

SCSI hard drives can be cleared using the overwriting software. Purging is possible through degaussing, provided reuse is not an issue. Zip drives also can be overwritten, and purging is possible through a degaussing device, provided reuse is not intended. The standard options for destruction apply to SCSI and Zip drives.

Magnetic tapes, whether for data storage on digital tape or for video on the VHS format, present a slightly different sanitization issue. Overwriting may not be practical due to the time involved to pass the tape across the tape head to complete the overwriting process. Re-recording the tape over with a nonsensitive signal may be the best way to clear sensitive data. In order to accomplish this process the same type of machine should be used that did the original recording of sensitive material. Degaussing can also accomplish the clearing action without harming the medium's ability to capture data or images in the future. It does not impair reuse. An NSA approved degaussing device accomplishes purging quite well for magnetic tape. The destruction of magnetic tape involves either shredding or burning.

Optical disks such as CDs and DVDs should be erased using the computer's burn-in software. Check the software manufacturer's documentation on any additional steps necessary to reduce the possibility of optical remanence. It is recommended, however, that due to their low cost and large storage capacity that optical disks be destroyed rather than be reused. Destruction processes include grinding off the surface layer, shredding the disk, and incinerating disk.

The heading, "Memory," includes quite a range of media:

- SD cards
- DRAM (Dynamic Random Access Memory)
- PROM (Programmable Read Only Memory)
- Flash cards
- PCMCIA cards
- USB removable media
- Smart cards

SD cards can be overwritten with new data using a wiping or clearing program. Destruction includes shredding, pulverizing, disintegration, and incineration. Clearing or purging DRAM is done by powering it off and by removing any backup battery. It can be destroyed by shredding, pulverizing, and disintegration. Flash cards can be overwritten with new data using a wiping or clearing program. Destruction includes shredding, pulverizing, and disintegration. PCMCIA cards cannot be cleared or purged, so the only option when removing them from service is destruction via incineration or disintegration. USB removable media can be overwritten with new data using a wiping or clearing program. They can be destroyed by shredding, pulverizing, and disintegration. Finally, smart cards are not designed for reuse. Cut up the internal memory and have the pieces incinerated.

Magnetic cards can be overwritten with new data using a wiping or clearing program. Use a degaussing device to purge the device prior to reuse. The destruction methods are shredding and incineration.

OTHER CONSIDERATIONS

Tom Olzak in his article, “Fundamentals of Storage Media Sanitation,” points out two levels of risk exist for computer media. The first level is the “keyboard recovery.” By entering search commands into the keyboard, a knowledgeable attacker can uncover data thought deleted or erased by the user. In a lab attack, technical methods recover data from storage media. This recovery is possible due to the remanence of data on magnetic media and in memory. The techniques are sophisticated but are well within the technical capabilities of intelligence agencies and law enforcement.

The responsibility for overseeing the sanitation process lies with the data owner. Establishing procedures based upon the sensitivity level of the data, security works with the data owner to ensure that information leakage does not occur. The appropriate action of clearing, purging, or destruction depends upon the guidance approved by the data owner.

Numerous overwrites may be necessary to prevent a keyboard attack. A single pass of overwriting may be ineffective due to the variations in the strength of the bits that end up recorded on the medium. Several passes of overwriting are usually the rule to avoid recovery of the information via the keyboard. A caveat is that on older disks bad sectors may form. Such bad sectors do not receive new data, so overwriting may not touch those sectors. Those sectors may harbor sensitive data. Destruction, rather than reuse, may be the best strategy on older media.

Overwrites should follow a certain sequence. The first pass is a single pattern of 1's and 0's. Then, the second pass is the complement of the first pass: just doing the reverse of the binary pattern. The next pass starts a new binary pattern, and the sequences continue until a sufficient number of overwrites have occurred.

Optical disks, while they can be overwritten, if they are of the read/write variety, may retain data and be prone to lab attacks. Destruction of optical media is best. Even RAM poses some dangers of data remanence. Although powering off RAM normally removes data, if the same data is constantly stored in the same memory area, then some data persistence is possible. If the RAM stores crypto-keys on a regular basis, then it should be destroyed after it useful life in the organization, rather than being reused.

Jason Andress in "Secure Data Deletion and Recovery" (*The ISSA Journal*, January 2007), emphasizes that the first step in document security is to redact metadata, for it often reveals far too much information about the document's creation and content. Oversight in this area is indicative of the "deletion" problem overall, for users only think they get rid of the sensitive data, when, in fact, they do not. As far as magnetic remanence goes, he notes that areas on the medium's surface that are not in the direct drive head path (ridges) can still hold data, even after overwriting. Keyboard attacks are thwarted by extensive overwriting. Hardware methods, however, using a Scanning Tunneling Microscope (STM) or spin stand imaging using an oscilloscope, can still recover sensitive data from the ridge areas. Destruction of the medium or high-level degaussing are the best solutions for preventing information leakage at this granular level.

ESTABLISHING MEDIA SANITATION POLICIES

Without established procedures, computers and their associated media will end up in the trash without being sanitized at the end of their organizational usefulness. Or, they will undergo reuse without the proper cleansing steps. Security should meet with the respective data owners and determine an orderly process for the disposal and the reuse of information assets.

The policy should cover:

1. The classification of sensitive materials and data.
2. Identification of sensitive data on machines, devices, and media.
3. Adoption of a standard such as NIST 800-88 for reuse and destruction methods.
4. Designating staff or vendors responsible for implementing the standard.
5. A checklist for implementing the reuse or destruction procedures.
6. An auditing procedure to ensure adherence to the standard.
7. Examination of media or computers after clearing or purging to ensure no sensitive data remains.
8. Verification of the destruction process.
9. Periodic review of the entire sanitation process to ensure that it remains current and fulfills the organization's security needs.

With regard to USB drives, a security policy and review is also in order. After all, military secrets were found on a flash drive for sale in a bazaar in Bagran, Afghanistan. A University of Kentucky professor had private information on 6,500 former students compromised via a USB drive. Nimrod Reichenberg in "Seven Steps to Secure USB Drives" (*The ISSA Journal*, January 2007) counsels that incidents like these two just described are best prevented through a good USB policy. The main seven points of an action plan to deal with USB problem are:

1. Establish a security policy for USB drives.
2. Allow only company-issued USB drives.
3. Use full encryption.
4. Make sure users cannot circumvent encryption and security policies.
5. Establish an audit trail for practices.
6. Have a data recovery plan.
7. Implement the solutions enterprise-wide.

Table 5.1: Overview of Sanitizing

Topic	Information
Three Types of Sanitizing	• Clearing • Purging • Destroying
Different Types of Media	• Hard Drives (Fixed and Removable) • USB Drives • Floppies (Magnetic) • Memory Cards and Sticks • Optical: DVD, CD-ROM, CD-RW • Magnetic Tape • PROM
Clearing	• Overwriting Data • Manufacturer's Reset
Purging	• Degaussing • Firmware purging
Destroying	• Shredding • Pulverizing • Disintegration • Incineration • Melting
Disposal	Discarding without sanitizing

Chapter 6

SECURING PAPER AND PHYSICAL DOCUMENTS

This chapter examines the security issues involving documents a user can see and touch rather than just existing in electronic form. Hopefully by the chapter's end, security specialists will appreciate that documents fall into several types, each with their own special protection requirements. Paper and physical document security requires the skills of the physical security specialist, of the historian or the detective, and of the media librarian. It can be as interesting and demanding as any efforts to protect electronic data or information in cyberspace. (See Tables 6.1, 6.2.)

DOCUMENT TYPES

Establishing differences between paper and physical documents at first seems an academic exercise. The differences, however, affect how the security professional implements security measures. Paper documents, usually, are reasonably easy to file away and to destroy. File cabinets with suitable locks and a commercial-grade shredder offer a fair amount of security. But, handling physical documents, such as information written on a whiteboard or markings on boxes in inventory, offers different security challenges.

Paper documents contain information written or printed on paper or on a similar material like transparency film or photographic paper. Common forms of these documents include: paper sheets, transparencies, slides, microfilm, microfiche, photographs, labels, and small packaging material. They all share the traits of being readable by the human

Table 6.1: Document Families

Document Types	
Category	**Characteristics**
Paper	On paper or a similar material, readable by the human eye
Physical	On other materials, but still readable by the human eye
Machine-readable	Requires electronic equipment to be read or understood
Examples	
Paper	Paper documents, transparencies, slides, microfilm, microfiche, photographs, labels, small packaging
Physical	Inscriptions, signage, medical imagery such as x-rays, chalkboards, whiteboards, storyboards, some forms of visual art and illustrations, storyboards, boxes and storage containers with exterior markings and information
Machine-readable	Videotape, audio tape, CDs, CD-ROMs, DVDs, computer media, files stored electronically

Table 6.2: Paper and Physical Documents

Paper and Physical Document Security	
Area of Vulnerability	**Countermeasures**
1. Papers thrown into trash	1. Shredding
2. Notes attached to computers and others areas in workspace	2. Institute a clean desk policy, backed up by inspections
3. Sensitive documents stored in unsecured areas	3. Establish a classification system and a set of secure storage procedures
4. Documents awaiting destruction not secured	4. Provide secure storage for documents scheduled for destruction
5. Printers open to everyone in the facility	5. Secured rooms for printers assigned to sensitive documents
6. Fax machines unsecured	6. Place fax machines handling sensitive traffic in secured printer rooms.
7. Information left on chalkboards, easels, wall displays, and whiteboards	7. Locked rooms for work-in-progress and projects requiring long-term wall graphics. Regular meeting rooms with boards should be erased daily.
8. No media library	8. Establish a media library

eye, sometimes with the aid of magnification, being easily filed away, and being reasonably easy to destroy by shredding, by pulverizing, or by burning. (Reading microfilm and microfiche may require a machine, but since such a machine acts as a magnifier and projector, without electronic processing of the information, classifying these forms in the "paper family" is reasonable.)

Physical documents are on materials other than those in the paper category. Again, they are visible to the human eye, albeit at times with some magnification. But, creating security through filing away such items or through easy destruction methods becomes more problematical. Common forms or media for physical documents include: inscriptions, signage, medical imagery such as x-rays, chalkboards, whiteboards, storyboards, some forms of visual art and illustrations, storyboards, boxes and storage containers with exterior markings and information. (Some may argue that transparencies and medical images, like x-rays, are similar in that they are both films. Yet, medical images tend to be larger in size than transparencies and have different storage requirements.)

Machine-readable documents require electronic equipment to be seen or heard. Often, these documents are computer based. They yield no information to the unaided human senses. Common forms for these documents include: videotape, audiotape, CDs, CD-ROMs, DVDs, computer media such as floppies, USB drives, memory, and hard drives, and files stored electronically by other means. (See Table 6.3.)

The question then becomes: What is the security impact of all these distinctions? Each document family has its own protection criteria. The savvy security professional understands the differences in the life cycles of each family. Factors to consider for all document types are:

1. Size of the document
2. Ability to store or to hide the document
3. Ability to destroy it
4. Reuse of the medium
5. Visibility to the human eye
6. Audibility to the human ear
7. Copying issues
8. Clandestine theft of the document
9. Mislabeling issues

Obviously, paper documents are reasonably easy to store and to hide from peering eyes, but boxes containing proprietary ingredients for an

Table 6.3: Document Types

Type	Security issues	Solutions
Paper	1. Intermixing sensitive and non-sensitive documents 2. Easy to copy 3. Easy to read or photograph	• Secure filing and storage • Destruction plan in place • Clean desk policy • Security patrols • Locked secure areas for printers and faxes
Physical	1. Destruction more complex 2. Special manufacturing equipment may need security 3. Higher visibility to outsiders 4. Difficult to classify	• Destruction plan in place • Chalkboard, whiteboard, storyboard policy • Cloaking of sensitive materials • Securing of production equipment
Machine-readable	1. Reuse issues 2. Misplacement and mislabeling 3. Easy to copy without detection 4. Theft issues 5. Specific destruction measures required	• Media library • Follow reuse guidelines (see Chapter 5) • Follow destruction guidelines (see Chapter 5) • Tracking tags or barcodes

industrial process may have markings that are not so easy to hide. Paper documents and photographs generally speak for themselves regarding their content, but a magnetic tape, if mislabeled, may render key data inaccessible. Understanding the security advantages and disadvantages of respective document types becomes essential to any document security program. (See Table 6.4.)

Recognizing the unique characteristics of the respective document types determines how the security specialist will store, retrieve, protect, allow reuse, and destroy documents. The security of any organization is best judged by how it handles its documents. Information is the lifeblood of any twenty-first century business or organization; slipshod document procedures open the concern to corporate espionage, violations of compliance to laws such as Sarbanes-Oxley (SOX), and to litigation by customers and clients for divulging private or proprietary data.

DOING OFFICE AND SITE INSPECTIONS

No document security program will succeed unless security personnel walk through the organization looking for violations of the document

Table 6.4: Document Characteristics

Characteristics	Affected Documents
Size of the document	A consideration for all types
Ability to store or to hide the document	Physical documents present challenges
Ability to destroy it	Physical and machine-readable documents present challenges
Reuse of the medium	Machine-readable documents present unique challenges
Visibility to the human eye	1. Machine-readable documents are opaque without electronic equipment 2. Physical documents may be visible to unauthorized eyes from a distance
Audibility to the human ear	Audio recordings and audio files are mute without electronic equipment
Copying issues	A consideration for all types; copying equipment in close proximity to sensitive documents is a particular concern
Clandestine theft of the document	A consideration for all types; however, paper in bulk or large physical documents may present difficulties for a thief
Mislabeling issues	Machine-readable documents have high risk

security policy. Careful observation uncovers violations of the security policy. Striking a balance between employees being able to personalize their workspaces and maintaining a reasonable clean desk policy is vital to the program's success. Demanding Spartan "monastic cells" (or cubicles) does nothing for morale, but asking that sensitive documents be locked away is reasonable. Desks and work areas need to be neat and orderly enough to permit quick visual inspection for sensitive materials.

Security officers must understand the document security policy. Their training must include how to inspect an area for violations of the policy. If they detect a policy violation, they should know the necessary procedures for handling the violation. An ideal program includes the elements of observation, inspection, protection, and documentation. Observing correctly means visiting all work areas and focusing on document-intensive areas such as fax, copier, and printer workspaces. Inspection requires close examination of monitors, keyboards, computers, and note-posting areas for messages and "sticky notes" containing passwords and login information. Protection means taking possession of sensitive documents and media. Security officers place these sensitive items in a secure area for latter retrieval by the user or owner of the document or media. Documentation is the creation of an incident report of the document seizure by the officer. Leaving a copy of the report at the work area lets the user or owner know where they can retrieve their document. Other copies of the incident report go to Security and to the offending business unit's manager.

Security personnel need to watch out for the following violations:

1. Sticky notes containing passwords and login information. (This is absolute "no-no" and should never be permitted under any circumstances.)
2. Unattended desktop computers that are not password protected. (Security officers need training on how to do the CTRL-ALT-DEL and screen lock command on a Windows system. If UNIX, LINUX, or Mac OS systems are in use, they need training on safe lockdowns of the computer.)
3. Unattended laptops and PDAs that do not have a physical locking device to prevent theft. (These items should be seized on second and third shift to prevent theft.)
4. Documents unsecured that have a sensitive security classification. (See "Classifying Documents" below.)
5. Desks cluttered with documents to the point where a quick visual examination cannot ascertain their security classification. (If a project demands a large number of documents as work-in-progress [WIP], a locked room or controlled access work area is necessary.)
6. Unsecured computer media or other machine-readable media that has a sensitive security classification. (See "Classifying Documents" below.)
7. Chalkboards, whiteboards, storyboards, or other visual displays that violate the document security policy as to duration and content. (If such visual aids are necessary on a continuing basis for a project, they should be in a locked or controlled access area.)
8. Unattended documents left at fax machines, copiers, and printers that violate the security policy as to duration and content.
9. Unsecured locked or controlled access areas.
10. Physical documents that require cloaking or hiding but are visible to public areas or to unauthorized individuals.
11. Documents awaiting destruction but are not secured.

12. Sensitive documents placed in unsecured trash.
13. Unauthorized cameras in sensitive areas. (Employees authorized to use cameras internally should have a special badge or permit.)
14. Unauthorized computer equipment or media in sensitive areas. (Employees authorized to bring in their own digital devices or media should have a special badge or permit.)
15. Unescorted visitors, vendors, and suppliers.

CLASSIFYING DOCUMENTS

Without a classification system, enforcement of document security policies becomes impossible. Unless security personnel can identify a document as being sensitive, they cannot take action to protect it if unsecured. Documents in the "paper family" require a sensitive label stamped or printed on the document's surface. Physical documents may not lend themselves to such marking. Often, these documents assume their classification by other means such as color coding or location. For example, whiteboards containing sensitive data are in a locked room or boxes containing proprietary ingredients are kept in a special secure area. Colored markings further identifies their sensitive nature. Machine-readable media require sensitivity labels, tags or barcodes to identify their classification.

The question then becomes, "What is sensitive?" Numerous classification systems exist for different organizations, and rubber stamping everything "Top Secret" is both impractical and overblown. Each organization must decide what constitutes a sensitive information asset. Government organizations have the advantage of relying upon well-documented classification system for sensitive data.

Business and other private organizations often have a less rigid system. The general classes for the private sector are:

- Confidential
- Private
- Sensitive
- Public

Confidential is the highest security classification in the private sector. For example, trade secrets, proprietary data of the firm or its customers, research and development information, and information required under law to remain confidential fall into this category. Release of this information would cause grave harm to the organization.

Private information involves sensitive operations such as the organization's finances, marketing plans, construction plans, and sales data. Release of this information creates serious harm.

Sensitive information includes employee directories and lists, internal communications, communications with customers and clients, and transactions not meant to be public. Release of this information may cause some harm to the organization or its customers.

Public information may be released to the public without harm to the organization. Data or information classified as public does not mean it gets handed out at the local street corner. Rather, if it does get released, the data does not cause direct harm to the organization.

These general baskets for classifying information contained in documents offer at best broad guidelines. Decision-makers within the organization must determine the security classification of a particular sensitive document. The data owner is primarily responsible for the classification; however, technical assistance may be obtained from the information security specialist where needed.

The criteria for determining the classification include:

1. The data's value
2. How useful the data is
3. The cost of replacing the data
4. Who handles the data
5. The level of damage caused if compromised
6. The level of damage caused if modified or corrupted
7. How is the data stored?

The first step in the process is to identify the categories of documents within the organization that require the confidential, private, and sensitive labels. The respective data owners should supply lists of documents for labeling and the security team should develop procedures for their protection.

Thinking in categories is essential because of the volume of documents produced by an organization. A committee cannot meet and decide on the classification of a document every time one is created. Rather, the role of the document in the organization should be a significant factor in its classification. Some common security-sensitive categories are:

- Intellectual property of high value
- Marketing plans
- Financial statements and reports
- Personnel records
- Research and development documents
- Customer lists
- Audit documents
- Sales data and plans
- Trade secrets

DEVELOPING SECURITY PROCEDURES

Primarily, an organization must have an effective information security policy in place to accomplish a reasonable security level for paper and physical documents. In our ever-increasing Digital Age, much emphasis goes to deterring network intrusion and attacks on databases in electronic format. Yet, an astute information security team must realize that at some point in the life cycle of sensitive information such data may end up in paper or physical form. Therefore, developing a comprehensive information security policy that includes paper and physical documents is essential for layered defense.

For guidance on developing information security policies, the professional should consult the SANS site (http://www.sans.org/resources/policies). This resource offers numerous templates and supporting documents on crafting security policies. In addition, "The SANS Policy Primer" by Michele D. Guel (2001) provides a concise overview on developing policies for protecting information.

Once the information security policy is in place, it serves as a framework on which to build procedures to ensure the policy's general provisions receive proper enforcement. Developing procedures for paper and physical document security require the joint participation of information and physical security professionals along with the input and cooperation of the information assets' owners. In working as a team, the members strive to cover the following principles:

1. Procedures should implement policies in a practical manner. Protecting sensitive paper and physical documents requires care, but the process should not be unduly burdensome. Otherwise, workers will constantly be looking for ways to shortcut the process.

2. Establish procedures for classifying documents as Confidential, Private, or Sensitive. Ensure that appropriate review prevents unnecessary classification. Everything cannot be a vital secret; otherwise, the whole protection process becomes unwieldy.
3. Make sure the classification system is comprehensible to all. The rationale should be easy to understand.
4. Define storage and handling procedures for the classified documents. Specify where, when, and how each type of classified document receives secure handling and storage.
5. Define destruction and reuse procedures for classified documents. (See Chapter 5.)
6. Define reclassification procedures. (A periodic review is necessary to ensure time, resources, and monies are not being devoted to protecting documents whose classification is obsolete.)
7. Define special procedures for physical documents.
8. Define how Security responds to incidents regarding sensitive paper and physical documents. How will the documents be secured? Who does Security notify? What reports does Security generate? Who handles the investigation?

Always allow for revision of procedures as changes in the organization dictate. Tradition is great, but modern security requires adaptability and flexibility. Make sure that your security procedures stay in "synch" with the realities of your work environment and meet on a regular basis with the owners of information assets. Obtain their feedback on procedures and what can be done to improve service.

Enforcing Paper and Physical Document Security

Regular inspections protect documents. If the security officer force must walk the premises on second and third shift, they can keep their eyes open for violations of document security policies and procedures spotting violations and taking corrective action should be a part of security officer training at the facility.

The security team of information and physical security professionals must design a checklist and a report form for document security incidents. In addition, the security team should establish a route for security patrols to ensure they visit all probable "trouble spots" within the facility.

These spots are places where documents tend to accumulate and deserve regular inspection. Ideally, security officers should have electronic watch stations that correspond to these spots along their route.

Key areas to investigate on a daily basis include:

1. Dumpsters and trash bins
2. Shredding collection points
3. Meeting rooms
4. Special project rooms
5. Copier areas
6. FAX areas
7. Printer areas
8. Breakrooms
9. Paper storage areas
10. Storage areas for sensitive physical documents
11. Loading and receiving docks
12. File cabinets and document storage equipment
13. Media Library
14. Document destruction area

Obviously, the security force looks for sensitive documents that are unsecured, unlocked doors to storage areas, excessive accumulations of documents, whiteboards that need to be erased, and the like. Constant vigilance cuts off many avenues of attack by the information thief. To use a buzz word, leveraging the physical security force extends the enforcement capabilities of the information security professional. Make sure that all uniformed security officers have training in securing paper and physical documents.

On the topic of training, the next best technique in guaranteeing a high level of document security is in educating the average worker. A multi-layered approach can have a significant impact. First, all new employees, as a part of the new hire orientation process, need to understand the importance of following document security procedures. Second, educational audits offer constructive criticism if a department has difficulties in implementing procedures. Unscheduled audits, if done in a friendly, nonconfrontational manner, can educate workers. Briefing workers and managers after an audit on the strengths and weaknesses discovered can have a positive impact, and, finally, periodic security training meetings will reinforce the need to protect sensitive paper and physical documents.

The basics of worker training should include:

A. Review of the information security policy.
B. Correct procedures in handling and securing sensitive documents
C. Correct procedures in destroying and disposing of sensitive documents.
D. The rationale of classifying documents.
E. The importance of a clean desk policy.
F. Policing work areas around copiers, fax machines, and printers to prevent excessive and prolonged document accumulation.
G. Why notes on a calendar or sticky notes on a terminal, logons, and access codes in one's workspace are a bad idea.

MEDIA LIBRARY

In a discussion of paper and physical document security, why cover computer and tape media? A document, even though in magnetic or optical format, is in a physical medium. When that medium is portable like a floppy, an optical disk, a USB drive, a magnetic tape cartridge, an audio or video recording, or a removable hard drive, it needs secure storage if it contains sensitive documents. A media library offers that solution.

In the media library, proper marking and labeling procedures insure that all media receive a correct content identification and security classification. Users must sign out for media and sign in upon its return. Ideally, barcode or RFID readers can facilitate the tracking and logging of media.

Physical security for the media library should require a locked area with access card entry by authorized staff. As indicated prior, security patrols need to check the media library to ensure all physical security access is working properly. In addition, the media library staff should have a database by which to track items. The same database serves as an auditing tool to evaluate any unauthorized shrinkage of the inventory.

An operation with a large amount of sensitive data on computer or magnetic media will find a media library to be an effective security control mechanism. It can avoid "lost" or "misplaced media that have created embarrassing headlines at various twenty-first century information businesses in the last decade. Additional protection is available by placing RFID or loss control tags on sensitive media. If the items are removed from the premises, perimeter detection equipment can alert security. Again, no one should be taking sensitive media home.

Chapter 7

FORENSICS

Where is the document? This question is the fundamental one for computer forensics examiners. Every user of electronic documents needs to understand that simply "deleting" a document within its application does not rid the computer of the data. The data becomes less convenient to use, for sure, but it is still there. Overwriting the data with new files is possible, but again, this overwriting may not be perfect, and data fragments may remain. The electronic document examiner possesses the skills and equipment to make data come alive again. Users, who wish to preserve the confidentiality of information, must understand the basics of data resurrection in the hands of a computer forensics expert. This chapter will not make one a computer forensics expert. The discussion will provide, however, an overview of what is possible in the field. Taking into account computer forensics should be a part of document security policy and planning. In Chapter 8 the discussion will cover anti-forensics, the countermeasures to prevent the successful examination of digital data.

FORGOTTEN DATA

The computer forensics examiner does not always have to possess the investigative acumen of Sherlock Holmes. Often, the careless user leaves all the information in the most accessible of places on the computer. These low-lying fruit dwell in the desktop's Recycle Bin, appear on the Document List ("My Recent Documents"), or reside in a common directory like "Documents and Settings" on the hard drive. These areas form the initial stopping points in a forensic sojourn through your computer.

People simply forget that they drop files in these places or that the computer tracks their activities automatically as with "My Recent Documents." A user may move a sensitive document to a USB drive or even create the file on the drive itself. Discovering the existence of the document, however, is not difficult if the user forgets to delete the entry in "My Recent Documents." How often have you "deleted" a document only to discover it a week later in the Recycle Bin? How many times have you saved a file to the wrong directory? Did you later discover this forgotten copy a week or more afterwards? These slips create few problems when dealing with information that is not sensitive. They become costly though when the data is confidential or sensitive.

Whenever one seeks to preserve confidentiality of digital data, a high attention to detail and constant vigilance are necessary, and for many, those levels of effort become difficult prices to pay in guarding confidentiality. Digital data quickly becomes hidden or forgotten and escapes our watchful eye. In an age of multi-tasking and fast-paced lives, the challenges of keeping secrets in the digital realm remain daunting. Any time a user decides to entrust sensitive files or documents to a computer, confidentiality can evaporate instantly without appropriate safeguards. What have I forgotten today? That question remains a troublesome "Iago" taunting us both day and night. Like Shakespeare's character, the question torments our sense of security because all the details are never completely understood. Computers are the ultimate stalkers, recording our every digital move. Unease is something we cannot escape completely with regard to digital data of significant impact if disclosed.

AN ELECTRONIC TRAIL REMAINS

Documents transferred to other media such as USB drives, SD cards, removable drives, or sent to other computers via e-mail or FTP (File Transfer Protocol) leave a trail on your computer. As previously mentioned, that trail may be explicit in "My Recent Documents," or it may be more subtle in the transactions and data created in the transfer. An examiner may not be able to see the file itself, but he or she can tell the file existed on your machine and the circumstances of its creation and transfer. Establishing a destination may also be possible. Internet temporary files reveal Web sites visited and even what the user did on those sites. Cookies identify sites visited and in some cases clues about your visit. E-mails reveal a sizeable amount of transaction history, whom you correspond with, and details about your personal and professional life.

These bits and pieces are easy for a computer forensics examiner to find and to analyze. How one transacts their life in cyberspace tells a great deal about the person: one's financial status, one's spending habits, one's love life, and so on. A single e-mail to a bank may reveal account information, current financial status, the need for a loan, and other details. Documents pertaining to finances may possess highly concentrated information that could unlock many keys to an individual's life. Probably the most vital information about the average American fits on two electronic pages:

1. The basic identifiers of name, address, date of birth, Social Security number.
2. Where one banks.
3. Where one works and the salary.
4. Relatives and associates.
5. Investments.
6. Assets such as a home, boat, vehicles, and even aircraft.
7. Credit affiliations and credit cards.
8. Businesses owned.
9. Professional affiliations.
10. Criminal history.

Give the average user six months to a year with a computer that has a connection to the Internet, and you will be able to examine the hard drive and to obtain answers to questions in most of the information categories listed above. The information may not be in any one location on the hard drive or in any one document or file, but it can still be located in most cases and summarized from the raw data fragments and clues. More information about us passes through our computer than we would like to acknowledge. It is the ultimate vacuum cleaner of facts and factoids about the users that tap away on its keyboard.

DELETED AND HIDDEN FILES

If one runs the inquiry "recovering deleted documents" in a Web search engine like Google, endless pages of hits will appear offering instructions, hints, and products for the reversal of the delete function.

Vast resources exist for undeleting files. Any computer forensics examiner possesses the requisite tools to locate all files on a hard drive, regardless of their naming convention. Most operating systems (OSs) slightly change the name of the file to render it invisible to the OS and permit reuse of the disk space. “Secretfile.doc” may become “!ecretfile.doc,” a trifle to restore for undelete programs.

Hidden files come in several flavors. Usually, the person seeking to hide a file has just the resources of the operating system to accomplish the task. The common technique involves setting the file attribute to “hidden.” In Windows the user just right clicks on the file icon and goes into Properties/General. The General Properties permit checking the “Hidden” attribute box. If the “Folder Options” utility in Control Panel has the radio button pressed for “Do not show hidden files and folders,” then files or folders with the “Hidden” attribute box checked will become invisible to users. This thin veneer of protection disappears when one presses the radio button in the “Folder Options” utility in Control Panel for “Show hidden files and folders.” This elementary form of “security through obscurity” can be effective when someone else is not aware it is being used, but this ploy would not fool a forensics examiner.

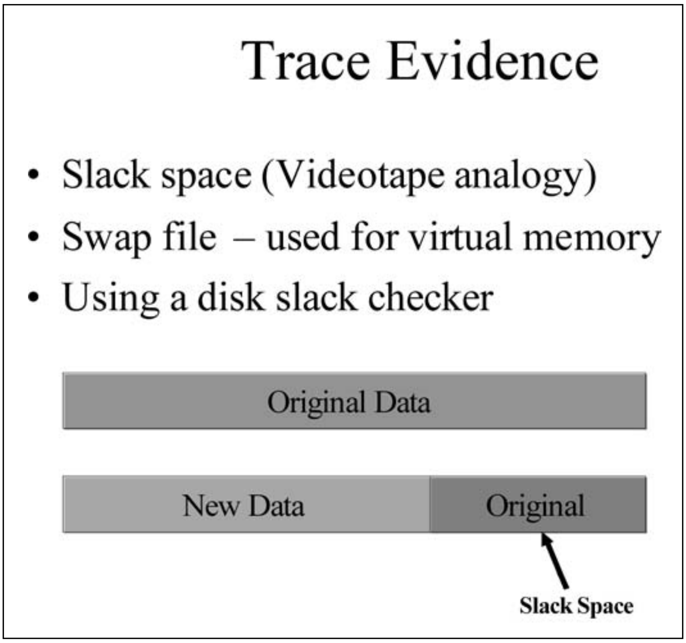
Altering file extensions is another gambit in the “hiding game.” Instead of the file being “jones.doc,” it becomes “jones.mht.” This tactic prevents the file from being visible to the word processing application or any of the normal office suite applications. Images of a sensitive nature can also be renamed to appear to be text files or spreadsheets. Burying the file in a large directory makes it difficult to find, and perhaps, the user wants to mask a file as a spreadsheet and gives the file the “.doc” extension. In any event, this simple ruse causes someone looking for a sensitive file a bit of a challenge. Carrying the deception to an additional level involves renaming the file from say “mysecretfinances.xls” to “aggregate.exe” and then placing the file in a directory filled with executables and program files. Fortunately for forensic examiners, a number of programs are available that examine the contents and structure of a file to uncover its real format despite any name or extension alterations. These forensic applications compile a list of suspicious files for review by the examiner.

Hiding directories themselves or placing files in obscure directories are other tactics employed by users to ward off prying eyes. Again, computer forensics examiners have tools to look for suspicious files in unusual locations on a hard drive. Searching for word patterns and document structure uncovers suspicious files where the content does not match the location and naming of the file. Placing files in the host protected area (HPA) of a hard drive is another hiding technique. Since the HPA is not readable by the computer's BIOS, data stored there may evade routine detection and analysis. Examining the HPA area, however, is now a standard computer forensics practice, so its value as a "stashing" area may be diminished.

Computer forensics is not a static discipline, for new developments occur constantly. At first appearance, data hiding through steganography seems impregnable. Steganography involves hiding one file inside of another. For example, a user has a sensitive text file on his drug-dealing activities and places it inside of a JPEG file, which is a photograph of a scenic vista taken during a vacation. This hiding is possible because not all the elements (bits, a "1" or a "0") of a byte (a group of eight bits) are significant in preserving the image. The steganography program switches out the least significant bit (LSB) in the image's bytes with a bit from the text file. Think of each byte as an egg carton. Steganography removes one "egg" from the carton (one bit from the host file's byte) and replaces it with an "egg" (the replacement bit) from the file the user is trying to hide. Bit by bit the entire sensitive text file becomes hidden in the host or container file.

This technology, however, has a major vulnerability. It requires a specialized software tool to accomplish the replacement process. These steganographic tools like Steganos, Hide and Seek, and White Noise Storm each have a unique signature in their software, which enables computer forensics experts to develop tools to detect them. They resemble malware in the sense of their susceptibility to detection. Once an examiner knows what steganography tool is in use, detecting files embedded with sensitive text or images becomes much easier. Computer forensics has adapted to this challenge.

Trace evidence is the last area of hidden data that needs review. Two activities create trace evidence on the computer. (See Figure 7.1.) First, when files get "deleted" the space becomes available on the hard disk drive again; however, subsequent writing over the space with new data is not always perfect. The new data may not cover the entire original space, so some of the original information remains. This gap between the new and old information is "slack space." Even though fragmentary in nature, this slack space contains valuable information. For example,



Trace Evidence

- Slack space (Videotape analogy)
- Swap file – used for virtual memory
- Using a disk slack checker

Original Data

New Data

Original

Slack Space

Figure 7.1: Trace Evidence

a drug trafficker keeps transactions about cocaine sales on his hard drive. Even finding bits and pieces of transactional information can be quite incriminating especially when words like “coke” and cocaine” keep reappearing. Other sensitive information faces the same scenario as to forensic capture and analysis. It does not take a large amount of information to reveal the larger picture. Computer forensics examiners have a wide range of tools, such as disk slack checkers and HEX editors, to detect “slack space” and to analyze its contents.

The term “slack space” does not refer to a laid-back attitude or a place where everything turns mellow. Rather, the residual data on a storage device, like a hard drive, or in RAM (random access memory) constitutes slack space. Visualize three scribes writing down what you say on chalk slates. Each scribe records a sentence on a rotating basis. The third scribe jots down every third sentence you say. After the speaking goes on for a few minutes, the first two scribes wipe their slates clean. Your next recitation is shorter, so only the first scribes take it down. The writing remaining on the third scribe’s tablet is a data fragment of your original message.

Think of your computer as a collection of scribes that capture your data and keystrokes flawlessly and place the information it sectors on a hard drive, floppy disk, or on other storage media. When you write over existing data, often the new data does not fill up completely all the previously used sectors. Some of the residual data remains in slack space on the storage medium, much like the scribe's tablet that has not been erased.

Since human language and data possess significant redundancy in structure and in repetitive words, such as connectors and common nouns and verbs, fragments can reveal a great deal of information about the original message. Reading a computer's slack space is the stock-in-trade of a computer forensics professional. Such experts take as an article of faith that most users do not realize how unerring the scribes inside a computer can be.

Some of the misconceptions about how computers store information include:

1. Deleting a file obliterates the data. (Unless you take extraordinary measures, much of your data remains available for forensic examination even after deleting and normal overwriting.)
2. "I can hide data on my computer by changing filenames, placing files in obscure directories, or by setting the file attribute to hidden." (Unfortunately, these clever moves will not deter even a minimally competent computer forensics person.)
3. Password protecting a file keeps it secure. (Actually, most passwords are easy to crack. Any conventional password of less than eight characters will succumb eventually to a brute-force attack.)
4. Storing all your secrets on a floppy, a USB drive, or on other removable media automatically protects those secrets. (The truth is mixed. Yes, the specific file contents do not appear on the hard drive within the PC. Data trails, however, pointing to those removable drives will be on the PC's hard drive. Unless you know how to erase those trails, strong clues exist as to what you did. You have to be ready to explain why you wrote data to another device.)
5. "I format the hard drive to erase everything." (Unfortunately, that trick really does not work. Unformat programs can restore the original data.)
6. "I encrypt everything." (Perhaps you do, but is the encryption effective? Encryption offers a fool's paradise unless you understand the technology. More about this issue in Chapter 8.)

The second activity computers do that creates hidden data automatically is the use of swap files. If the available memory in the RAM

(Random Access Memory) on microchips is not sufficient for processing, then the computer creates virtual memory on the hard drive to accommodate the processing's needs. Swap files facilitate the use of this virtual memory; therefore, the swap file on the computer has pages of information recently processed by the machine. Searching swap files is a normal part of the forensic examination of the computer. Data otherwise erased, overwritten, or expunged may be there in the swap file data. It is a place to consider with caution when processing sensitive data.

TECHNIQUES OF COMPUTER FORENSICS

Any forensic examination of a computer or digital device starts with the physical "crime scene analysis." A savvy computer forensics examiner treats the targeted computer as a crime scene with all relevant evidence being preserved as a part of the examination process. Observing the locus is essential to develop clues to solve problems such as:

1. Finding passwords.
2. Locating filenames.
3. Determining URLs for Web sites visited.
4. Identifying e-mail addresses.
5. Identifying associated media such as USB drives, floppies, etc.
6. Locating peripheral devices used with the computer.

Surveying the locus includes looking for notes, written logs, computer media, peripherals, computer connections, content on the computer screen, and items on the computer's desktop. (Users often focus only on the content of electronic documents with regard to protecting confidentiality. Yet, all of the auxiliary information on notes, on paper documents, and on computer media yields secrets too, or the information points to where sensitive information resides.) The forensic examiner exploits this auxiliary information to aid in searching the contents of the computer.

Beyond eyeballing what is around the computer, the computer forensics expert photographs the computer and its immediate environment. Digital photographs, for example, document all the above listed in the initial observation of the locus. Additionally, recording through photography all of the computer connections and even the internal configuration of components with in the computer case may be essential to the investigation. If digital photographs capture critical information, the examiner may take a digital hash of the photograph to establish later that it has not been tampered with by the examiner or by the examination process.

Diagramming the scene may also establish the location of the computer and its connections relative to other components or peripherals. Cataloguing written materials, software, instruction manuals, books and booklets, and associated computer media at the scene will further aid the examiner in understanding the processing done on the computer and will help derive clues as to what to look for in content on the computer. Again, users must understand that many information pieces illustrate what the user has done with files and documents. Mere deletion or even elaborate expunging of data as outlined in Chapter 5 may not completely erase the digital trail left by the user's actions or his or her involvement with sensitive information. Document security must be comprehensive and must address all elements of the information creation and storage process in order to be effective.

Low-level forensic examination of the computer would look for "low-lying fruit." The examiner would check the Recycle Bin, the document list ("My Recent Documents"), common directories such as "Documents and Settings" and "Program Files," Internet temporary files and cookies, Internet browser "Favorites," and any stored e-mails for evidence of the suspected activity by the user or the owner of the computer. Any forensic examination of a computer should be based upon a theory of the case. In other words, the examiner should know what his or her client or agency is looking for on the machine. The suspected activity may be drug dealing, child pornography, organized crime activity or hidden financial records. If the purpose of the examination is to gather business intelligence, the intelligence analyst would have specific information targeted in the search such as proprietary research data, marketing plans, customer lists, and so on. If the evidence discovered is intended for use in the legal system, the examiner needs to be careful not to alter the content of files and directories or to cause "writes" to the computer's data.

Beyond a basic, quick look for "low-lying fruit," the examiner needs to search for text relevant to the theory of the case. Before examining data on the hard drive, however, the forensics expert needs to make a bit stream copy of the targeted disk. He or she copies the drive bit for bit, capturing not only files, but also the slack space and swap file area. For example, "dd" is one bit by bit copying utility used by the computer forensics profession. During the copying procedure, the examiner employs a "no write," usually a hardware device, to prevent any accidental writing to the disk being copied. After the copying is complete, the examiner takes a "hash" of the hard drive and the copy. A hash is a digital fingerprint of the data expressed as a number. The hash acts as a means of demonstrating that the copy's content matches that of the original. It also serves as evidence that the copy's contents have not been altered. (See Figure 7.2.)

The search for relevant text to the investigation occurs on the copy of the hard drive not on the original. By following this standard procedure, the original will not be corrupted by the forensic examination. Text searching includes looking for keywords, words patterns, or by employing regular expressions (REGEX) to pick up common patterns such as Social Security numbers, dates of birth, and drivers' license numbers. Again, the word patterns the examiner looks for are based upon the theory of the case. Word frequency analysis can provide an overview of what is on a disk in terms of subject areas and topics.

Password guarded files need not be an obstacle to the forensics examiner. The cracking of passwords is definitely part of the stock-in-trade of the computer forensics expert. With tools like L0phtCrack and John

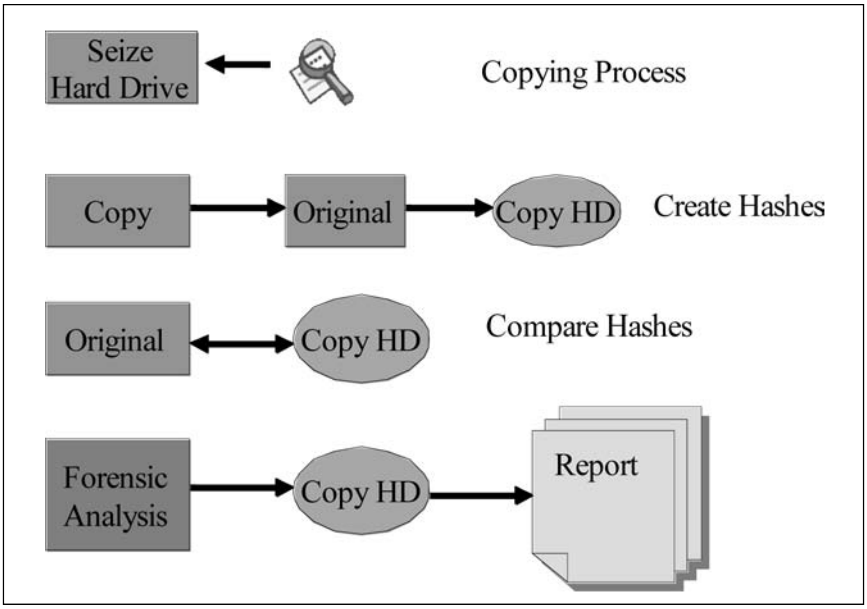


Figure 7.2: Copying

the Ripper, the examiner may attack passwords in one of three ways. (Other password cracking tools include Legion, NTInfoScan, and KerbCrack. These password cracking tools are available on the Internet through a Web search using Google or a similar search engine.) First, a brute force attack simply tries combinations of letters and characters until it hits upon the correct password. Brute force attacks work due to the weakness of most passwords. Passwords are usually too short, less than eight characters. Many are regular words containing only letters. More robust passwords, however, are eight characters or longer and have a mix of letters (upper and lowercase), numbers, and other characters.

The question arises then, what do we make of a password like "DallasCowboys2007!"? At first look, it appears robust, for it is reasonably long, it has a mix of upper and lowercase letters, and it has numbers and a nonalphanumeric character in the password. Yet, the password is too recognizable and not random. The second method of attack, known as a dictionary attack, could exploit this password's design and crack it fairly quickly. A dictionary attack comes in many flavors. On the Internet many different types of dictionaries are available for downloading. These dictionaries can be for a given natural language like English, French, Spanish, and so on, or they can be specialized covering areas of knowledge such as American football, Classical mythology, popular music, Star Trek®, and many other areas of lore. Combined with the ability to create combinations and variations from the base or specialized dictionary, the hybrid attack, which is the third method, is quite effective in producing password matches for the likes of "DallasCowboys2007!" Hybrid attacks reveal the principle that the more an investigator knows about the user, the more effective the attack on that user's passwords can be. Knowing one's hobbies, interests, relatives, associations, and basic identifiers such as date of birth will help discover passwords by knowing which dictionaries to use in the attack. People have an "infosphere" about them of information relations that they are comfortable with. This infosphere, when properly exploited, provides the key to sensitive documents.

Searching for images is another stock-in-trade activity for the computer forensics examiner. The tools include a thumbnail viewer to visualize images quickly in a picture sorter format, a forensic suite like EnCase to view images, and an extension searching tool to discover images renamed to other formats like text to hide the their true content. Current technology permits the location of images fairly quickly on a hard drive; however, eyeballing them all still falls to the examiner. In addition to images, the examiner will look for steganography tools on the computer, as this information will aid to detecting host files. Recent developments in technology also permit examining the frequency distribution of colors in suspected images as a detection tool for steganography.

The examiner will also look for deleted files and will restore them with an undelete program. The examination will proceed with locating hidden information on the computer with the previously discussed file extension searching program, a tree searching program to examine directories, a slack space detection program, and by setting the computer's control panel so that files with a hidden attribute become visible.

If the computer has stored e-mails, the examiner will search the headers for those e-mails to gather IP addresses and e-mail addresses. That information is useful for cross-referencing against data found in Internet temporary files, cookies, browser favorites, and other URLs found on the machine, and of course, the content of the e-mails themselves may offer key information to the investigation.

EXAMINING PDAS AND OTHER MOBILE DEVICES

When thinking of document security, the information security professional must go beyond considering only files residing on a desktop computer or on a server. Digital information has become mobile on a wide range of digital devices including PDAs, cellular telephones, laptops, and other personal information tools such as digital watches that store data. Mobile digital devices due to their compactness and in some cases special operating systems receive special focus by the computer forensics community.

Some of the tools evolved to examine PDAs include:

1. Palm dd or "pdd," which examiners use on Palm OS (operating system) devices to do a bit-by-bit data acquisition from the data storage on the device.
2. Pilot-link retrieves RAM (Random Access Memory) and ROM (Read Only Memory) from the device. The RAM is the work area for processing on the device. ROM contains permanent software used by the device. ROM may be on a removable card so that the user may switch out different programming packages.
3. POSE, which is a Palm OS emulator, allows an examiner to emulate a Palm OS device on a PC. It is mainly used for

demonstrations to other investigators or in court. An examiner can make a bit stream copy of the device's data and then run it on POSE.

4. PDA Seizure, which is Paraben's forensic software tool for PDAs that permits capturing and analyzing PDA data.
5. EnCase, a full forensic suite of tools, has support for Palm OS devices.
6. The duplicate disk utility, "dd," allows binary data dumps from PDA devices.

Digital devices like PDAs and cellular telephones have a wealth of data on them that can be of great advantage to criminals and industrial spies. PDAs may contain extensive contact information, sensitive business files, and personal financial information. Many of these palm-size computers contain software comparable to the word processing and spreadsheet programs of desktops and laptops. Cellular telephones today are not far behind PDAs in terms of their capability. They can contain not only telephone records but also Internet capability and text messaging.

Table 7.1: Computer Forensics Tools

Tool	Description
Coroner's Toolkit	For analyzing Unix systems. http://www.porcupine.org/forensics/tct.html
dd	Duplicate disk utility for binary data dumps from a device or a computer
dtsearch Desktop	Text-searching tool. http://www.dtsearch.com
EnCase	A full-featured forensics suite for the Windows environment, which accommodates Palm OS examination. http://www.encase.com
Forensics Tool Kit (FTK)	Disk search capabilities and e-mail analysis. http://www.accessdata.com
PDA Seizure	A Palm OS forensics tool by Paraben. http://www.paraben-forensics.com
pdd	A bit stream copying program for Palm OS devices. http://www.atstake.com/research/tools
WinHex	A disk editor to search hard drives, memory cards, CDs, DVDs, USB drives, and so on. http://www.sf-soft.de/winhex

The process for forensic examination of these devices is quite similar to that for desktops. Specialized software is available to accommodate the differences in operating systems, but fundamentally, the process is the same. The examiner makes a bit stream copy of the device's data and takes a hash of it to authenticate its completeness and accuracy. Then, the examiner uses the various forensic suites to analyze the data. If need be, he or she then runs the data through an OS emulator to demonstrate its content to others.

THE FORENSIC CHARACTERISTICS OF ELECTRONIC DOCUMENTS

As indicated in Chapter 1, Metadata is one of the chief characteristics of electronic documents. Many users and authors forget the circumstances of creating and editing a sensitive document, and in most cases, the circumstances recorded in the metadata give powerful clues about the contents of the document. The first step in guarding electronic documents is to ensure that no unwanted metadata passes through in the published electronic document. Otherwise, computer forensics people will find some answers to their questions there with minimal effort on their part.

Form is another vital characteristic of electronic documents. The structure of the document itself speaks powerfully: a word processing document has an entirely different structure than a spreadsheet, for example. No matter what name changing occurs, the structure is a giveaway as to the document's actual content. Key words and patterns of words are other signals as to a document's true content and purpose. It may be very difficult to visually look at an entire hard disk's data, but keyword searches make that process manageable. Word frequency is another sign of what is really in a document: marketing plans will have a higher frequency of certain words than say a business report from the facilities department. A computer program written in "C" will have a different form from one written in Java. Form tells a great deal to computer forensics programs.

Searchability is another strong characteristic of electronic documents. Searching for specific text, word patterns, regular expressions, and image patterns are all feasible given current technology. Wild cards permit searching when one only knows a portion of the keyword or if one wants to see all examples of a given pattern. Security by obscurity does not work with electronic documents. If adversaries know what they are looking for, they will find the sensitive text or image, even if you bury it in the middle of a boring text or in some obscure directory.

History speaks from electronic documents. Metadata records the authorship and revision lineage of a document. The Internet browser records Web sites and pages visited. Logs and event monitoring software track activity on the computer. E-mails in their header information trace the route an e-mail took over the Internet. The modern digital device is a recorder of all the events and transactions regarding its users.

Location is the last major factor to consider regarding electronic documents. They are easily moved from one electronic location to another. Users can forget where they have left sensitive files. Sensitive documents can end up misplaced, out of their proper security zone as discussed in Chapters 2 and 3. They can be copied without the knowledge of the owner and the compromise can remain unknown for long periods of time. Corporate spies exploit the portability of electronic information and the sometimes carelessness that results from that portability. Forensic computer professionals find documents and clues pointing to documents in all the places that users forgot to remember to clean up.

Chapter 8

ANTI-FORENSICS

Those who seek to block computer forensics efforts fall into four categories. First, there are individuals and organizations with legitimate security concerns. They are interested in protecting personal, governmental, or proprietary documents or information. Such individuals and organizations seek confidentiality as a socially acceptable goal. Second, criminals obviously do not want their electronic records and information pertaining to their activities to undergo computer forensic examination. They have a socially and legally unacceptable goal of keeping their criminal activities hidden and undetected.

Third, political dissidents against an authoritarian regime seek to protect their activities from scrutiny by those in power. The rebellious elements in society are viewed with disdain by those in power. In authoritarian or totalitarian societies, the rebellious members receive the label of criminals, which results in imprisonment or harsher punishments. True, the moniker of freedom fighter or dissident is a relative one, but in repressive societies keeping secrets is a justified survival mechanism in many cases. And fourth, intelligence agents, whether corporate or governmental, seek to keep their sensitive information from counterintelligence agents that employ computer forensics as tool for detection of their activities. The ethical consequences of the intelligence business would be a long discussion, but without taking a stand for or against this activity, one has to concede that protecting information's confidentiality is its utmost concern.

ENCRYPTION

The primary defense against forensics is encryption. Robust encryption prevents someone from seeing your sensitive information. What constitutes robust encryption? Using an encryption algorithm that has undergone public testing and analysis is best. The algorithm should be public, because after public scrutiny, users can trust the encryption to be strong and resistant to attack. The other side of encryption is using a strong key. Long keys are hard to break because the keyspace is so large that brute forcing the key would require years if not decades. A 128-bit key would be a bare minimum, but a 256-bit affords even greater protection. A tested algorithm and strong keys are the two pillars of good encryption.

What should one encrypt? It sounds like a simple question, but there are options to consider. First, user acceptance is always crucial to obtaining good cryptographic security. If encryption procedures are too cumbersome, users will avoid them. So, encryption needs to be as transparent to the user as possible. Second, choosing which documents and files to encrypt can be challenging one for a user to decide. A sound document classification policy needs to be in place to make encryption decisions straightforward and clear to all users. Third, should you encrypt files or drives? The safest solution is to encrypt the entire drive. In other words, every time the user saves information to the hard drive, the information becomes encrypted. The process for this type of encryption usually is fairly transparent to the user. There are not any steps for the user to do. Encryption is automatic. In addition, having all the data on the drive encrypted prevents slack space and metadata analysis of the drive. That level of protection can be a real bonus because "data about your data" exists on other places on your hard drive. Full encryption blocks that window of opportunity for the forensics document examiner.

Unfortunately for users, pitfalls remain even with full encryption of the disk. In order to access the disk, the user must still employ a key, password, or passphrase. Security for these data structures is vital for preserving security. Storage must be robust and be resistant to both information-based and physical attacks. Keys and passwords need storing on a database that is itself protected by a robust password. The database is also physically secured, whether it is on a USB drive or on a server. Lock up those assets when they are not in use, or in the case of a server, have it in a locked room. Leaving notes about the work area containing the passwords is foolhardy to say the least. Careful handling of cryptographic tools ensures security for sensitive documents.

The final issue in cryptographic security is the red flag it sends up for certain users. Someone with a legitimate business use for cryptography does not have to worry about creating any red flags. However, those individuals engaging in questionable activities, some of which may be illegal, are definitely drawing attention to themselves by using encryption. If their information assets come under scrutiny, then they will have to answer the questions regarding their use of cryptographic tools. The discovering of encryption technology use by suspicious individuals is a strong sign that they are covering up activities. Again, once the tools become identified, cracking the encryption often rests upon having as much intelligence on the individuals as possible. Such intelligence may lead to breaking passwords; as people are creatures of habit and convenience, their passwords may arise from their personal knowledge and backgrounds.

Since robust encryption creates serious barriers to computer forensics, a short recap of the safeguards is in order:

1. Choose an encryption tool that has undergone robust public testing and commentary. You want the algorithm to be a public one, not proprietary. The product you purchase should state the name of the algorithm used, so you can research it on the Internet.
2. Use strong passwords and keys according to the encryption software manufacturer's recommendations. Consider the life of the sensitive information you are protecting. The longer the sensitive document's information will be of value to an adversary, the stronger the keys and passwords need to be.
3. Especially on mobile digital devices, use hard drive encryption. Encrypt the entire drive. This principle also holds for USB drives, memory cards, and flash drives if they contain sensitive information. The same policy should apply to CDs and DVDs holding similar information.
4. Make sure the encryption tool is user-friendly and easy to operate. The ideal arrangement is complete transparency to the user. Information and files automatically become encrypted as the user saves the data.
5. Have a security policy for cryptographic key management and for passwords used by cryptographic tools. Make sure the policy covers the complete life cycle of these data structures from creation to secure storage and distribution to destruction.
6. Put into place auditing controls and supervision of the cryptographic tools and their implementation to ensure conformance to policies.

7. Create physical security policies to protect servers, desktop computers, and mobile digital equipment from theft and compromise. (See Chapter 4 for recommended safeguards regarding mobile equipment.)

THINKING ABOUT YOUR COMPUTER

Before placing sensitive information on a computer system, several pieces of information provide the margin of secrecy. First, how does your computer process information when you create it in an application? For example, you write a letter on a controversial subject, say something critical about your employer. How will your PC store that information? If you delete the file later, does that action remove all traces of the data?

Many users hold misconceptions about the nature of file erasure. They assume pressing one button fixes everything. Frequently, they confuse the concepts of wiping, erasing, and deleting. Deleting a file removes it from the visibility of the operating system. Usually the first character of the file name changes in a simple deletion. For example, if the filename is "criticizeboss.doc" the deleted file name becomes "?riticizeboss.doc". It becomes unreadable by the operating system's normal file detecting process. The operating system becomes blind to the file's existence. All of the data within the file still resides on the hard drive or other storage media. That data remains indefinitely, until written over by other data or removed by other means.

Erasing involves removing the actual contents of a file from the storage media. If you erase a file, the data cannot be recovered from a hard drive or floppy by ordinary techniques. Using a special utility program like a hex editor will not uncover the data.

Wiping the file responds to the physical factor known as magnetic remanence. Usually, the removal of the magnetic flux in the storage medium during erasing is not perfect. With specialized equipment a computer forensics specialist can recover erased data. Wiping, by using numerous overwriting and degaussing techniques, effectively removes most of the remaining magnetic flux of the original data. (See also Chapter 5.)

Degaussing is a fancy term for removing the magnetism from an object by applying an electrical field. Overwriting simply means placing 1's and 0's on the medium many times (usually one hundred or more), and raising the temperature of the magnetic medium to a certain level, known as the Curie temperature, removes all magnetic flux. As a practical matter, data owners rarely use heating the data medium like a hard drive or floppy. The common protocol is overwriting followed by degaussing. A wiping program does the overwriting procedure. The degaussing step occurs by passing the medium through a device that generates the electrical field. Magnetic flux is the measure of the number of lines of magnetic force in storage media. Neutralizing a magnetic field in a storage medium using a degaussing tool may not remove all of the flux. Sophisticated detection equipment may still recover information from the disk. If data overwriting is not used, heating the disk to its Curie temperature or burning the disk insures the total destruction of magnetic flux.

To draw an analogy, "deleting" places the unwanted quilt into the attic, where it may be found in the future only by chance. Erasing means cutting out the quilt squares. Wiping the data involves giving the quilt a long bath in industrial strength bleach.

Removing data from computer media requires deliberate action and special effort. It can be a hassle. For example, if you go to the local stench-ridden trash disposal site armed with a shovel and a facemask, some archaeology on your part will yield "tossed" floppy disks. After you clean the filth off the case, opening the plastic exterior case requires only an ordinary knife to separate the two plastic halves.

Inside the case, a brownish thin disk of magnetically sensitive film with a silver-colored metal button disk in the center will be visible. Washing that disk with household detergents (noncorrosive and nonabrasive) will remove any dirt or dust. Placing the cleaned dry disk into a new case will enable you to run the disk in a drive. Almost invariably, you will be able to read the contents of that disk. Yet, its previous owners thought the disposal of the disk was "secure." Physically cutting up a disk may not render it unreadable. Forensic labs can reassemble the segments (if large enough) of the magnetic film and read the disk. Careless disposal of floppies, hard drives, and other computer media only creates a silent witness to knowledgeable investigators. Wipe then burn magnetic media. Crush and burn optical media like writeable CD-ROMS and DVDs. When you do less than those steps, treasures may be available to prying eyes.

Remember you are the hunted. Every game embracing the hunt theme happens at the same time here. The digital equivalent of paintball hunters, laser tag aficionados, and all the ghouls and zombies populating electronic actions games at the local arcade seek your data. Digital detectives are tenacious and resourceful. You always have to be vigilant.

Before discussing the specifics of how your computer processes data, let's first explore its character. What are your specific goals in securing data? Protecting information on your computer eats time and challenges the endurance of the Biblical character Job, and it can make you paranoid. Did I follow all the steps to wipe that file? Did I check all the possible locations on my machine? If you store everything on your computer out of habit, kill the habit. Clear goals on what needs safeguarding become essential. Otherwise, the fruit of your unfocused security efforts will drive your fear with a mental cattle prod of endless questions. Formulate those goals by asking the following:

1. What is the worst that can happen if you have sensitive information compromised?
2. Can you rate the sensitive information on a scale based upon potential damage? (For example, a low level of damage is a "1", and grave damage is a "10.")
3. For the information that causes significant damage, do another rating based upon damage control. (An "A" rating means you can manage any harm fairly well, but a "D" rating indicates an uncontrollable aftermath.)
4. Examine your "10D" combined ratings. Are you being realistic in your ratings? For example, losing your driver's license from your wallet can be a hassle, and yes, it may lead to identity theft. Exerting a high level of control, however, is possible through the issuing state agency. The agency can generate a license with a new number and note on its database the loss of the other license. That category of information, easy-to-fix losses, still requires some security but not extreme security.
5. What threats against your highly sensitive information are realistic? Will disclosure cause serious harm or merely just embarrassment? If you are a movie star with a strong heterosexual image, the disclosure of your secret gay life may cause serious damage to your career. Gossip journalists are a real threat, but, if as an ordinary citizen, you live a secret gay lifestyle, your neighbors may not relish the fact, the local clergy may frown about it, but the press will not be frenzied if you are gay.

Create a threat table or matrix for your particular situation. (See Table 8.1.) The first column describes the sensitive information. In the second column, give the information its combined rating. In the third

Table 8.1: Threat Matrix Example

Sensitive Information	Rating	Threats
Customer list for your business	7C	Competitors want this list.
Personal credit information	8B	Identity thieves
Details of previous divorce settlement	9C	Current domestic problems

column, list the realistic threats. This tool should provide an objective measuring stick on what information in your life or business is truly a "secret" worthy of special protection.

What information should you not place on your computer regardless of who you are? Consider the following taboo:

1. Social security number.
2. Date of birth, place of birth.
3. Other key identifiers like your driver's license number or medical insurance ID number.
4. Account numbers at financial institutions.
5. Credit card account numbers.
6. Passwords to enter financial accounts.
7. Detailed financial transactions.
8. Detailed medical information.

Software is available on the market that manages all your financial information. The same is available for medical and health data. Unless you have a computer that is physically secure in a locked room and is not connected to the Internet, do not use those programs. Even with those safeguards, always encrypt such information completely. When you go "digital" with those details, you provide an information thief a neat little package. While it may offer convenience, it is a bad move from a security standpoint. Later in the chapter we'll have some tips on alternative storage measures. Yet, keep in mind that whenever you hear someone boast about their whole life or business being on their PC, take a moment to reflect. Anyone loses control when they put their whole life on a PC, personal digital assistant (PDA), or cellular telephone. Concentrated information makes you easy prey. Never transmit any of that data on the "taboo list" through a terminal in a cyber café or at a public business center. These items made the taboo list because if someone else gets their hands on them they can own your life or exploit your business. Be respectful of their power.

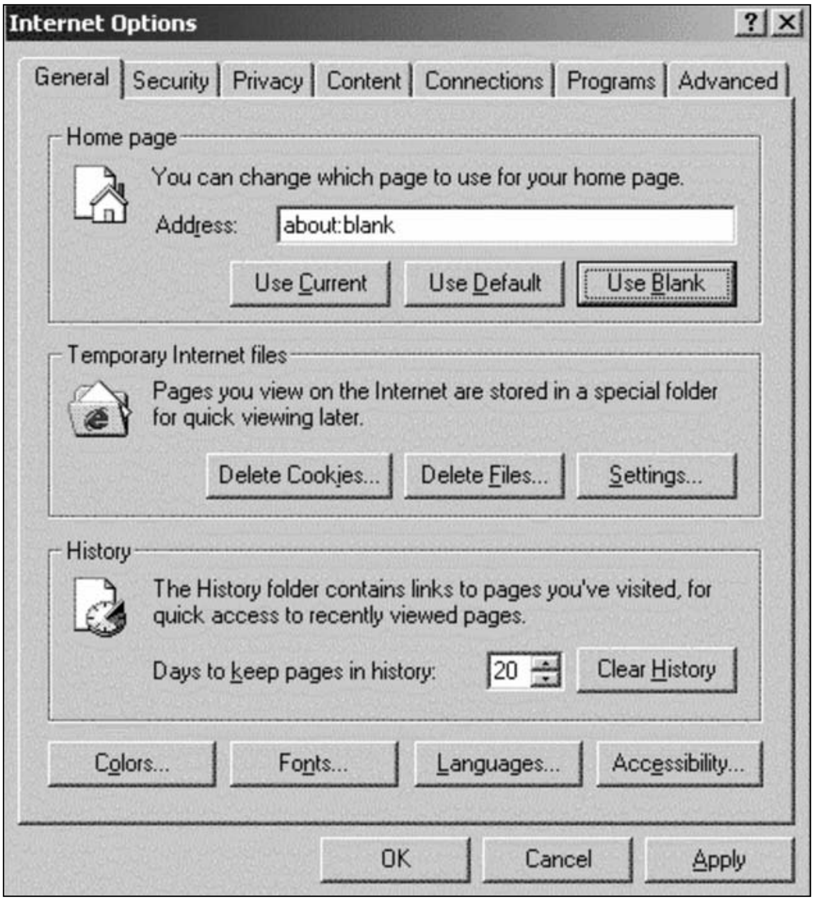
Now, let's get back to learning how your computer processes files. As we now know, deleting does not destroy files. Otherwise, operating systems would not offer an "Undelete" command. Formatting does not wipe all the data off a disk. Otherwise, the "Unformat" command would not be available on many operating systems. If you really want to see what is on a disk beyond what the operating system reveals, use a viewer that shows ASCII text and the corresponding hexadecimal code. You'll begin then to understand the computer as a silent witness. ASCII stands for American Standard Code for Information Interchange. An ASCII character is any letter or number that can be typed from the keyboard, including spaces, punctuation, and special characters like the "@" sign. These characters lack any word processing formatting, such as bold, Italics, or font styles, and they have a hexadecimal counterpart. Hexadecimal numbers are the base 16 number system. The "@" sign is 0100 0000 as a binary, which translates to 0x40 in hexadecimal, a shorthand for the binary number. When viewing the hard drive with a hex editor, the user sees a split screen, where one side shows the "@" character in the ASCII text and the other side shows the corresponding 0×40, representing the actual 1's and 0's on the hard drive. "Hex" numbers are less cumbersome to work with and to view than binary numbers, because they are shorter.

Viewing data at that level is like removing a top coat of paint with a solvent and discovering the painting hidden underneath the covering layer. A hex editor or utilities program weaves a tale of data created in the past, which forms the painting's story. A palimpsest comes to life.

The silent witness also operates at a higher level of the operating system's desktop. A computer forensics expert will check the "Trash Can" or "Recycle Bin" on your desktop. Many users forget what they have moved to the trash. Another area experts examine is the folder containing "cookies." Unlike the baked goods you savored as a child, "cookies" contain data, not white fillings. When you visit most Web sites, the site places a small piece of text on your computer to identify you when your computer returns to the same site. Cookies act as a trail of where you went on the Web. Deleting cookies only requires pressing on the "Delete Cookies" button in "Internet Options." (See Figure 8.1.) A police officer at a seminar I attended on computer crime commented that he had broken many cases just based upon the files and data recovered from the desktop trash and details from the cookies folder.

Most browsers create a record of the sites visited called "History." This history again reveals the sites on your Web itinerary and provides access even to the individual pages. Browsers also store individual pages as temporary files. Unfortunately, many users do not know how to clear their cache and delete their history. What is cache on your browser? How do you clear it? How do you clear the history?

Cache is simply a storage area for temporary files created from Web pages a user visits. If you fail to clean out this cache after surfing the Net, others can see the pages you visited. For example, to clear the cache in the Windows Internet Explorer (IE) browser, you press on the top toolbar item, "Tools" and then press on "Internet Options." Under the "General" tab for "Internet Options" (See Figure 8.1.) you press on



Internet Options

Home page

You can change which page to use for your home page.

Address: about:blank

Temporary Internet files

Pages you view on the Internet are stored in a special folder for quick viewing later.

History

The History folder contains links to pages you've visited, for quick access to recently viewed pages.

Days to keep pages in history: 20

Figure 8.1: Internet Options

"Delete Files" within the "Temporary Internet Files" area. Also, be sure to check the box "Delete all offline content" before you press the "OK" button. You have now successfully cleared your browser's cache.

To clear the history, press on the "Clear History" button in the History section in Internet Options. Newer versions of Internet Explorer® (IE) permit clearing all elements of the cache (history, cookies, and temporary pages) in one action, but the principle remains the same.

When writing text documents in word processing applications, the program may generate temporary files without any intervention by the user. Even if you store this document directly onto a floppy or a removable drive, a temporary file may still save to the hard disk. This behind-the-scenes operation can be very dangerous when trying to keep sensitive data secret. Only a close examination of the hard drive at the ASCII and hexadecimal level will reveal this fugitive data is in hiding.

If the main plot derives from the major programs available on a machine, the back story depends upon the operating system's hidden abilities. In learning about the operating system, the computer's general housekeeper and manager, focus on two themes. Get out the manual that comes with the computer. (It may be on a CD-ROM or built within the computer's "Help" software.) The first concept to research describes the different privilege levels available to users. Current operating systems like Windows XP® or Vista® allow configuration management. That term is security parlance for being able to allow different levels of access.

When you first use your computer, it generally defaults to administrator mode. As the computer's administrator, you have the highest level of privileges. (In the Linux/Unix world, this level of privilege would be "root" access.) By conducting all of your business as administrator or root, you allow anyone who can sign on as administrator to anything on the machine. As a sound security practice, you want to operate your machine with the *least* amount of privileges as possible.

In other words, two roles should exist on your PC. When you make major changes on the machine such as adding a new program, you log in as the administrator or as root. Any other time, you want to run the PC as a simple user. That user would be able to run and access only certain files and programs. Access to the computer's registry, the set of internal rules by which it operates, for example, would not be allowed. If you take your automobile for an oil change, is it wise to give the mechanic all your keys? A prudent customer gives up the ignition key only. This key allows just enough access to get the job done. The trunk key, the house keys, and other personal keys stay on your person. The single key represents limited user privileges, the whole key ring, administrator privileges.

Why create this hassle? The only danger to sensitive information on your computer is not just someone sitting at the keyboard trying to hack in. Attacks at a distance via the Internet are a real possibility. If your PC has an Internet connection, someone can send a Trojan horse via an e-mail attachment to your system. A Trojan horse, like in the Greek myth from the Trojan War, masquerades as something benign, a game, an image, or an application. Once inside your machine, the Trojan horse writes its own ticket. Unknown to you, it transmits your sensitive information over the Internet to others.

When you protect your administrator account with a long, robust password and run your machine as an ordinary user, you limit what a Trojan horse has access to. That user level, of course, has to possess very limited privileges if the defense is to work. In addition to guarding against offsite attacks by Trojan horses and viruses, least privilege makes an attack from the keyboard much harder. Unless the attacker can bypass the protection of the administrator account, he or she may only crack the lower level user account, and that account won't let him or her see much. Learning about setting different levels of access should be at the top of your list.

The second theme or concept you need to research is the operating system's file system. Do you have the option of using NTFS over FAT32? These "techie" terms stand for different approaches to file access and organization. For the moment, let's say that NTFS offers greater security than FAT32. With a FAT32 file system in place, it is much easier to bypass system security. Pay attention to what your manual has to say about NTFS. Other topics to research in your manual are references to "swap files," to log files, and temporary files.

FAT stands for File Allocation Table. It is a roadmap to the location of all the data on a disk. A FAT32 file system has a 32-byte data structure giving the name of the file and its location. The maximum filename in FAT32 format is eleven (11) characters in an 8.3 format, for example "filename.doc." NTFS is NT File System named after Windows NT. It allows filenames up 254 characters long. The NTFS uses a different, smarter allocation program. It also has an elaborate permission structure, which makes NTFS far more secure that FAT32. In addition, someone can boot your computer from a floppy or USB drive to bypass a password login if the computer uses a FAT32 file system. An intruder cannot do that with the NTFS system installed.

As indicated in Chapter 7, passwords are the main defensive tool for keeping unauthorized persons out of your computer. Strong passwords make an intruder's job difficult and sometimes nearly impossible. Weak passwords make the job of intrusion child's play.

Passwords need to be long *and* random. For example, "DALLAS-COWBOYSAREGREAT" is a long password, but it is not a random one. A better password would be "dAllAs1456cowBoyZ." The second password illustrates a few general principles of good passwords:

1. Use a mix of alphabetical and numerical characters.
2. Use a mix of upper and lower case letters.
3. Never use a word or phrase found in any dictionary.
4. Use a password over twelve (12) characters.
5. Even a specialized dictionary of Dallas Cowboys' football terms would not contain this "word."

Numerous password-cracking tools are available on the Internet. (Check out http://www.atstake.com/research/lc3/ the L0phtcrack site to see an example.) When you create short, easy-to-guess passwords, the chances of these cracking tools working on your machine become high. These tools use brute force dictionary attacks. A dictionary attack works in one or two ways. A limited dictionary attack uses words or phrases that you might commonly employ based upon your background. For example, if you are a science fiction fan, the dictionary would contain terms and words frequently found in science fiction stories. A general dictionary attack would use the contents of a standard collegiate dictionary. The attacks are called "brute force" because the cracking program tries one word after the other in the dictionary until it finds the one that works. Someone with physical access to your machine can employ a cracking tool directly from the keyboard with help from the floppy or CD-ROM drives. In addition to creating good passwords, establishing strong physical security for your machine is vital.

AN EXAMPLE: THE SCARFO CASE

Laws pertaining to privacy do not evolve usually from actions by pillars of the community. Sometimes, activities by questionable individuals bring privacy and constitutional issues before the courts. This recent case before the federal court speaks to the difficulties of trying to protect sensitive information, and in this case, evidence of illegal activity.

According to "Big Brother and the Bookie" by George Anastasia in *Mother Jones* (Jan-Feb 2002), Nicky Scarfo considered himself to be a computer savvy mobster who kept his bookmaking records on his PC. He protected those records by using Pretty Good Privacy (PGP) with its 128-bit encryption. (The term "128-bit" refers to the length of the encryption key. Cryptographers recognize that the longer the key, the harder it is to break. A key 128 bits of data long possesses robustness and strength on a grand scale.) When the FBI seized his bookmaking file, they could not crack the PGP encryption. So, pursuant to a court order allowing surreptitious entry, the FBI penetrated Scarfo's office and placed a keystroke recorder in his personal computer's (PC) keyboard. Eventually, Scarfo's keystrokes yielded the sought-after pass phrase.

Security is a people business, an enterprise dependent upon human psychology and perceptions. Misconceptions and narrow-focused thinking affect security outcomes. The craft of security relies on more than employing a simple set of technical protocols or methods. Countermeasures are more than cookie-cutter solutions applied generically where as the cliché goes "one fix cures all."

Nicky Scarfo's security plan had several flaws. First, he assumed that some familiarity with computing made one a computer expert. He had enough knowledge to make decisions about protecting his system, even though his adversary was the FBI. In addition to using computers in his "business," Scarfo did a stint working at a friend's software company. Feeling at ease in front of a PC may not enable one to run a marathon against computer forensics experts with extensive resources.

Storing records of illegal activity on in-house computers was not a wise strategy for this Mafioso. He did not develop a plan for offsite storage at a location that could not be tied to him or his operations. Employing paper storage at a third-party site or using removable drives or storing sensitive data on Universal Serial Bus (USB) "cigar" or stick drives were all viable alternatives.

Second, Scarfo assumed that encryption alone afforded total security. True, 128-bit PGP encryption affords a high level of protection, but as the noted encryption expert, Bruce Schneier, counsels in a white paper on his Counterpane Web site (http://www.counterpane.com/whycrypto.html), encryption does not guarantee complete or absolute security. For security is not just an engineering problem according to Schneier. User behavior often betrays encryption. In Scarfo's case, his inability to protect the pass phrase created an avenue for compromising the encryption. He literally typed out the answer for the FBI.

Finally, Scarfo also demonstrated a real ignorance of general security principles. Like many business people he adopted a one-dimensional view of protecting information. He felt that all that was necessary was just encrypting the data. He adopted a comfort zone in his office without evaluating whether the area was secure. His failure to establish basic physical security measures allowed the FBI to penetrate the area and to install the keyboard recorder. He failed to employ electronic surveillance to monitor the area for physical intrusion, and no emanations (radio frequencies) surveillance was in place to detect latent transmissions from the computer. Hidden traps, like applying clear nail polish to seams on the back of the keyboard or the computer, were not in place to detect tampering with the keyboard or the computer's main unit.

The FBI suffered from technological myopia. They viewed the encrypted file as strictly a technological challenge. When they could not succeed in a brute force attack on the ciphertext (encrypted file), they chose another technology. They employed a highly secret, classified keystroke-capturing device. Their decision worked. Within two weeks of planting the device, the FBI had its answer.

In a subsequent court challenge, Scarfo's attorneys threatened the classified tool with exposure. Scarfo alleged that the failure of the FBI to obtain a wiretap order before installing the device violated his Fourth Amendment rights against unreasonable searches. His lawyers argued that their client needed to know the technical details of the device to properly craft his challenge. The judge eventually ruled in the government's favor. Yet, the government ran a very real risk of losing a powerful investigative tool.

Power tends to fill the available vacuum. If the government has the technical capability, it will use that power frequently even if it may not be appropriate for a particular case. Indiscriminate use will occur. Wiretaps of telephone conversations are by law "selective snooping." If a monitored conversation does not deal with the substance of the investigation, the monitoring must cease. However, keyboard strokes constitute a stream of data that investigators must examine character by character. Since personal information is leverage, the government gets a greater yield of collateral data from this kind of surveillance. As the article in *Mother Jones* observes, a tool ideal for monitoring terrorists ends up prying into the doings of a bookmaker. Using a cannon to annihilate a mosquito is unacceptable overkill no matter how effective.

The irony that smacks one in the face stems from the content of the pass phrase. It was Nicky Scarfo's father's federal prison identification number. Most texts on computer security warn about users employing passwords and phrases based upon information from their daily lives. Easy to remember facts like birth dates, street addresses, license plates, and so on all serve as easily guessed passwords. The FBI failed to investigate this side of the case. The Bureau's investigators could have checked into likely associations the suspect might use in a pass phrase. Instead they pursued a strictly technological solution.

Scarfo and the FBI shared a common perspective. They both embraced technology as a solution to the secrecy issue. Undoubtedly, the criminal sector will continue to employ computer technology as a record-keeping tool and even as a means for committing crimes. The lure of convenience and portability insure its continued use as a weapon. The increasing high profile by computers as a vector for crime will be a boon for computer forensics experts. Expect a continually escalating war where one side makes technological progress and the other responds with a stronger hand. However, the public may find itself the loser in such a war. We may evolve into a society where secrets in digital form will be very difficult to maintain and where privacy becomes only a faint memory.

A *Washington Post* article in the August 6, 2003 *Austin American-Statesman* echoes the reality of this concern. Florida has a new counter-terrorism database called the "Matrix" (no relation to the film) which stands for "Multi-State Anti-Terrorism Information Exchange." Possessed of high granularity (locating by fine details) in search capability, the database can locate a "brown-haired owner of a red Ford pickup in a 20-mile radius of a suspicious event" to quote the article, and the Homeland Security Department appears to be jumping on the global database bandwagon with a "pilot data-sharing network in Virginia, Maryland, Pennsylvania, and New York." John Poindexter of Iran-Contra fame continues to advocate for the Defense Department an international database to achieve the same results overseas.

What will be the private sector's counterpoint to the new onslaught by government to vacuum up every available bit of data about people? Are we heading for George Orwell's dark vision in *1984* or something worse?

As a counterpoint to growing power in the public sector, I foresee a new security professional entering the arena. Infamous persons like

Scarfo and the legitimate individual alike may retain the services of a private digital security consultant. Such a professional would not be reliant upon the traditional governmental/corporate structures for a livelihood. They would design security programs for private parties. The harbingers of this trend are already evident in the literature becoming available on outlets like Amazon.com. Titles such as *Invasion of Privacy: How to Protect Yourself in the Digital Age*, *Y our Secrets Are My Business*, *Desktop Witness*, and *Secrets of Computer Espionage* offer step-by-step instructions on evading prying digital eyes whether they are from the private or public sector. *Desktop Witness*, is a manual, albeit at times a bit technical for the average reader, for hiding information from a repressive government. *Secrets of Computer Espionage* supplies detailed instructions on how to do surreptitious computer forensics.

Perhaps, individuals employing private investigators and security consultants will act as a bulwark against overreaching government. Technology that unveils can also hide. Our constitutional framers never foresaw the power of information. Their vision was one of knowledge. Knowledge, which is information with understanding and wisdom added, can produce an enlightened society. The misuse of raw information, however, often leads to autocracy and tyranny. Total access to a person's information carries the threat of control, the power to intimidate and to coerce. We could slip into a shadow government where the public institutions become irrelevant, and the true power lies with behind-the-scenes investigative agencies.

A passive information underground would use unconventional methods for storing data. The "book-people" in Ray Bradbury's *Fahrenheit 451* serve as a metaphor. They internalized books by memorizing, so the government could not seize and burn the contraband texts. In our world this action would translate as rendering sensitive data into non-digital form or into unconventional media. Bookmakers prior to Nicky Scarfo's generation kept records on flash paper. If law enforcement raided an operation, records could be destroyed in an instant with a match. Nicky Scarfo could do the following in the future:

1. As Michael A. Caloyannides states in *Desktop Witness* (John Wiley & Sons, Inc. 2002), "most substantive criminals intentionally do not use encryption because it is too alerting." Instead, they use book codes, messages hidden within normal correspondence, and "double-talk" in ordinary language.

2. Place more voluminous data on paper or on removable digital media and hide it at a "spy-drop," which is a physically secure location that cannot be tied to the data's owner.
3. Use objects in the physical world to store information. The location and sequence of chalk marks of light poles or exterior walls can hide significant amounts of data. At the very least, these markings serve as a clandestine medium for leaving messages. ("War chalking" of WiFi access points (APs), to signal wireless eavesdroppers of listening spots, is an example of this method.)
4. Hiding information on other people's computers or servers. Plenty of File Transfer Protocol (FTP) servers allow guest logins and make excellent refuges for rogue data. (FTP servers allow businesses, schools, and universities to store and to transfer large data files. Think of them as digital libraries.) Even if someone discovers the rogue data, it lacks context and ownership labels to prove who the owner is.
5. Coding information using games like chess, bridge, poker, and puzzles. Any game with sufficient complexity can act as a communication or storage medium. Less infamous persons could employ similar techniques to protect sensitive information against digital detectives.

UNCONVENTIONAL THINKING

As indicated prior, analyzing threats is key to the successful management of sensitive information. To defend everything is hopeless. Where are you vulnerable? Answering that question involves risk assessment. A risk assessment creates a sound approach to safeguarding secrets.

If you fall into any of these categories below, you have a high level of vulnerability:

1. Involved in a divorce or family law dispute such as child custody.
2. Engaged in an unconventional lifestyle. For example, you worship an Earth religion such as Wicca in a religiously conservative community.
3. Politically active and in the limelight.
4. Ongoing legal problems, whether civil or criminal.
5. Involved in a dispute with a major corporation or a government agency. Someone powerful is upset with you.

6. You and your business are a potential target for industrial espionage.
7. You are the target of investigative journalists.
8. Private investigators are conducting an inquiry about you.

A divorce or a family law case sparks intense emotions and concerns about what the other side is hiding. Computer specialists offer services to family law attorneys to break into the opposing side's PC. Placing sensitive information on a computer in this situation runs a real risk of exposure.

Unconventional lifestyles can draw the interest of an opposition bent on driving you out of the community. Your opponents going through the trash or gaining access to your computer is quite possible. They will be looking for "dirt" to embarrass or to ruin you.

The private lives of the prominent in the political area are fair game in today's world. If your private life is less than impeccable, be very careful of what falls on your computer's hard drive.

Those chronically in trouble with the law or engaged in litigation do not need their computer as a silent witness. Avoid the digital world. Seek other ways to store information.

If someone powerful is angry with you, be circumspect in how you store information. They can afford to place private investigators on your information trail. In addition they may set law enforcement or journalists to look into your personal and professional life.

If your business is one where intellectual property drives the business, then always be on guard for industrial espionage. Whether you are at home or at work, your computer needs to be physically secure, encrypted, and locked down with good password protection. Computer media needed to be wiped and destroyed prior to disposal.

An important question becomes: why are you storing a given piece of information on your PC? If the benefits do not outweigh the risks, rethinking a digital format may be worth considering. Even if you create files, and write them directly to a removable disk such as a floppy or a USB drive, forensic examination may detect the transfer. USB stands for Universal Serial Bus. It is among other new industry standards that allow the connection to a computer of a wide range of memory devices including removable microprocessor drives and cards.

An examiner will see that you created a file and placed it on a removable disk. That bit of information may cause problems for you or it may not. Just keep that potential in mind.

Small trails left in the digital record provide stepping stones from nonsensitive data to sensitive information. Think about and visualize an electronic trail being forged every time something happens on your PC or other digital device. Web sites visited, e-mails sent and received, document created and transferred, and other transactions all leave clues.

Possible solutions to this problem are as follows:

1. Use different computers for different levels of sensitive information. In the computer security field, there is the concept of a sacrificial host. Such a machine is one you can live with being compromised. If someone breaks into it, they are not going to learn anything of real value. You would keep this PC in say your home and use it only for mundane purposes. For more sensitive information, you would use another computer in a location not known to others.
2. As indicated, removable hard drives, and USB drives are an option. Just be able to explain any transfer trails. Watch out for unintentional "temporary" copies being made on your PC by the application.
3. Use encryption, but remember it attracts attention. For highly sensitive information, encryption should not be the only means of defense. Think in terms of layers of protection. Good physical security in a location not known by others greatly enhances the encrypted file's secrecy.
4. Keep certain facts in your head. Obviously, you cannot memorize an entire database, but you can memorize the password for it.
5. Forget about hiding tricks such as files within files or within images on your PC. Computer forensics people know about these tricks. Searching for them would be a normal part of any forensic examination. You are far better off hiding data physically elsewhere.
6. Use a false fortress PC. Remember that creating the fortress involves several steps. Erect a portcullis with a screensaver password. Dig a moat by adding a long, random, very hard to guess, wildly chaotic password for the administrator account. Build a formidable wall by changing the name of the administrator account to something else and run the machine as a highly restricted user. But even if an attacker gets through these defenses, not having sensitive information on the machine provides the ultimate defense.

7. Finally, do not carry your whole life around with you. Men have everything on their laptops and in their briefcases: from business plans to medications to emails to their sweethearts. Women shoulder large, fashionable bags to tote around PDA's, laptops, appointment books, correspondence, and even diaries. True, these portages offer convenience, but they also serve as haute cuisine for information thieves.

OTHER STEPS IN PROTECTION

Overwriting data is another way to foil computer forensics. As indicated in Chapter 5, overwriting takes time and numerous passes to be effective, but it does remain an option in anti-forensics.

Bram Shirani, CISSP comments in his 2002 presentation on "Antiforensics" that Alternate Data Streams (ADS) is another method of concealing sensitive files. The technique works on Windows systems and allows the linking of a secret file with a normal one. (See http://www.diamondcs.com.au/streams/streams.htm.) Computer forensics examiners now have programs to detect ADS, but it still remains an effective technique against a less than rigorous examination of the PC.

Shirani also suggests that if one wishes to keep "hacking," steganographic, or forensics tools on their machine, storing them as an encrypted tar archive makes them highly resistant to computer forensics attack. He states: "tar up all the tools they have used, and encrypt them. This may be noticeable, but if they anticipate being noticed, they don't care. You will probably never find what they have hidden in their crypto-tar."

Using forensic tools to detect slack space and to examine swap files is another step in the cleaning of one's "digital tracks." The user can then overwrite the discovered sensitive areas.

A general clean-up of a computer would include:

1. Remove hidden files (including ADS) and directories.
2. Eliminate unwanted metadata from documents. (See Chapter 1.)
3. Remove sensitive e-mails.
4. Remove steganography tools and host files.
5. Remove hacking and forensic tools.
6. In the alternative, do encrypted archives of tools.
7. Erase and purge "deleted files" containing sensitive information.

Chapter 9

EVALUATING WEB PAGES

The previous eight chapters have a common theme: protecting the confidentiality and integrity of physical and electronic documents against internal and external attacks. Guarding assets, however, differs from relying on information provided by others. When users depend upon information external to the organization, risks do exist, especially when the information tendered serves as the foundation for decisions. External information affects decisions concerning:

- Employment
- Purchasing
- Marketing and Distribution
- Mergers and Acquisitions
- Contracts and Joint Ventures
- Corporate Strategy

External documentation has two forms: Web-based information and physical or electronic documents submitted by outsiders. This chapter deals with evaluating Web-based information for truth and reliability, which are two values that create trustworthiness. In the next chapter, the discussion turns to spotting document forgeries or fraudulently altered documents, whether they be in electronic or physical form. Bad information or documents coming into the organization's decision making process causes as much harm as sensitive information falling into the wrong hands. A balanced defense in both directions (data coming in and going out) is essential for a complete document security program.

The Web is far from a benign environment. Yet, many presentations to decision makers rely upon research gathered from the Internet. An information war of international proportions rages in cyberspace. Deception seeps into mundane information. Spammers constantly repackage their e-mails to circumvent anti-spamming software. Hucksterism and outright misleading statements dominate certain categories of online advertising. Legitimate sites face counterfeiting, which ends up enticing the unwary into using the bogus site with the users thinking they are contacting the legitimate one. Unbridled by the editing and review procedures of traditional media, some sites traffic in disinformation or irresponsible gossip through the ever-growing numbers of blogs, Weblogs, personal sites, and unregulated Wikis. Combating fraud and disinformation becomes a part of the information security professional's lot.

PERSUASION

Aristotle's *Rhetoric* argues that logic, character, and emotion all factor into persuasion. Those ancient perceptions on the ability of information to persuade hold true in the twenty-first century. The logical dimension of Web-based information is the information's accuracy. Do the Web pages make arguments based upon facts? Are those facts correct, based upon confirmation by other sources? Are the arguments sound logically? Do the premises really support the conclusions?

Character issues arise from reputation. What are the motivations of the author or authors of the site? How reliable is the information that the site offers? What other sites link to it? What are the opinions about the site based upon others that use it? What bias does the site have?

Emotion is a powerful force. The emotions stir us to action. They can serve, however, either good or bad ideas. What kinds of emotional appeals does the site make? Are there appeals to fear or other negative emotions? Does the site seek a balanced approach to the issues at hand, or is it merely engaged in hucksterism? Always be aware of the level of emotional manipulation that a Web site engages in with the visitor. If the user's intent is to simply rally to a cause by visiting the site, the emotional appeal may be the way to go, but in gathering information that is reasonably objective, strictly emotional appeals should raise questions in the user's mind.

Madsen Pirie in *How To Win Every Argument: The Use and Abuse of Logic* (Continuum, 2006) observes that multiple forms of emotional appeal exist in everyday discourse. They include:

- Appeal to fear.
- To envy.
- To hatred.
- To superstition.
- To pride.
- Emotion is better than reason

Emotion is what starts us to do something, which is good. However, upon motivating us, emotion needs to take a backseat to reason. Reason helps us make the right decisions as to acting on our feelings. Emotion should not hamper our judgment or be the sole factor in making decisions. Is the Web site being fair with the user, or does it seek to manipulate rather than to inform?

While these questions may seem straightforward, do they get asked enough in daily commerce on the Web? Web-based information boils down to trust. How trustworthy is a given site? Criteria for evaluating Web sites are essential when researching for information on the Internet. Trustworthiness depends upon these criteria:

1. Source Bias.
2. Reliability, the history of accuracy.
3. Validity, correspondence to reality.
4. Authority, the level of credibility.
5. Ability to verify claims presented on the site.

Is computer-based information more trustworthy than printed sources, or is it less trustworthy than print sources? The answers to these questions depend on the criteria listed above. In some cases, computer-based information will be a more trustworthy source depending upon how it ranks on the five criteria.

Probably most important, does the site reveal or indicate its biases? Are the sources for the site's explained? Are contributors identified? Are the viewpoints of particular causes, organizations, and groups clearly delineated? Is fact separated from opinion on the site? If the site has a commercial interest in the information that it presents, is that interest divulged?

Finding sources that are consistently reliable is not always easy. News sources, for example, do contain errors. Reliability is more, however, than just etching "truth" in stone. Information fluctuates in value, relevance, and importance. A source's determination to update information and to make corrections is just as important as getting a particular fact correct. The willingness to revise information and to investigate further is a core factor in evaluating a source's reliability.

In judging a Web site's reliability, a user must determine if the site has mechanisms for feedback from users. Does the site acknowledge factual errors and does it explain how the various hosts respond to inquiries regarding errors? Is feedback available through an e-mail address provided, or is there a feedback portal on the site? Are errors posted on the site along corrections and comments by users? These questions are all important in deciding on a site's reliability.

What mechanisms does the site have for determining whether its information is true? How does it gather information? Does it have a staff of fact gatherers, investigators, or journalists, or does it merely report what others have said, without corroboration? Are claims presented on the site investigated? Are links provided on the site to gain access to other sources of information? Does the site have a track record on its information having a correspondence to reality?

As far as the authority of the site goes, what other sites have links to it? Do news sources cite the Web site as an authoritative source? Do the sponsors of the site have creditability? Are they recognized as authorities in their field? What level of documentation does the site offer for its information? For example, does it provide supporting evidence, white papers, links to other sites, testimonials from customers, or reviews or comments from independent parties to substantiate its claims?

Is the site specific in its claims? Does it state supporting facts for its positions that can be independently confirmed by investigation elsewhere? Does the site offer the means to contact its staff by providing a physical address, contact telephone numbers, and e-mail addresses? In contacting the staff, do they respond to inquiries fairly and in a reasonable time frame?

A site may be quite persuasive in its content, in its design and graphics, and in the enticements that it offers visitors, but the fundamental question becomes:is the site fair with its users? Does it acknowledge where it is coming from and what its mission is? Are its promises lived up to by its management and staff? Asking questions of a Web site takes time and energy, and in the casual user contact with the site, probably intensive questioning is not necessary. When Web-based information serves as the foundation for major decisions, then careful questioning and evaluation of the site and its information is an essential survival skill.

The Web is a global theater that provides entertainment and human interaction. Understanding how that theater appeals to our emotions should trigger our thinking capabilities to investigate situations where we sense we are being manipulated. Our bedrock in conducting investigations should be other persons that we trust for their advice and opinions, information sources that bring a different perspective to the issues at hand, and our own good sense and judgment, when we take the time to step back and to think.

Observing maxims such as "Investigate before you invest" and "Extraordinary claims require extraordinary evidence" furnish pause for reflection. Developing a checklist for evaluating a Web-based source is another effective line of defense. Here is an example of a possible checklist:

1. Age of the site?
2. Date of the document on the site?
3. Dates last updated for the site and the document?
4. List of contributors provided?
5. Checked the registry of the site's IP address to determine ownership? (Open an MSDOS window and run the inquiry "nslookup <site's name>." The inquiry will provide the IP address. Go to http://www.arin.net/whois to enter the IP address to obtain the registry information.)
6. Contact information (physical address, telephone number, and e-mail address) provided on the site? Do those contact points actually work? (Do directory assistance or the online yellow pages confirm the telephone number and physical address? Does the site's staff respond to telephone calls and e-mails?)
7. Check for links on the site.
8. Are the links active and unbroken?
9. Are links to and from the site informative? Reputable? Authoritative?
10. What is the depth of content on the site?
11. Does the site provide reader or user helps such as graphics, tables, supporting documents, fact sheets, text boxes, a glossary, a bibliography, and an index?
12. In what ways does the site inform the user well?
13. In what areas is the site misinformed or lacking in information?
14. What is the experience and background of the contributors? Do Web engine searches produce any background information on the contributors?

15. How is editorial oversight done for the site?
16. What accreditation has the site received?
17. What is the site's privacy policy?
18. In what ways does the site benefit a user?
19. In what ways should a user be concerned about the site?
20. What steps should a user take to investigate the site further? What independent sources can confirm the legitimacy of the site's information?

If users get into the habit of cross-examining Web sites and online documents, they will develop skills to resist unethical persuasion in cyberspace. The art of cross-examination in cyberspace requires understanding the different categories of online predators. In the following paragraphs, the text describes the personas of the common predators one is likely to encounter.

Panderers appeal to the appetites and to greed. Promises of sex, romance, or easy money are their stock-in-trade. Usually they deliver on none of these promises. A common scam is to send an e-mail from a female to an unsuspecting male target. The e-mail claims that the sender thinks she has met the male target before online. To contact her, the recipient of the e-mail gets a Web site URL to click on. Upon arriving at the site, the target has to enter a credit card number for "identification purposes only." The site states no charge will be against the credit card; of course, the truth is just the opposite. Charges will appear on the card. It is a stupid thing for the target to do, giving the credit card information, but the promise of romance makes people do stupid things. Schemes to work-at-home for thousands of dollars a week and for getting rich quick buying and flipping real estate with no money down are in the same vein. They prey upon a fantasy. Their attacks affect individuals and employees with a weakness, which may be exploited to the detriment of the individual and the organization.

Charlatans are con-artists. Rather than just appealing to the most basic of human instincts, their "pitches" have superficial logic. The "Nigerian scams" are a prime example. The target receives a letter or an e-mail claiming that sender lives overseas and needs assistance to obtaining an inheritance in the United States. For a small investment, the recipient can aid in recovering the funds and receive a portion of the estate. Many variations exist on the scam, but the theme is helping some poor foreigner out. Charlatans come in many varieties. They range from "boiler room marketers" sending out spam to lure targets to their Web sites to purchase inferior or bogus products to those that scam merchandise, services, and payments off of auction sites like eBay®. The common trait of this persona is the lack of contact information. They take great pains so you cannot get a hold of them. Businesses and individuals must shun doing any commerce with companies or individuals that do not have alternate means of contact outside of the Internet.

Information thieves want an individual's marketable personal data or a business' proprietary information. They haunt chat rooms, newsgroups, and collaborative forums on the Web "bottom feeding" for information. In addition, they have Web sites that promise certain goods and services or prizes if the individual completes a questionnaire. They send out impressive e-mails to targeted individuals within an organization or business that appear to be from a professional or trade organization offering a "free" subscription. Just go to their Web site and fill out the "profile." Do not give out information about yourself or your organization to strangers, unless you first verify who they really are. Use the Web site checklist given above to investigate them, provided what they are offering is worth your time to do so. Many times just ignore them, for you will not be losing anything.

Spammers bombard businesses and individuals with inferior products, inflated claims, bogus remedies, pornographic ads, and just plain nonsense. Unfortunately, these merchants of foolishness thrive because some people respond to their spammed information and ads. Never respond to spam. Report spam to your internal anti-spam team or to the appropriate anti-spam site. Never open attachments to a spammed e-mail. Treat spam like the garbage that it is, and put it in the trash where it belongs. Handle all e-mail where you do not know the sender with the greatest skepticism. If you cannot verify who they are, trash that e-mail.

"Griefers" are online bullies. Found usually in the gaming forums on the Web, they engage in underhand tactics and the taunting or baiting of other gamers. These individuals thrive on belittling others, and this persona is not just limited to online games. Professional articles on computing published to the Web attract "know-it-alls" who post diatribes anonymously against the author or who send the author flaming e-mails denouncing the article. Ignore these "jerks." They are too cowardly to reveal who they are and to take responsibility for their actions and words. If you respond to their baiting, you may reveal information about yourself or your organization that you'll regret later, and you could set yourself up for online stalking.

Commercial hustlers turn everything into an advertisement. Their Web sites exist solely to sell you something; whether that product or service is something that one actually needs is another story. If a user does not encounter the Web site directly, usually a pop-up ad tied to another site will direct one to them. The high-octane commercial sites may contain inflated claims and artificial time deadlines for purchase. The rule of thumb is if you sense any "hustle" going on, move on.

Extremists use the Web for their own nefarious ends. These activities include recruiting, fund raising, communicating among members, and gathering intelligence against targets. Extremists range from domestic terrorists to international terrorists to fringe political groups advocating violence and confrontational politics. Some of these groups will have an open Web presence, which can provide some useful intelligence about their activities. Other groups may operate on the Web through front organizations, a cover that aids in "false flag" recruitment and fund raising. Unsuspecting individuals may end up contributing to what they think is a benign cause without realizing who actually receives the funds. Hate sites are usually fairly obvious. But individuals and businesses need to check out any "charitable" organizations on the Web with the Better Business Bureau and the GuideStar.org Web sites prior to contributing. (See http://www.bbb.org and http://www.guidestar.org.)

Security professionals also need to be constantly aware that extremists gather intelligence via the Web, so keeping a client's information footprint on the Web as small as possible is a good practice. If the client may be the target of extremists, it is unwise, for example, if the CEO's travel itinerary appears on the Web. (See the article, "www.terror.net: How Modern Terrorism Uses the Internet" at http://www.usip.org/pubs/specialreports/sr116.pdf.)

Identity thieves are a special category of information thieves. The tactics they employ are quite similar. Usually, they are the masters of URL obfuscation and in copying the design and look of legitimate sites. If they send a link in an e-mail purporting to be from say the target's bank, the URL will resemble, at least in part, the banks' URL, but the active portion of the address will be in digital or Unicode format to obscure the actual destination. The destination Web has the appearance of the legitimate one, but it is only a portal to gather personal and financial data on the unlucky users who end up there. Most sophisticated attacks include

DNS cache poisoning and cross-site scripting. In these attacks the thief redirects the user's IP traffic either to a site of his or her choosing or through the thief's own server in order to intercept communications between the user and their financial institution.

"Hatemongers" are a special case of the extremist category. They may not attempt direct violence. Instead, they foment racial, religious, sexual, or ethnic bigotry. While these purveyors of prejudice and hatred do not pose a direct risk to most business operations, employees should not be allowed to visit these sites from the company's computers and networks. The company's acceptable use policy for accessing the Internet should prohibit visiting hate sites, due to the internal strife they can create in the work environment and their basic unfairness to certain classes of employees.

Online criminals round out our discussion. The individuals are predators using the Web to facilitate their criminals. The most infamous is, of course, the online pedophile, but other sexual offenders also use cyberspace to lure victims into situations where a sexual assault can occur. Common-law criminals such as robbers and burglars do research and gather intelligence through the Internet. (So, the rule of thumb is not to reveal sensitive information about one's self to strangers online. Once you identify where you live and when you will be out of town, a burglar has all he needs.) Other criminals include gangs and drug traffickers who use the Internet for communications and to gather intelligence. The common theme or observation regarding these criminals is that security professionals should be vigilant to make sure that their organization's networks are not exploited to commit these crimes. Insiders like employees, consultants, suppliers, and vendors, who have access rights, may try to use the organization's information resources. Deploying adequate content surveillance and monitoring on your organization's Internet connections can help detect this criminal activity.

DISINFORMATION

Disinformation seeks to persuade others with falsehoods that fundamental problems exist with a targeted organization, enough so to raise serious questions of trust with the public. In campaigns waged against a business or organization, several platforms serve this purpose. The rumors may start in a blog, in a Wiki, in a chat room or in a discussion forum, or on a Web site. Since the Web is largely unedited in the sense of traditional media, virtually any thought or opinion makes its way into cyberspace. The facts may be spurious, but the contagion spreads if it catches the attention and imagination of the Web community.

Recently, numerous Web sites sprouted in connection with the 9-11 tragedies. They alleged U.S. government complicity in the attacks and the subsequent "cover-up." These conspiratorial sites and discussion groups constantly begged the question by examining photographs and video from their point of view only. Arguing, for example, that the hole on the side of the pentagon was not large enough for the commercial airliner crashing into the structure, they offered little more than opinion, or internal explosive devices brought down the World Trade Center towers. Again, the theorists did not venture beyond conjecture with their "evidence." Later investigations by both governmental agencies and private organizations discredited these claims of an internal American conspiracy. Yet, the ability to capture public attention via the Web, even when the claims border on the fantastical, offers a powerful lesson on what messages can disseminate through the Web.

Disinformation attacks on an organization concentrate in three major areas: questioning the firm's reputation, disputing product or service reliability, and interfering in the firm's business relationships. Allegations of law breaking or questionable business ethics rob a company of its credibility in the marketplace. When these allegations are not substantiated in the traditional media or by regulatory agencies or by law enforcement, they become highly questionable.

Readers may recall a number of years ago, certain microprocessors produced by Intel® had a flaw in their arithmetic logic units (ALU) that resulted in a large number of calculations being incorrect. Intel acknowledged the problem and recalled machines with the flawed microprocessor. The fact was most users would have never noticed the problem in doing ordinary tasks like word processing. Yet, even after the machines and chips were recalled and fixed, potential customers would come into retail stores demanding to do testing on the PCs to see if they could detect a calculation error. Allegations of product or service problems linger in the marketplace long after the events that precipitated them. In Intel's case, the problem was real, but overblown, and it posed no safety risk to the consumer. Allegations of safety risks, such as contaminated food or impurities in drug products, can cause immediate consumer reaction and avoidance of the product.

Interference in business relationships involves putting pressure on suppliers, vendors, and customers not to do business with the targeted organization or company. These allegations run the gamut from the company not paying its bills to being involved in behavior unacceptable in the marketplace such as racial or ethnic prejudice or unfair labor practices. Again, a ripple can grow very quickly in cyberspace into a tidal wave. Spreading lies about an organization can have a rapid, crippling effect on that company.

Social engineering is a form of disinformation. Its schemes include inducing fear, appealing to authority, appealing for help or pity, and declaring a crisis when there is not one. Whether in response to a telephone call or to an e-mail, the employee is afraid of getting into trouble if he or she does not provide the requested information. Much in the same way customers fear doing business with a company under a "dark cloud," fear becomes a powerful tool in the disinformation toolbox. Sounding important or powerful hopefully makes employees kowtow, so the social engineer gets what he or she wants. Citing authority or sources that appear authoritative is another tactic of disinformation. Merely because a Web site is well-footnoted does not mean what it has to say is true. Appeals for help or to pity are another clever tactic for most people have a natural inclination to want to help others. The social engineer manipulates an employee into thinking the person asking for information is helpless or in peril. Purveyors of disinformation against an organization create the impression that victims are suffering at its hands, which creates pity in the general public.

Finally, manufacturing a crisis or an emergency provokes an employee to act quickly in supplying information to the social engineer. Unless a password becomes immediately available, the marketing campaign will be lost or a key customer will close its account. The variations on the theme of crisis are myriad. Traffickers in disinformation against an enterprise stress that the target is creating a crisis or an emergency situation, and unless swift counteraction occurs, the consequences will be dire.

Handling disinformation cases requires a coordinated effort by the affected operating units, the security team, the legal team, and the public relations staff. Rumor control should be the first step. Getting correct information into the conventional media and on the Web can dampen the effects of disinformation. Restoring public confidence and that of customers, vendors, and suppliers is essential to preventing an avalanche of hysteria.

In investigating disinformation cases, determining the primary source or channel for the false information will provide a correct focus for the inquiry. Locating the point of origin will mean following site links and threads in online discussion groups. It may also mean determining the blog referenced by other Web sites and news sources. Finding the epicenter for a disinformation campaign may boil down to creating a list of those individuals and organizations or businesses that have a motive for waging information warfare against your client. The usual suspects include disgruntled former employees, overzealous watchdog groups, extremists with a grudge against your client, unethical competitors, and customers, vendors, or suppliers who feel "wronged" by your client. Checking the Web presence of the leading suspects through online search engines should uncover their participation in blogs, Wikis, collaborative spaces, Webinars, and Web sites. Always determine the Web "footprint" of the disinformation. Who is saying what information about your client, and where are they saying it online?

Contacting sources at the secondary or tertiary level may provide valuable leads. These sources may be able to inform an investigator from which site they gained the information that the client was having a "problem." Interviewing these sources via telephone or sending an e-mail helps in the gathering of information about how the disinformation has rippled across the Web. In addition, following links from these secondary or tertiary sites should lead an investigator to the primary site of origin. Uncovering the places on the Web where the disinformation originated is essential.

Once investigators identify the responsible parties, appropriate legal counsel should be able to determine an effective course of action. The actual site hosting the information may not be aware of the disinformation. Someone could post false information on a Wiki or in a collaborative space with little or no editorial supervision. Often, placing the site on notice regarding the disinformation posting will result in prompt action to remove it. If a site's administrative staff removes a posting, then the client's media or public relations department needs to issue a press release regarding this corrective move. Additional legal action against the parties responsible will depend upon the advice of legal counsel. The factors considered for additional action would be the amount of damages caused by the disinformation, the laws or statutes violated, and the ability to construct an electronic or paper trail establishing the responsibility of those involved in the dissemination.

FRAUD

Scams and swindles continue to be a serious problem in cyberspace. Obviously, goods and services paid for but not delivered is common when one is dealing with strangers on the Internet. Unregulated auction sites and commercial sites without appropriate safeguards are the arenas for these crimes. The solution lies in using sites that protect both sellers and purchasers. They have reasonably secure payment processing systems to protect sellers and channels to handle complaints by purchasers. Dealing directly with a seller or vendor without a trusted intermediary requires caution. Check with other customers that have used the site. Run search engine queries using the vendor's name and site information to locate complaints about them posted on the Web. Make sure the site has contact information in addition to an online e-mail address. Can you call them on the telephone? Do they have a physical address?

Information thieves that run sites merely to gather sensitive information from the unsuspecting, we have already discussed. Volunteering information about oneself or one's business is unwise when the party requesting the information is a stranger. Find out who you are dealing with before you provide any sensitive information.

Bogus charities operate on the Internet, so check them out prior to contributing. Even more serious are the impersonators of legitimate businesses. They may mimic a financial institution in an e-mail or on a bogus Web site. Conducting bogus online surveys to obtain personal identifiers or sensitive business information is another ploy. Reverse social engineering tactics may have impostors representing themselves as law enforcement or as corporate security to perpetrate identity or information theft. Their alarmist e-mails often state that your account has been compromised and that you should respond immediately with the needed information.

Hoaxes waste time and resources. Have a policy in place to screen out hoax e-mails, and be sure to train employees in how to evaluate hoaxes. A good site to aid in this effort is Hoaxbusters at http://hoaxbusters.ciac.org/.

The best protection against online frauds is to use the checklist provided in the "Persuasion" section of this chapter. (See also "Reviewing and Verifying Documents" in Chapter 10.) Most important, always seek confirmation from other sources on the claims made by Web sites. Search the Web using online search engines to find comments about the

Table 9.1: Evaluating Web Pages

Criteria	Notes	Importance
Links	Other sites linking	The regard other sites have for the site you are evaluating.
Facts/Content	Accuracy, validity, and reliability	Cross-check facts against other sources to determine their accuracy.
Contact Information	Can a user get a hold of them by other than e-mail?	Alternate contact points must be available.
Testimonials	People who have used the site or purchased from the site	How do others evaluate the site?
Site Ownership	Registered owner and physical location of the server	Is the actual ownership and location of the site the same as what is represented publicly?
Metadata	Are there any inconsistencies?	Look for the misrepresentation of facts such as authorship and document dates.
Emotional Appeals	What does that tell you about the site?	Be on the lookout for manipulation.

site that you are considering doing business with online. (See Table 9-1: Evaluating Web Pages.)

SUMMARY

The greatest dangers from relying upon Web pages are their potential to persuade us unfairly, to disseminate disinformation, and to deceive us into yielding either our money or our sensitive information. The critical evaluation of online information will be a vital skill in the twenty-first century. Integrating the examining of documents, the analyzing of content, and the cross-checking of facts presented in documents and on Web sites will be the hallmark of the information security professional.

Chapter 10

DOCUMENT FORGERY

Forgery detection has been the province of the security professional in the banking and financial sectors for over one hundred years. The game of forging negotiable instruments remains with us, aided by the revolution in desktop publishing technology. Check forgery continues to be a major loss category to businesses and financial institutions.

Negotiable instruments, however, are not the only vector for forgery or document alteration facing the twenty-first century enterprise. Bogus invoices and billing documents, fake identification, fraudulent academic records and degrees, and fraudulent business records also pose serious risks to organizations if relied upon to make decisions. Depending solely upon documents, without independent verification of their authenticity and reliability, carries dangers to staff conducting due diligence investigations of prospective employees, customers, vendors, suppliers, and potential business acquisitions.

Today's charlatans know about the dangers of ineptitude in their craft. Savvy enough to create convincing documents, professional forgers avoid the common mistakes that are the hallmarks of amateurs:

1. Misspellings of common words.
2. Obvious errors, like giving the wrong name of the governmental agency issuing the document.
3. Not using the correct formatting scheme for the document, such as the wrong color background for a driver's license.
4. In a paper document using the wrong paper stock.
5. Ineptly substituting photographs, images, or logos where the alteration becomes obvious.
6. Poor lamination techniques for wallet or badge documents.

7. Forging an older document, but the forgery looks like a brand new document, or making same mistake of creating a new document, but it has elements that make it look old.
8. Inconsistent data in the document, such as dates do not correspond with other information within the document. (The photo is of a fifty-year-old man, but the date of birth is March 2, 1985.)
9. Not including the necessary logos, images, holograms, or magnetic striping expected for the document.
10. Border detail work or other detail features are not of sufficient quality to judge the document as genuine on first inspection.

Amateurs often do not check for inconsistencies after transferring information from one document to another. Today's desktop publishing technology, however, can make up a number of the deficiencies listed above, provided the forger exercises careful proofreading of the forged document prior to passing it. Needless to say, checking any document for the above-listed danger signs is the first line of defense in external document security.

IDENTITY DOCUMENT COUNTERFEITING

Producing identification to gain employment has become a standard procedure in the United States of America. Businesses generally have to complete the I-9 form to verify the applicant's right to work in the United States. Applicants lacking the necessary documentation often seek illegal documentation either to hide their nationality or to protect their real identities. While the average person may lack the expertise or resources to fabricate the documents, he or she often can purchase them through professional forgers.

Forged identity documents also facilitate identity theft. The criminal obtains the basic identifiers on an individual: full name, date of birth, Social Security number (SSN), and address. Then, the thief forges identity documents, which are used in purchases, applying for loans, and in other financial transactions. The process is not all that difficult, nor does it require exceptional investigative skills. The next time you stand in the checkout line at the store, see what you can observe with a little "shoulder surfing." The person in front of you has their driver's license displayed, which shows their name, address, date of birth, and driver's license number. Someone writes a check, which reveals his or her name, address, telephone number, bank and bank account number, and even driver's license number. Some checks even have date of birth and the person's SSN on them. Information thieves shoulder surf for this information all the time, and cellular telephones with cameras in them make taking a picture of the data fairly easy. The thief just pretends he is using his cell phone for a call.

Tools that enable the creation of forged identity documents include: scanners, card printers, image editing software, and desktop publishing programs. And, the passing of the document does not always have to be done with a physical form of identification. An applicant can "forget to bring" a needed piece of information. Many organizations permit the applicant to fax or e-mail the document to them later. In electronic format, the physicality of the document is usually lost, which works to the forger's advantage. Physical characteristics such as paper stock, document aging, wallet wear on the identification's lamination, seals, holograms, and magnetic striping are not available for inspection. Forgers use electronic technology to create and to disseminate fraudulent documents. In the Cyber Age, we tend to trust what is in electronic form, and we lose the hands-on feel for documentation. Forgers exploit that tendency.

Techniques include data swapping at the pixel level, data alteration such as changing a name or date, image swapping such as one person's photograph for another, and the complete fabrication of a document from scratch. Current desktop publishing programs and imaging suites like PhotoShop® make any of these approaches within the technical capabilities of most persons. High craftsmanship with inks, pens, and hand drawing equipment is no longer necessary. Even creating a document from scratch is not that difficult if a genuine original is available for scanning.

Often the Web provides images, examples of documents, logos, and other graphic materials to create impressive looking identification. Try this experiment: go to the Google Web page and click on the "Images" search. Search for a driver's license from the state of your choice. Search for an "American Passport." Search for a corporate logo of your choice. Enter in the name of a company along with the phrase "identification card." Search for "military identification card." Search for "security badge" along with the name of a chosen institution. The results will produce numerous examples of images that can serve as the templates to create impressive looking identification.

Common identification formats that forgers seek are:

1. Driver's licenses
2. Passports
3. Military identification
4. Corporate ID cards
5. Corporate badges
6. School and university IDs

True, these formats are starting to contain certain security elements such as magnetic striping, holograms, optically variable images, and microdots of pictures and authentication information. Unfortunately, many workers who ask for identification as a part of conducting business do not have in-depth knowledge on verifying credentials or in spotting fraudulent identification. Most identification receives a "salute" rather than close examination.

Passports are a case in point. Most lay persons see the documents as being very secure and reliable. While some of them are difficult to forge, it takes a trained, knowledgeable eye to spot forgeries. If one is trying to get through American Customs with a forged American passport, then the level of sophistication in the forgery has to be high. Using a forged American passport to gain employment, to apply for a loan, or to cash a negotiable instrument requires less sophistication in the forged product.

A story by Jeff Goodell in the *New York Times Magazine* of February 10, 2002 highlighted the problems of passport forgery. The article focused on fake Belgium passports. It noted that a large number of Belgium blank passports are available on the black market. With the use of a good laser printer and knowledge of which fonts for text in the passport, faking Belgium passports is not difficult. Links between organized crime, white collar criminals, and terrorist organizations flourish in the trafficking of stolen and forged passports.

In response to this illicit trade, Belgium investigators employ several countermeasures. First, they consult a database of about 1.4 million stolen documents to cross-check the validity of presented travel documents and other identification. They also employ ultraviolet light examination of passports to examine for telltale inks and watermarks. The article notes that forgers try to add an element of authenticity to forged passports by transferring visa and entry stamps from a legitimate passport to a fake one by the use of a raw potato cut in half as the transfer medium.

In addition, the new Belgium passports have a laser-cut pinhole image of the passport holder, a watermark of King Albert II, and an optically variable image of Belgium. These measures raise the barrier for forgers quite high, especially when trained eyes examine the passports. How well these countermeasures fare with the untrained public is difficult to say. Undoubtedly, forgers will continue to traffic in clever imitations for use in the private sector to fraudulently obtain goods and services.

The United States of America also employs a database of identifying data and brief biographical information for U. S. Customs personnel to check against when reviewing passports. The U.S. Passport, while having a small numbers of stolen blanks in circulation, is far from tamper-proof. Goodell's article states that in about five minutes time, knowledgeable hands can swap out a photo on a passport by rolling back the plastic covering.

The bulk of forged identification traffic in America lays in driver's licenses, bogus birth certificates, and faked college and university IDs. Again, some law enforcement officers are highly knowledgeable in spotting the fakes, while many workers in the commercial sector have scant knowledge of forged materials. The security departments of colleges, universities, and corporations can take steps like maintaining a database of stolen or lost identification documents. Such a database aids officials in any subsequent investigations should the identification emerge as part of a fraudulent scheme. In addition, timely reporting the loss or theft of such credentials to law enforcement adds to the intelligence resources of government agencies investigating identity theft cases and scams involving fake identification.

Obtaining manuals describing and providing pictures or photographs of common forms of identification such as driver's licenses is another aid to employees who have to review identification as a regular part of their job. The Drivers License Guide Company publishes a guide to all U.S. drivers' licenses with photographs of genuine licenses from each of the fifty states. (See their site at http://www.driverslicenseguide.com.) Retailers use the guide to spot bogus drivers' licenses presented when paying by check.

Limiting the documents acceptable verifying identity is another way to increase security. Not accepting faxes, copies, or e-mailed images of identification should be standard procedure when correctly validating the subject's identity is critical. Requiring a photographic identification is another safeguard. Identification issued by a governmental agency should be the standard. Corporate and university or college identification cards have their place, but the lack of uniformity in their design and format makes it difficult for training those who examine identification regularly. Consider having employees receive training from a representative of the governmental agency. This individual can be from the state drivers' license bureau, the military, or the U.S State Department. Also, provide employees documentation available to evaluate genuine identification documents issued by the agency. Again, publications like the guide for drivers' licenses described above can accomplish this end.

COUNTERMEASURES

Forgery-resistant identification has been evolving over several decades. Two dynamics play in its evolution. First, security professionals want identification documents that are difficult to forge. The amount of detail required becomes a bar to the document appearing genuine. And second, forgeries should differ from the genuine identification document in ways that make them easy to spot. Detection of fakes becomes a critical factor.

In order to achieve these two aims, understanding how forgers do their craft is important. Several strategies are available to the twenty-first century forger:

1. Copy a genuine document after making key data changes.
2. Swap out images or photographs and key data.
3. Treat the genuine document chemically to remove certain data and then substitute the fraudulent data.
4. Remove security features like taking a hologram from a genuine document and transferring it to the forgery.
5. Create a complete forgery from a facsimile of a genuine document.
6. Pass a genuine document of another person as one's own. (When a close resemblance exists between the passer and the document's owner, this approach is possible.)
7. Pass a forgery as a copy, an electronic image of the "original," or a faxed document. (This "second generation" approach gets around the tell-tale physical characteristics such as paper stock or lamination design.)

Modern countermeasures seek to combat all of these strategies whether the forgery is in paper or electronic from. Before enumerating those countermeasures, one must remember that no matter how effective the security technology is, the detection of forgeries depends upon human observation. A lack of motivation on the part of the reviewer may prevent good observation and let even a clumsy forgery pass. or the reviewer may succumb to fatigue, to social engineering, or to other psychological manipulation by the passer. Vigilance becomes the hallmark in examining identification.

Digital watermarks enable digital documents to have a marking that indicates their true origin. A digitally watermarked photograph or image, even if transferred to another document will still retain its digital stamp of identification as to origin and authorship. This digital identification is embedded in the document or image and, it is difficult for the human eye to detect, but it is readable by computers, so it acts as a strong countermeasure against data swapping and alteration. Also, the whole document can have a digital identifier, which can be recorded in a database. This measure also prevents the outright fabrication of false identification electronically because the false document would not have the needed document identification code within it. Checking identification against a database improves the vigilance factor and reduces human error.

Paper documents have numerous options for enhancing security. Laid lines in the paper give it a unique set of watermark lines. Color prismatic printing also makes the substitution of other paper stock extremely difficult due to the unique color background on the genuine paper. Void pantographs cause the words "VOID" or "UNAUTHORIZED COPY" to appear on a black and white or color copy made of the document. Warning bands indicate that if they do not appear in color, the document is not an original. High-resolution borders are common on negotiable instruments such as stocks and bonds. Holograms contain an image that is difficult to replicate and act as a seal of authenticity on the document especially on identification documents. This technology is also used on software packaging to prove the software is genuine.

Microprinting places very small text or identifiers on the document, which are difficult for forgers to see and to fake. Secure number fonts are fonts with unique bordering around them or with unique patterns within the numbering or letters. Any tampering with the numbers or letters is immediately visible. Paper watermarks, visible if the document is held up to light, are a method of identifying genuine paper stock.

Chemical voids in the paper stock make the word "VOID" appear if the document receives a washing in a chemical solution to remove pen ink. Check washing is a means for a forger to change the amounts on a check, as the washing does not affect the printer's ink.

Plastic cards used for identification and for credit cards also have several technologies to prevent forgery. Security designers employ microprinting, holograms, embossed characters, tamper-evident signature panels, ultraviolet inks, and magnetic stripes to increase security, and the ability to check the card against a database of stolen or lost cards and invalid card numbers is a key part of the validation process.

Electronic media have several options for increasing security. Smart cards, access cards, token cards, and special security badges fall into this category. These cards may have RFID tags embedded in them. Radio Frequency Identification requires a reader to detect the validating information on the tag. Secure ID cards require the card to synchronize with a security server before access is granted to say a network. Challenge-response technology can be a part of a card's authentication. The authenticating server or host computer sends a digital challenge to which the card's or token's chip has to respond. Only a genuine chip can make the correct response. Biometric security is another option. Usually, verifying the holder of a card via a fingerprint scan is the easiest method. Location-based security is also possible. An identification card or access card can only be validated from certain readers or computers. These points of validation have secure physical access, so someone using the card has to clear through a layer of physical security first. Even then the cardholder would still have to pass a biometric test or password authentication or both. The more layers of authentication the better, depending what is at risk. Actually, people having to present what they know (a password), what they have (the card or token), who they are (the fingerprint), and where they are (at a secure location) undergo four-factor authentication, an extremely secure method.

Computer forensics is also entering the campaign to fight digital forgeries. Professors Alin C. Popescu and Harry Farid at Dartmouth University have developed a method for scanning digital images to determine if elements or regions of the image have undergone tampering. While not yet foolproof, this technology shows promise for developing a whole range of tools for verifying the authenticity of digital documents.

Regardless of the technologies employed, any security staff that examines identification and other critical documents as a part of the security mission must become thoroughly familiar with the security features of the documents accepted on a regular basis. The best source of information on the validity of identification documents or other critical documents like automobile titles or birth certificates is the agency that issues them. Check with the agency on obtaining information on how to spot forgeries.

Determining the vectors of attack helps to protect your organization. Consider the documents that your enterprise uses on a regular basis in commercial transactions. If it is a matter of verifying identification for employment, follow the counsel of the previous paragraph and learn how to detect forgeries in the classes of identification documents that you will accept. Under no circumstances should you accept copies or electronic images of documents without seeing the originals first. Only accept identification issued by a governmental agency with a photographic likeness.

Academic records require production directly from the institution's registrar via mail and must bear the registrar's seal. If, for some reason, time constraints prevent mail delivery prior to the hire date, the applicant may provide a copy of his or her transcript, but the registrar's office must be contacted to verify the authenticity and correctness of the transcript. The applicant will have to sign a release to allow the registrar to provide the information. Academic degrees and diplomas must be confirmed in the same manner.

Medical records must be obtained directly from the institution or doctor's office by using a release executed by the patient. Records provided by the patient should not be the basis of important business decisions unless the institution or doctor confirms their content.

If legal records form a part of the regular business documents used by your organization, then a records service should obtain copies of real estate documents, Uniform Commercial Code filings, regulatory documents, court filings, and other public documents directly from the records depository. Copies or electronic images submitted by an interested party should not be accepted as the basis for business decisions unless a trusted records service provides confirming copies directly from the repository.

Know the negotiable instruments that your enterprise accepts. Learn how to spot forgeries based upon information provided by the issuer. For example, if your firm accepts American Express® credit cards and traveler's cheques, obtain the documentation from American Express on how to check for genuine instruments. Financial records such as financial statements, transaction journals, and bank statements need review and confirmation in an audit process by a qualified accountant if such documents serve as the basis for major business decisions. Never accept such records at face value without contacting independent third parties that can verify their content. Remember that someone with fundamental financial knowledge will have no problem crafting very impressive financial documents using desktop publishing software.

Invoices, purchase orders, bills, and accounts receivable or payable documents are all easily faked. Have in place a procedure for processing these purchasing documents received by mail. Only process documents from vendors on your firm's approved vendor list and only when there is a corresponding internal purchase order to authorize the expenditure. Only send payment to the approved address for the vendor, not to a different one appearing on the invoice.

Multiple avenues of attack are available to forgers when your firm does not confirm independently the content, validity, and genuineness of the documents presented. Taking documents at face value is a prescription for being a victim of fraud. Question and investigate all documentation submitted by outsiders when such documents serve as the basis of major business decisions.

REVIEWING AND VERIFYING DOCUMENTS

What follows are guidelines for examining documents in either electronic or physical form. Again, a copy or electronic version of a document is acceptable as preliminary information, but before a critical business decision occurs, the original of the document must undergo review, or independent sources must confirm its content. Here are the items to consider in review:

1. In electronic documents, check the document for metadata. (See Chapter 1 on "Metadata.) Look at inconsistencies between what the metadata says and the document's text. Evaluate oral statements made by the presenter in light of the metadata. If the document is an original composition according to the presenter, but the metadata indicates other authorship, then warning alarms should go off.

2. In reviewing electronic documents, check to see if templates or exemplars for that type of document are available on the Web. (Recall the discussion of drivers' license and other identity card exemplars in the section, "Identity Document Counterfeiting.") If an electronic document closely mimics the exemplar, with only the dates, names, and address different, ask for the original. Also, keep in mind many legal documents, ranging from Power of Attorney to deeds, have templates and exemplars available online, making their forgery quite easy.
3. Is it possible to confirm the content of the electronic document online? Researching names, addresses, and other identifiers through a Web search engine is one approach. Examining the archive of a Web site is another method. Web site archives are available through http://www,archive.org.
4. Check for a digital watermark in the document if you know that a genuine electronic document has one.
5. Check with the issuing authority or with a recognized guidebook as in the case of drivers' licenses.
6. With paper documents, check the paper for a watermark.
7. Are there other anti-forgery technologies in use on the paper document such as laid lines, color prismatic printing, void pantographs, warning bands, high-resolution borders, holograms, or microprinting? Documents, where the risk of forgery is a significant factor, such as birth certificates, automobile titles, negotiable instruments, and drivers' licenses or passports, should have one or more of these countermeasures in use.
8. If the document contains a considerable amount of information as in a legal document, is it possible to contact sources referenced in the document for confirmation of the content?
9. Are the dates in the document consistent?
10. Are names, descriptions, and other facts in the document internally consistent?
11. Are persons or businesses cited in the document verifiable by cross-checking them using a Web search engine, or by public record checking, or by telephone calling those cited?
12. Are the assets and liabilities listed in a financial statement or document verifiable? For example, if vehicles are listed as assets, can they be found in the motor vehicle registration database?

13. Are exemplars available online for the identification documents? What does the comparison tell you about the genuineness of the document?
14. Are document templates available online for the legal documents? What does the comparison tell you about the genuineness of the documents?
15. If you are skeptical about the validity of financial documents, arrange for a review by a forensic accountant.
16. Check with the issuing financial institution when you have questions about the genuineness of a negotiable instrument.

AN EXERCISE

When forgers make mistakes, those errors fall into these categories: mistakes in content, errors in form, and inferior physical quality. Anyone can download the template for a legal document from the Internet, but turning it into a convincing forgery requires a degree of mastery over content. Errors in form happen when the forger does little research on what the actual genuine document should look like. “Near misses” simply will not cut it, and they are the mark of a real amateur. If the physical forgery cannot match the genuine document in quality, then a knowledgeable examiner will spot the forgery right away. Very rarely will a forger hit the mark in all three areas. When he or she does accomplish that feat, it is the trademark of a real craftsman, albeit a rogue one.

This exercise concentrates on reviewing content in a questionable document. The forger seeks to lower his automobile insurance rate by submitting to his insurance agent an altered police report. In the original report, the investigating police officer has the forger as the at-fault party in an automobile accident. After obtaining the original report from the police department, the forger scans the two-page document into his desktop publishing program.

The first page of the document is the standard front sheet for a Texas Motor Vehicle Accident Report. Here the forger changes a few details on the form to indicate that the other party was at fault. The second page, the supplemental sheet, is the location for the investigating officer’s narrative. On this page, the forger, Carl Jaspers, needs to exercise his creativity. He crafts the following two paragraphs to replace the police officer’s narrative:

> This two-vehicle collision occurred at the intersection of Ruther and Cross streets at 8 p.m. The intersection is controlled by a traffic light. The conditions were nighttime darkness, but street lights provided normal city lighting, and the weather was clear. Carl Jaspers was northbound on Ruther and Peter Jackson was westbound on Cross. Mr. Jaspers stated at the scene he had the green light prior to the collision. (The traffic light appeared to be functioning normally.)
>
> An independent witness, Jeremy reader, DOB: 4-12-61, of 1219 Westheimer in Houston , Texas 77010, (713-555-5555) confirmed Mr. Jaspers' account. The witness was at the Quicky Fuel gas station at the corner of Cross and Ruther and saw the collision as he was fueling his vehicle. Mr. Jackson was cited for disregarding a traffic control signal.

Based upon content alone, the narrative sounds plausible. (Obviously, we have not supplied an actual telephone number for the witness to eliminate the possibility of publishing a private number.) Jaspers, however, made several mistakes. Can you suggest some avenues of inquiry to discover them?

The first problem with the narrative is that it does not include a statement from Mr. Jackson. Normally, a police officer would include statements from all involved drivers in a supplemental accident narrative. Jaspers did not want to put words in Jackson's mouth that he could repudiate later with a telephone call if someone contacted him. What is missing from a document is sometimes just as important as what it contains. Be aware of omissions that do not make sense when reviewing a document, whether in electronic or physical form.

Jaspers makes two factual errors in the statement. Ruther is actually a one-way street running from north to south, so Jaspers could not have been northbound. He got his directions turned around when trying to adapt the original report's language to suit the needs of the forgery. A city street map, available on the Web, would uncover this inconsistency. There is a Quicky Fuel station in the area, but it is not at the actual intersection. Rather, it is about one and a half blocks away, so anyone at the gas station would have had a problem witnessing the collision. A Web search engine check for Quicky Fuel stations in the city would reveal the various locations.

The witness information is pure fabrication. The zip code for 1219 Westheimer in Houston, Texas is 77006, not 77010, and 1219 Westheimer does not appear to be an address associated with a building or a residence. If we imagine the telephone number to be a real one, dialing the number reveals a number not in service, and running the telephone number on Google comes up with nothing. Finally, running Jeremy Reader with DOB of 4-12-61 comes up with nothing on Google. Running the same information on a public records database on the Web yields nothing in Texas.

A call to the Municipal Court reveals no ticket on file for Mr. Jackson for the date of the accident. However, checking under Mr. Jaspers' name reveals a ticket for the same offense on the date of the accident. Obviously, this exercise is a simple one to demonstrate that content in documents is not that difficult to check, and content can be a major stumbling block for a forger because it takes time and considerable effort in some cases to acquire the necessary knowledge.

Consider, for example, if someone wanted to forge a handwritten letter from Shakespeare to a theatrical friend. The difficulties would be immense. First, one would have to obtain paper from the period and inks appropriate to the time period. This would be the physical challenge. Obtaining the materials would be very difficult to be sure, but possible. Marketing an electronic image of the document may be feasible on the Internet. Yet, at some point, the physical characteristics of the forgery would have to pass expert review.

More difficult would be the form of the letter. The forger would have to understand the conventions and style of Elizabethan correspondence, including spelling, word usage, and the scripting of the letters. The forger would also need to copy successfully William Shakespeare's handwriting. Exemplars of his signature exist, but creating a script would be an immense undertaking requiring remarkable skill.

If these problems were not enough, a forger would have profound content issues. He or she would have to understand Shakespeare's vocabulary and style, and most difficult, have Shakespeare's understanding of his times, the Elizabethan stage, and the slang and colloquialisms of his era. This level of knowledge simply does not happen after reading a paperback book or two on Shakespeare's life.

Content is always difficult, especially when the document has a narrative structure or a number of interrelated details. Pay close attention to content details in reviewing any document used for a major business decision. Look for errors and inconsistencies, and obtain confirmation of the contents from independent sources whenever possible. (See Table 10-1: Spotting Forgeries.)

Table 10.1: Spotting Forgeries

Characteristic	**Details**
Metadata (Electronic documents)	Inconsistent authorship, inconsistent dates. Good for spotting the use of templates.
Content and Facts	Confirmable by other sources? Any internal inconsistencies?
Templates and Exemplars Online	Look for them being copied from online and the forger just filling in the blanks.
Paper Document Countermeasures	• Laid lines • Color prismatic printing • Void Pantographs • Warning bands • High-resolution borders • Holograms • Microprinting
Plastic Cards	• Matching account numbers against a database • Microprinting • Holograms • Embossed characteristics • Tamper-evident signature panels • Ultraviolet inks • Magnetic stripes
Issuing Agency Guidelines	Obtain the training and information on how to spot a forgery.

Appendix

SECURITY POLICIES FOR DOCUMENT SECURITY

Network Security has its own outline or schema for security policies that range in topics from Router Security to Acceptable Use to Firewall Configuration. This outline concentrates solely upon topics appropriate to the major issue of Document Security. To organize the topics conceptually, the outline moves from the outermost area of security concern, the world external to the organization's security boundary, to the innermost assets of the organization.

Policies

External

- Documents Permitted in External Environment
- Document and Media Disposal
- Encryption Policies
- Mobile Devices
- Travel Security for Sensitive Information

Perimeter

- Public Security Zone
- Business Channels Monitoring and Security

Internal

- Internal Security Zone
- Document Classification
- Physical Security for Documents
- Media Reuse

Inner Core

- Sensitive Security Zone
- Confidential Zone
- High Security Zone

The general concerns in drafting security policies for Document Security are as follows:

1. *Clarity.* Policies must have a high level of readability. Employees must understand what is expected of them without jargon or highly technical language.
2. *Accessibility.* Policies should be on an Intranet or a Wiki where employees may access them quickly and search them using keywords.
3. *Relevancy.* If policies are not kept current, they become irrelevant and only hinder security understanding.
4. *Standards must be defined.* Adherence to a policy requires setting standards by which employees can conform. Failing to adopt agreed upon standards to measure adherence creates confusion and a lack of credibility for the security effort.
5. *Responsibility.* Persons or job roles responsible for implementing a policy must be identified. Saying it is everyone's responsibility means no one will take responsibility. Define who is responsible for which components of the policy.
6. *Tools and resources.* Describe all tools and resources necessary for the correct implementation of the policy.
7. *Criteria.* Establish the criteria for discriminating between documents and information assets. By which criteria will the security team classify documents? What factors determine the assigning of documents to particular security zones?
8. *Auditing and testing.* Describe the methods the organization will use to verify compliance to the Document Security policies on a periodic basis.

BIBLIOGRAPHY

Articles

Akapose, Wole: "E-mail security: A Review of Available Technologies," *The ISSA Journal*, February 2006. (Contains a section on CAPTCHA.)

Anastasia, George: "Big Brother and the Bookie," *Mother Jones* (Jan–Feb 2002).

Andress, Jason: "Secure Data Deletion and Recovery," *The ISSA Journal*, January 2007.

Lambrecht, Bill: "Discarded U.S. Computers for Sale in Nigeria, along with their Secrets," *St. Louis Post-Dispatch*, December 17, 2006.

Manning, Stephen: "The Biggest Threat to Computer Security? Carelessness," *Austin American-Statesman*, June 19, 2006.

Meinel, Carolyn: "How Hackers Break In . . . and How They Are Caught," *Scientific American*, October 1998.

Mendell, Ronald L.: "And, the Floppies Spoke for the Victim," SecurityPortal, September 1999. (The article is no longer on the Web. Article covered floppies and magnets.)

Mendell, Ronald L.: "Intelligence Gathering for ITSEC Professionals," *The ISSA Journal*, December 2005.

Mendell, Ronald L.: "Internet Rhetoric for Security," *The ISSA Journal*, August 2006.

Pham, Alex: "Bullies Invade Even the Virtual World," *Austin American-Statesman*, September 23, 2002.

Reichenberg, Nimrod: "Seven Steps to Secure USB Drives," *The ISSA Journal*, January 2007.

Tyson, Dave: "Geeks and Guards: Leveraging the Corporate Guard Force." *The ISSA Journal*, August, 2006.

Washington Post staff: Article on "Multi-State Anti-Terrorism Information Exchange." *Austin American-Statesman*, August 6, 2003.

Wired staff: "Foil a Snooping Boss," August 2006.

Books

Aristotle: *The Art of Rhetoric.* (translated by H.C. Lawson-Tanced), New York: Penguin, 1991.

Brown, Christopher L.T.: *Computer Evidence: Collection & Preservation.* Hingham, MA: Charles River Media, 2006.

Caloyannides, Michael A.: *Desktop Witness.* Indianapolis, IN: John Wiley & Sons, Inc., 2002.

Capaldi, Nicholas: *The Art of Deception.* Buffalo, NY: Prometheus Books, 1979.

Casey, Eoghan: *Digital Evidence and Computer Crime,* 2nd Edition. San Diego, CA: Academic Press, 2004.

Graves, Kimberly: *CEH Official Certified Ethical Hacker Review Guide.* Indianapolis, IN: Sybex, 2007.

Harris, Shon: *CISSP.* Emeryville, CA: Berkeley, McGraw-Hill/Osborne, 2002.

Mendell, Ronald L.: *How To Do Financial Asset Investigations,* 3rd edition. Springfield, IL: Charles C. Thomas, 2006.

Mendell, Ronald L.: *Investigating Computer Crime in the 21st Century,* 2nd edition. Springfield, IL: Charles C. Thomas, 2004.

Mitnick, Kevin D. and Simon, William L.: *The Art of Deception.* Indianapolis, IN: Wiley Publishing, 2002.

Mitnick, Kevin D. and Simon, William L.: *The Art of Intrusion.* Indianapolis, IN: Wiley Publishing, 2006.

Pirie, Madsen: *How To Win Every Argument: The Use and Abuse of Logic.* New York: Continuum, 2006.

Schifreen, Robert: *Defeating the Hacker.* West Sussex, England: John Wiley & Sons, 2006.

Scott, Robert: *The Investigator's Little Black Book 3.* Beverly Hills, CA: Crime Time Publishing, 2002.

Solomon, Micheal G., Barrett, Diane, Broom, Neil: *Computer Forensics Jump Start™.* Sybex, 2005.

The Knightmare: *Secrets of a Super Hacker.* Post Townsend, WA: Loompanics Unlimited, 1994.

Vacca, John R.: *Computer Forensics: Computer Crime Scene Investigation.* Hingham, MA: Charles River Media, 2002.

Data Sheets

PortAuthority Technologies,"Information Risk Assessment: Information Leak Prevention for the Enterprise," 2005.

Reconnex, "Feature Overview," 2006.

Reconnex, "Information Protection. Always.," 2006.

Reconnex, "iGuard 3600 Data Sheet," 2006.

Electronic Documents

Ayers, Rick and Jansen, Wayne: "Guidelines on PDA Forensics." NIST, August, 2004.

Guel, Michael D.: "A Short Primer for Developing Security Policies." SANS Institute, 2001.

Keller, Alex: "Google Hacking: A Crash Course," (PowerPoint file), Alex Keller is a Network/Systems Administrator for BSS Computing at San Francisco State University.

Shirani, Bram, CISSP: "Anti-forensics." PowerPoint presentation, HTCIA Spring Training, 2002.

In Electronic Format on the Web

Appligent White Paper, "The Case for Content Security," located at http://www.appligent.com/docs/tech/ContentSecurity.pdf.

Barron, Anne, "Three Easy Steps for Gathering Intelligence at Trade Shows," located at http://www.scipstore.org/scipstore.org_asp//news/cimp/v3i4article1.asp

CAPTCHA Tutorial, located at http://www.captcha.biz/.

Digimarc®, "Combating Identity Document Counterfeiting," located at http://www.digimarc.com/govt/docs/dmrc_wp_combating.pdf.

Digimarc, "Enhancing Personal Identity Verification with Digital Watermarks," located at http://csrc.nist.gov/piv-program/FIPS-201-Public-Comments/digimarc.pdf.

Goodell, Jeff, "How to Fake a Passport," New York Times Magazine, February 10, 2002, located at http://www.globalpolicy.org/nations/citizen/2002/0210fake.htm.

Guel, Michele D., "The SANS Policy Primer," (PDF format), located at http://www.sans.org/resources/policies.

Kissel, Richard et al: "Guidelines for Media Sanitization," (PDF file), National Institute of Standards and Technology (NIST) Special Publication 800-88, September 2006 located at http://csrc.nist.gov/publications/nistpubs/800-88/NISTSP800-88_rev1.pdf

Long, Johnny, "Google Hacking Mini-Guide," located at http://www.informit.com/articles/article.asp?p=170880&rl=1.

Masse, Robert and Wang, Jian Hui, "Hacking with Google for Fun and Profit!," located at http://www.gosecure.ca/SecInfo/library/WebApplication/GOOGLE-HACKING-GS1004.ppt

Metadatarisk, "The Dangers of Document Metadata: The Risks to Corporations," located at http://www.metadatarisk.org/collateral/content_security_risks/US_Brief_Dangers_of_Document_Metadata.pdf .

Microsoft Knowledge Base, "The Remove Hidden Data Tool for Office 2003 and Office XP," located at http://support.microsoft.com/kb/834427.

Microsoft Knowledge Base, "How to Minimize Metadata in MS Word 2000 Documents," located at http://support.microsoft.com/kb/237361.

Microsoft Knowledge Base, "How to Minimize Metadata in Excel Documents," located at http://support.microsoft.com/kb/223789.

Microsoft Knowledge Base, "How to Minimize Metadata in PowerPoint Documents," located at http://support.microsoft.com/kb/314797.

Microsoft Knowledge Base, "How to Minimize Metadata in PowerPoint 2002 Documents," located at http://support.microsoft.com/kb/314800.

Mitek Systems, "Combating Check Forgery," located at http://miteksystems.com/pdf/FPS%20White_paper.pdf.

National Security Agency., "Redacting with Confidence: How to Safely Publish Sanitized Reports Converted from Word to PDF," located at http://www.fas.org/sgp/othergov/dod/nsa-redact.pdf.

Olzak, Tom, "Fundamentals of Storage Media Sanitation Part One," June 2006, located at http://blogs.ittoolbox.com/security/adventures/archives/fundamentals-of-storage-media-sanitation-part-1-9407.

Olzak, Tom, "Fundamentals of Storage Media Sanitation Part Two," June 2006, located at http://blogs.ittoolbox.com/security/adventures/archives/fundamentals-of-storage-media-sanitation-part-2-9559.

Olzak, Tom, "Fundamentals of Storage Media Sanitation Part Three," June 2006, located at http://blogs.ittoolbox.com/security/adventures/archives/fundamentals-of-storage-media-sanitation-part-3-9680.

Payne, Donna, "Control Metadata in Your Legal Documents," located at http://office.microsoft.com/en-us/assistance/HA011400341033.aspx.

Payne, Donna and Lewis, Bruce, "EDD Showcase: Metadata: Are You Protected?," Law Technology News, August 2004, located at http://www.lawtechnews.com/r5/showkiosk.asp?listing_id=430591.

Popescu, Alin C. and Farid, Harry, "Exposing Digital Forgeries by Detecting Duplicated Image Regions," located at http://www.ists.dartmouth.edu/library/tr-2004-515.pdf.

Popular Science, "Debunking 9/11 Myths: Introduction with Forward by John McCain," http://www.popularmechanics.com/technology/military_law/3491861.html.

Smith, Russell G., "Criminal Exploitation of New Technologies," Trends & Issues, Australian Institute of Criminology, July 1998, No. 93, located at http://www.aic.gov.au/publications/tandi/ti93.pdf.

Smith, Russell G., "Identity-related Economic Crime: Risks and Countermeasures," Trends & Issues, Australian Institute of Criminology, September 1999, No. 129, located at http://www.aic.gov.au/publications/tandi/ti129.pdf.

Sollicito, Michelle Johnston, "Securing Your Laptop," located at http://www.informit.com/articles/article.asp?p=174137&rl=1

Taylor, Laura, "PDA Security 101," located at http://www.intranetjournal.com/articles/200304/ij_04_07_03a.html

Ward, Mark., "The Hidden Dangers of Documents," 18 August 2003, (BBC Web page), located at http://news.bbc.co.uk/2/hi/technology/3154479.stm.

Wikipedia, Article on CAPTCHA, located at http://en.wikipedia.org/wiki/CAPTCHA.

Unpublished Manuscript

Grayhat Research Corporation, Advanced Information Security Course, October 2006.

Useful Web Sites

Adobe (the security of PDF files) located at http://www.adobe.com/security.

"A Handbook of Rhetoric Devices" located at http://www.virtualsalt.com/ rhetoric.htm.

"Antisemitism on the Internet" located at http://www.jugendschutz.net/pdf/osce_berlin.pdf.

ARIN (for checking IP addresses in North America) located at http://www.arin.

net/whois.

Better Business Bureau located at http://www.bbb.org.

Guidestar (for charitable organizations) located at http://www.guidestar.org.

Hoaxbusters at http://hoaxbusters.ciac.org/.

Metadatarisk.org (on the dangers of Metadata) located at http://www.metadatarisk.org.

Stealth Products located at http://www.computersecurity.com/stealth/computer_tracker.htm and http://www.computersecurity.com/stealth/data_protector.htm.

"www.terror.net: How Modern terrorism Uses the Internet" located at http://www.usip.org/pubs/specialreports/sr116.pdf.

INDEX

Charles C Thomas

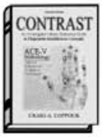
P.O. Box 19265
Springfield, IL 62794-9265

Book Savings* Save 10% , Save 15% , Save 20% !

- Ax, Robert K. & Thomas J. Fagan—**CORRECTIONS, MENTAL HEALTH, AND SOCIAL POLICY: International Perspectives.** '07, 402 pp. (7 x 10), 5 il., 16 tables.
- Campbell, Terence W.—**ASSESSING SEX OFFENDERS: Problems and Pitfalls. (2nd Ed.)** '07, 362 pp. (7 x 10), 46 tables.
- Hicks, Wendy, L.—**POLICE VEHICULAR PURSUITS: Constitutionality, Liability and Negligence**. '07, 136 pp. (7 x 10), 8 tables.
- Jurkanin, Thomas J., Larry T. Hoover, & Vladimir A. Sergevnin—**IMPROVING POLICE RESPONSE TO PERSONS WITH MENTAL ILLNESS.** '07, 220 pp. (7 x 10), 4 il., 17 tables, (spiral) paper.
- McEntire, David A.—**DISCIPLINES, DISASTERS AND EMERGENCY MANAGEMENT: The Convergence and Divergence of Concepts, Issues and Trends from the Research Literature.** '07, 368 pp. (8 x 10), 8 il., 14 tables.
- Mendell, Ronald L.—**DOCUMENT SECURITY: Protecting Physical and Electronic Content.** '07, 174 pp. (7 x 10), 8 il., 19 tables.
- Schmidt, Linda M. & James T. O'Reilly—**GANGS AND LAW ENFORCEMENT: A Guide for Dealing with Gang-Related Violence.** '07, 200 pp. (7 x 10), 21 il.
- Violanti, John M.—**POLICE SUICIDE: Epidemic in Blue. (2nd Ed.)** '07, 252 pp. (7 x 10), 7il., 2 tables.
- Violanti, John M. & Stephanie Samuels—**UNDER THE BLUE SHADOW: Clinical and Behavioral Perspectives on Police Suicide.** '07, 222 pp. (7 x 10), 1 il.
- Williams, Howard E.—**TASER® AND SUDDEN IN-CUSTODY DEATH: Separating Evidence from Conjecture.** '07, 294 pp. (8 x 10), 15 il.

- Campbell, Andrea and Ralph C. Ohm—**LEGAL EASE: A Guide to Criminal Law, Evidence, and Procedure. (2nd Ed.)** '07, 376 pp. (7 x 10), 31 il., $79.95, hard, $59.95, paper.

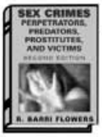
- Colaprete, Frank A.—**INTERNAL INVESTIGATIONS: A Practitioner's Approach.** '07, 358 pp. (7 x 10), 61 il., $82.95, hard, $55.95, paper.



- Coppock, Craig A.—**CONTRAST: An Investigator's Basic Reference Guide to Fingerprint Identification Concepts. (2nd Ed.)** '07, 210 pp. (7 x 10), 66 il., $53.95, hard, $33.95, paper.

- Ellis, John W.—**POLICE ANALYSIS AND PLANNING FOR HOMICIDE BOMBINGS: Prevention, Defense, and Response. (2nd Ed.)** '07, 372 pp. (7 x 10), 68 il., $69.95, hard, $49.95, paper.

- Means, Kevin P.—**TACTICAL HELICOPTER MISSIONS: How to Fly Safe, Effective Airborne Law Enforcement Missions.** '07, 136 pp. (7 x 10), 53 il., $34.95, paper.



- Flowers, R. Barri—**SEX CRIMES: Perpetrators, Predators, Prostitutes, and Victims. (2nd Ed.)** '06, 312 pp. (7 x 10), 19 il., 18 tables, $63.95, hard, $43.95, paper.

- Franklin, Carl J.—**THE INVESTIGATOR S GUIDE TO COMPUTER CRIME.** '06, 328 pp. (7 x 10), 19 il., 2 tables, $63.95, hard, $43.95, paper.

- Mendell, Ronald L.—**HOW TO DO FINANCIAL ASSET INVESTIGATIONS: A Practical Guide for Private Investigators, Collections Personnel, and Asset Recovery Specialists. (3rd Ed.)** '06, 240 pp. (7 x 10), 13 il., 13 tables, $54.95, hard, $34.95, paper.

- Paton, Douglas & David Johnston—**DISASTER RESILIENCE: An Integrated Approach.** '06, 344 pp. (7 x 10), 22 il., 9 tables, $68.95, hard, $48.95, paper.

- Rivers, R. W.—**EVIDENCE IN TRAFFIC CRASH INVESTIGATION AND RECONSTRUCTION: Identification, Interpretation and Analysis of Evidence, and the Traffic Crash Investigation and Reconstruction Process.** '06, 324 pp. (8 x 10), 175 il., 13 tables, $86.95, hard, $56.95, paper.

- Haggin, Daniel J.—**ADVANCED DUI INVESTIGATION: A Training and Reference Manual.** '05, 236 pp. (8 1/2 x 11), 10 il., $34.95, (spiral) paper.

- McDevitt, Daniel S.—**MANAGING THE INVESTIGATIVE UNIT.** '05, 210 pp. (7 x 10), 2 tables, $51.95, hard, $31.95, paper.

5 easy ways to order!

PHONE: 1-800-258-8980 or (217) 789-8980

FAX: (217) 789-9130

EMAIL: books@ccthomas.com
Web: www.ccthomas.com

MAIL:
Charles C Thomas •
Publisher, Ltd.
P.O. Box 19265
Springfield, IL 62794-9265

Complete catalog available at ccthomas.com • books@ccthomas.com

Books sent on approval • Shipping charges: $7.75 min. U.S. / Outside U.S., actual shipping fees will be charged • Prices subject to change without notice

*Savings include all titles shown here and on our web site. For a limited time only.