Computer Forensics

In This Issue

Computer Forensics: Digital Forensic Analysis Methodology. 1
By Ovie L. Carroll, Stephen K. Brannon, and Thomas Song

Vista and BitLocker and Forensics! Oh My!. 9
By Ovie L. Carroll, Stephen K. Brannon, and Thomas Song

Demystifying the Computer Forensic Process for Trial: (Is My Witness Dr. Jekyll or Mr. Hyde?). 29
By Martin J. Littlefield

Managing Large Amounts of Electronic Evidence. 46
By Ovie L. Carroll, Stephen K. Brannon, and Thomas Song

Rethinking the Storage of Computer Evidence. 60
By Tyler Newby and Ovie L. Carroll

January
2008
Volume 56
Number 1

United States
Department of Justice
Executive Office for
United States Attorneys
Washington, DC
20530

Kenneth E. Melson
Director

Contributors' opinions and statements should not be considered an endorsement by EOUSA for any policy, program, or service.

The United States Attorneys' Bulletin is published pursuant to 28 CFR § 0.22(b).

The United States Attorneys' Bulletin is published bimonthly by the Executive Office for United States Attorneys, Office of Legal Education, 1620 Pendleton Street, Columbia, South Carolina 29201.

Managing Editor
Jim Donovan

Program Manager
Nancy Bowman

Internet Address
www.usdoj.gov/usao/reading_room/foiamanuals.html

Send article submissions and address changes to Program Manager, United States Attorneys' Bulletin,
National Advocacy Center,
Office of Legal Education,
1620 Pendleton Street,
Columbia, SC 29201.

Computer Forensics: Digital Forensic Analysis Methodology

Ovie L. Carroll
Director, Cybercrime Lab
Computer Crime and Intellectual Property Section
Criminal Division

Stephen K. Brannon
Cybercrime Analyst, Cybercrime Lab
Computer Crime and Intellectual Property Section
Criminal Division

Thomas Song
Senior Cybercrime Analyst, Cybercrime Lab
Computer Crime and Intellectual Property Section
Criminal Division

I. Introduction

In comparison to other forensic sciences, the field of computer forensics is relatively young. Unfortunately, many people do not understand what the term computer forensics means and what techniques are involved. In particular, there is a lack of clarity regarding the distinction between data *extraction* and data *analysis*. There is also confusion about how these two operations fit into the forensic process. The Cybercrime Lab in the Computer Crime and Intellectual Property Section (CCIPS) has developed a flowchart describing the digital forensic analysis methodology. Throughout this article, the flowchart is used as an aid in the explanation of the methodology and its steps.

The Cybercrime Lab developed this flowchart after consulting with numerous computer forensic examiners from several federal agencies. It is available on the public Web site at www.cybercrime.gov/forensics_gov/forensicschart.pdf. The flowchart is helpful as a guide to instruction and discussion. It also helps clarify the elements of the process. Many other resources are available on the section's public Web site, www.cybercrime.gov. In addition, anyone in the Criminal Division or U.S Attorneys' offices can find additional resources on the new intranet site, CCIPS Online. Go to DOJ Net and click on the "CCIPS Online" link. You can also reach us at (202) 514-1026.

II. Overview of the digital forensics analysis methodology

The complete definition of computer forensics is as follows: "The use of scientifically derived and proven methods toward the preservation, collection, validation, identification, analysis, interpretation, documentation and presentation of digital evidence derived from digital sources for the purpose of facilitating or furthering the reconstruction of events found to be criminal…." A Road Map for Digital Forensic Research, Report from the First Digital Forensic Research Workshop (DFRWS), *available at* http://dfrws.org/2001/dfrws-rm-final.pdf.

Defining computer forensics requires one more clarification. Many argue about whether computer forensics is a science or art. *United States v. Brooks*, 427 F.3d 1246, 1252 (10th Cir. 2005) ("Given the numerous ways information is stored on a computer, openly and surreptitiously, a search can be as much an art as a science."). The argument is unnecessary, however. The tools and methods are scientific and are verified scientifically, but their use necessarily involves elements of ability, judgment, and interpretation. Hence, the word "technique" is often used to sidestep the unproductive science/art dispute.

The key elements of computer forensics are listed below:

- The use of scientific methods
- Collection and preservation
- Validation
- Identification
- Analysis and interpretation
- Documentation and presentation

The Cybercrime Lab illustrates an overview of the process with Figure 1. The three steps, Preparation/Extraction, Identification, and Analysis, are highlighted because they are the focus of this article. See Figure 1, page 5.

In practice, organizations may divide these functions between different groups. While this is acceptable and sometimes necessary, it can create a source of misunderstanding and frustration. In order for different law enforcement agencies to effectively work together, they must communicate clearly. The investigative team must keep the entire picture in mind and be explicit when referring to specific sections.

The prosecutor and forensic examiner must decide, and communicate to each other, how much of the process is to be completed at each stage of an investigation or prosecution. The process is potentially iterative, so they also must decide how many times to repeat the process. It is fundamentally important that everyone understand whether a case only needs preparation, extraction, and identification, or whether it also requires analysis.

The three steps in the forensics process discussed in this article come after examiners obtain forensic data and a request, but before reporting and case-level analysis is undertaken. Examiners try to be explicit about every process that occurs in the methodology. In certain situations, however, examiners may combine steps or condense parts of the process. When examiners speak of lists such as "Relevant Data List," they do not mean to imply that the lists are physical documents. The lists may be written or items committed to memory. Finally, keep in mind that examiners often repeat this entire process, since a finding or conclusion may indicate a new lead to be studied.

III. Preparation/Extraction

See Figure 2, page 5.

Examiners begin by asking whether there is enough information to proceed. They make sure a clear request is in hand and that there is sufficient data to attempt to answer it. If anything is missing, they coordinate with the requester. Otherwise, they continue to set up the process.

The first step in any forensic process is the validation of all hardware and software, to ensure that they work properly. There is still a debate in the forensics community about how frequently the software and equipment should be tested. Most people agree that, at a minimum, organizations should validate every piece of software and hardware after they purchase it and before they use it. They should also retest after any update, patch, or reconfiguration.

When the examiner's forensic platform is ready, he or she duplicates the forensic data provided in the request and verifies its integrity. This process assumes law enforcement has already obtained the data through appropriate legal process and created a forensic image. A forensic image is a bit-for-bit copy of the data that exists on the original media, without any additions or deletions. It also assumes the forensic examiner has received a working copy of the seized data. If examiners get original evidence, they need to make a working copy and guard the original's chain of custody. The examiners make sure the copy in their possession is intact and unaltered. They typically do this by verifying a hash, or digital fingerprint, of the evidence. If there are any problems, the examiners consult with the requester about how to proceed.

After examiners verify the integrity of the data to be analyzed, a plan is developed to extract data. They organize and refine the forensic request into questions they understand and can answer. The forensic tools that enable them to answer these questions are selected. Examiners generally have preliminary ideas of what to look for, based on the request. They add these to a "Search Lead List," which is a running list of requested items. For example, the request might provide the lead "search for child pornography." Examiners list leads explicitly to help focus the examination. As they develop new leads, they add them to the list, and as they exhaust leads, they mark them "processed" or "done."

For each search lead, examiners extract relevant data and mark that search lead as processed. They add anything extracted to a second list called an "Extracted Data List." Examiners pursue all the search leads, adding results to this second list. Then they move to the next phase of the methodology, identification.

IV. Identification

See Figure 3, page 6.

Examiners repeat the process of identification for each item on the Extracted Data List. First, they determine what type of item it is. If it is not relevant to the forensic request, they simply mark it as processed and move on. Just as in a physical search, if an examiner comes across an item that is incriminating, but outside the scope of the original search warrant, it is recommended that the examiner immediately stop *all* activity, notify the appropriate individuals, including the requester, and wait for further instructions. For example, law enforcement might seize a computer for evidence of tax fraud, but the examiner may find an image of child pornography. The most prudent approach, after finding evidence outside the scope of a warrant, is to stop the search and seek to expand the warrant's authority or to obtain a second warrant.

If an item is relevant to the forensic request, examiners document it on a third list, the Relevant Data List. This list is a collection of data relevant to answering the original forensic request. For example, in an identity theft case, relevant data might include social security numbers, images of false identification, or e-mails discussing identity theft, among other things. It is also possible for an item to generate yet another search lead. An e-mail may reveal that a target was using another nickname. That would lead to a new keyword search for the new nickname. The examiners would go back and add that lead to the Search Lead List so that they would remember to investigate it completely.

An item can also point to a completely new potential source of data. For example, examiners might find a new e-mail account the target was using. After this discovery, law enforcement may want to subpoena the contents of the new e-mail account. Examiners might also find evidence indicating the target stored files on a removable universal serial bus (USB) drive—one that law enforcement did not find in the original search. Under these circumstances, law enforcement may consider getting a new search warrant to look for the USB drive. A forensic examination can point to many different types of new evidence. Some other examples include firewall logs, building access logs, and building video security footage. Examiners document these on a fourth list, the New Source of Data list.

After processing the Extracted Data list, examiners go back to any new leads developed. For any new data *search* leads, examiners consider going back to the Extraction step to process them. Similarly, for any new *source* of data that might lead to new evidence, examiners consider going all the way back to the process of obtaining and imaging that new forensic data.

At this point in the process, it is advisable for examiners to inform the requester of their initial findings. It is also a good time for examiners and the requester to discuss what they believe the return on investment will be for pursuing new leads. Depending on the stage of a case, extracted and identified relevant data may give the requester enough information to move the case forward, and examiners may not need to do further work. For example, in a child pornography case, if an examiner recovers an overwhelming number of child pornography images organized in user-created directories, a prosecutor may be able to secure a guilty plea without any further forensic analysis. If simple extracted and identified data is not sufficient, then examiners move to the next step, analysis.

V. Analysis

See Figure 4, page 7.

In the analysis phase, examiners connect all the dots and paint a complete picture for the requester. For every item on the Relevant Data List, examiners answer questions like who, what, when, where, and how. They try to explain which user or application created, edited, received, or sent each item, and how it originally came into existence. Examiners also explain where they found it. Most importantly, they explain why all this information is significant and what it means to the case.

Often examiners can produce the most valuable analysis by looking at when things happened and producing a timeline that tells a coherent story. For each relevant item, examiners try to explain when it was created, accessed, modified, received, sent, viewed, deleted, and launched. They observe and explain a sequence of events and note which events happened at the same time.

Examiners document all their analysis, and other information relevant to the forensic request, and add it all to a fifth and final list, the "Analysis Results List." This is a list of all the meaningful data that answers who, what, when, where, how, and other questions. The information on this list satisfies the forensic request. Even at this late stage of the process, something might generate new data search leads or a source of data leads. If this happens, examiners add them to the appropriate lists and consider going back to examine them fully.

Finally, after examiners cycle through these steps enough times, they can respond to the forensic request. They move to the Forensic Reporting phase. This is the step where examiners document findings so that the requester can understand them and use them in the case. Forensic reporting is outside the scope of this article, but its importance can not be overemphasized. The final report is the best way for examiners to communicate findings to the requester. Forensic reporting is important because the entire forensic process is only worth as much as the information examiners convey to the requester. After the reporting, the requester does case-level analysis where he or she (possibly with examiners) interprets the findings in the context of the whole case.

VI. Conclusion

As examiners and requesters go through this process, they need to think about return on investment. During an examination, the steps of the process may be repeated several times. Everyone involved in the case must determine when to stop. Once the evidence obtained is sufficient for prosecution, the value of additional identification and analysis diminishes.

It is hoped that this article is a helpful introduction to computer forensics and the digital forensics methodology. This article and flowchart may serve as useful tools to guide discussions among examiners and personnel making forensic requests. The Cybercrime Lab in the Computer Crime and Intellectual Property Section (CCIPS) is always available for consultation. CCIPS personnel are also available to assist with issues or questions raised in this article and other related subjects.❖

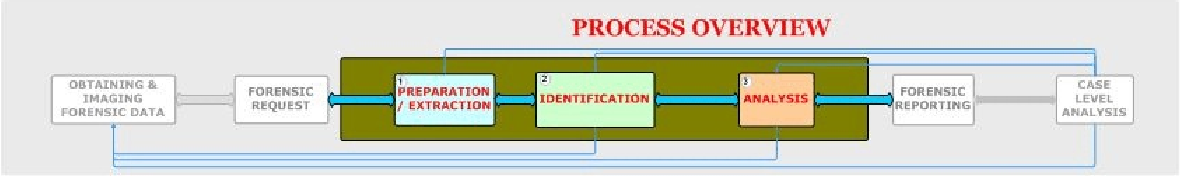

PROCESS OVERVIEW
OBTAINING & IMAGING FORENSIC DATA
FORENSIC REQUEST
1 PREPARATION / EXTRACTION
2 IDENTIFICATION
3 ANALYSIS
FORENSIC REPORTING
CASE LEVEL ANALYSIS

FIGURE 1

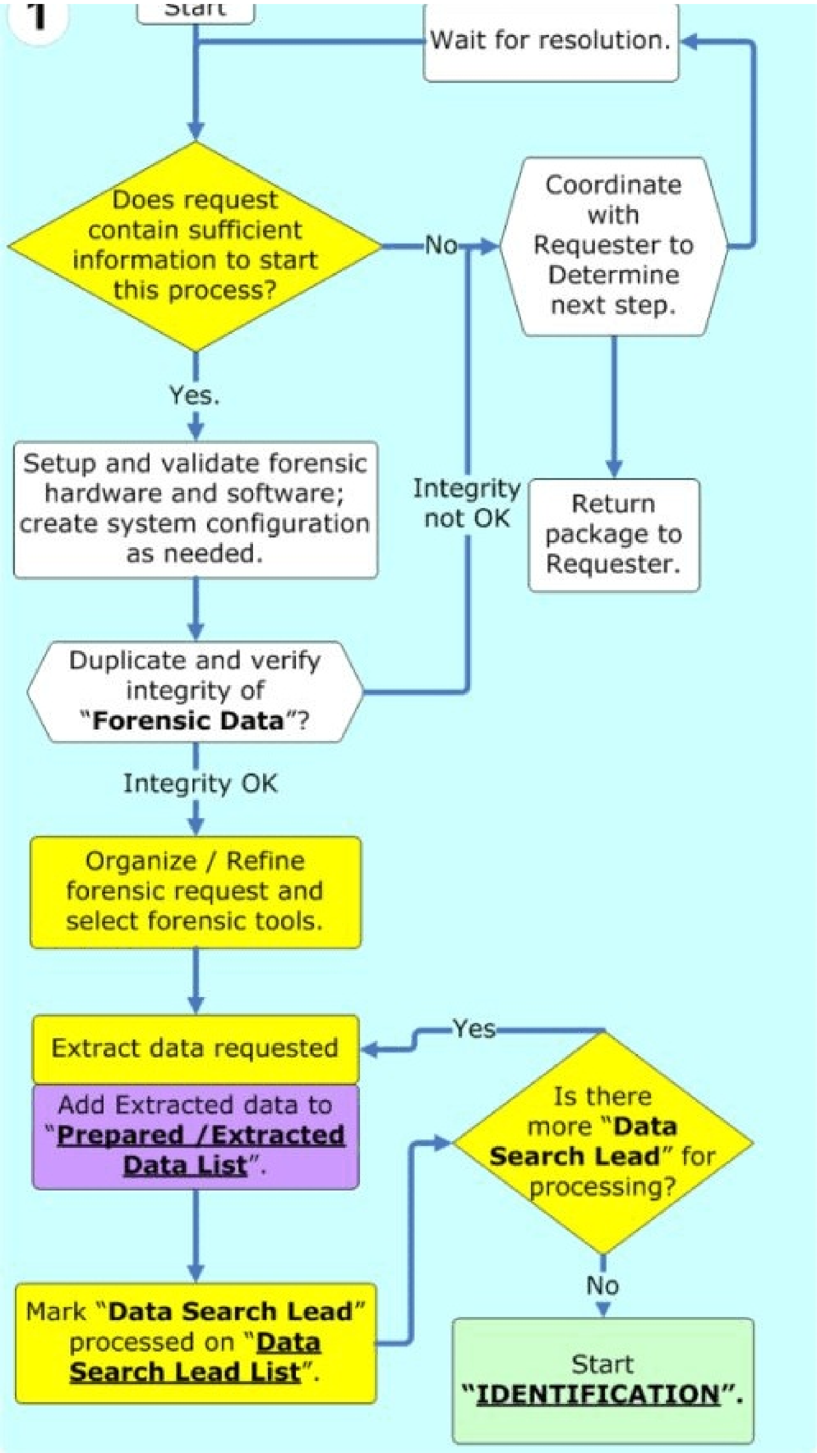

1
Start
Wait for resolution.
Does request contain sufficient information to start this process?
No
Coordinate with Requester to Determine next step.
Yes.
Setup and validate forensic hardware and software; create system configuration as needed.
Integrity not OK
Return package to Requester.
Duplicate and verify integrity of "Forensic Data"?
Integrity OK
Organize / Refine forensic request and select forensic tools.
Extract data requested
Yes
Add Extracted data to "Prepared /Extracted Data List".
Is there more "Data Search Lead" for processing?
No
Mark "Data Search Lead" processed on "Data Search Lead List".
Start "IDENTIFICATION".

Figure 2

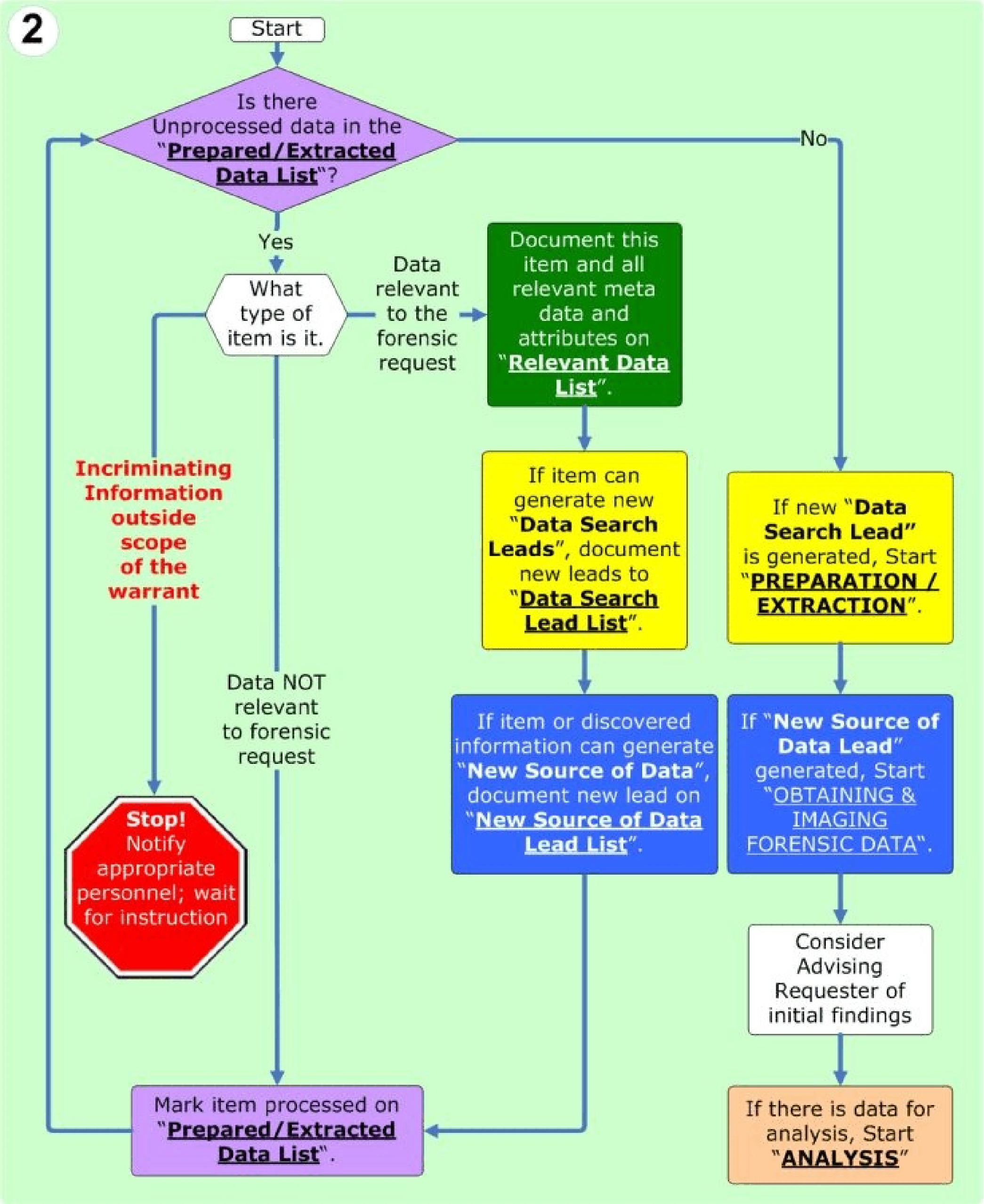

2
Start
Is there Unprocessed data in the "Prepared/Extracted Data List"?
No
Yes
What type of item is it.
Data relevant to the forensic request
Document this item and all relevant meta data and attributes on "Relevant Data List".
If item can generate new "Data Search Leads", document new leads to "Data Search Lead List".
If new "Data Search Lead" is generated, Start "PREPARATION / EXTRACTION".
Incriminating Information outside scope of the warrant
Data NOT relevant to forensic request
If item or discovered information can generate "New Source of Data", document new lead on "New Source of Data Lead List".
If "New Source of Data Lead" generated, Start "OBTAINING & IMAGING FORENSIC DATA".
Stop! Notify appropriate personnel; wait for instruction
Consider Advising Requester of initial findings
Mark item processed on "Prepared/Extracted Data List".
If there is data for analysis, Start "ANALYSIS"

Figure 3

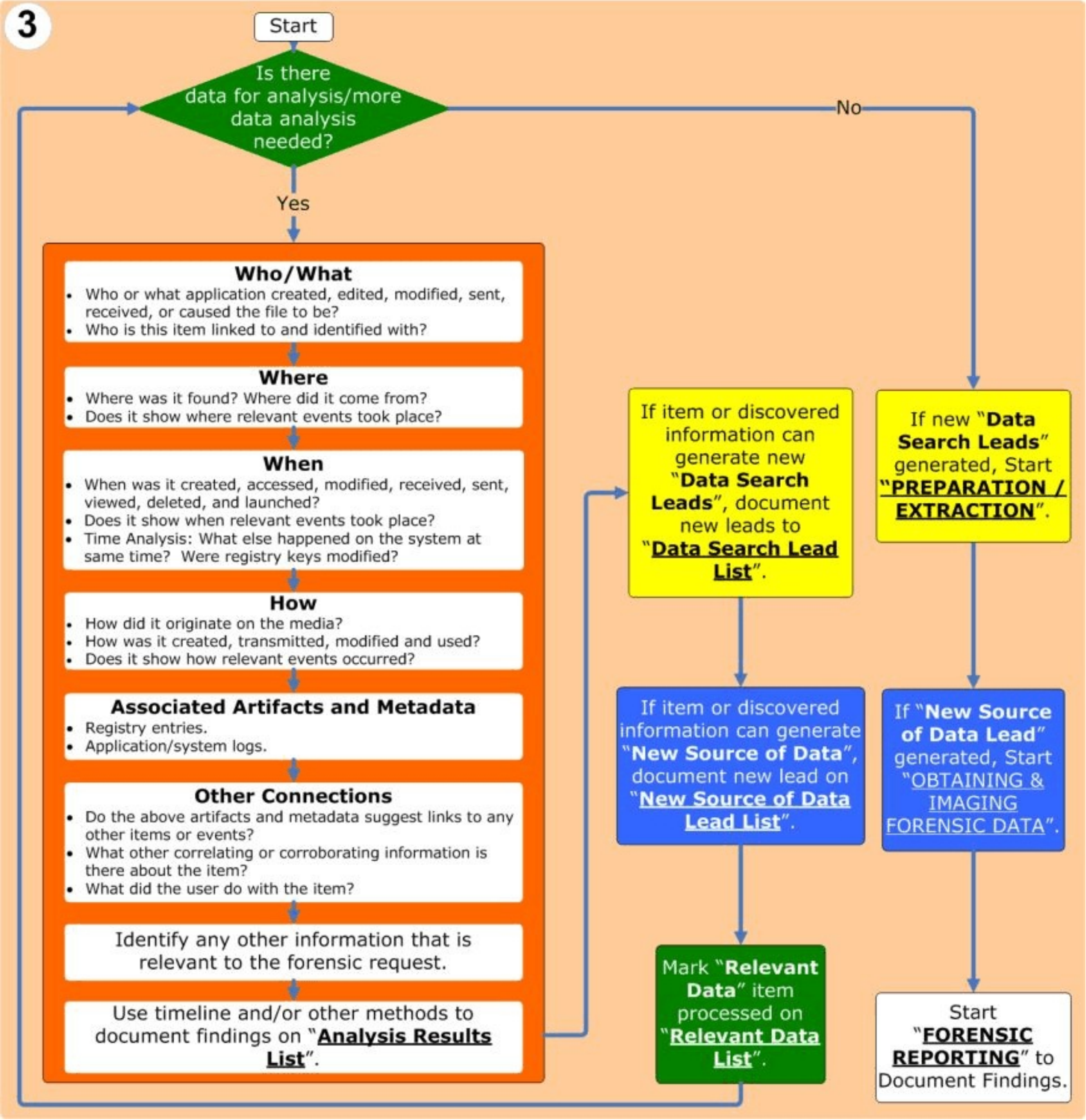

3
Start
Is there data for analysis/more data analysis needed?
No
Yes
Who/What
Who or what application created, edited, modified, sent, received, or caused the file to be?
Who is this item linked to and identified with?
Where
Where was it found? Where did it come from?
Does it show where relevant events took place?
When
When was it created, accessed, modified, received, sent, viewed, deleted, and launched?
Does it show when relevant events took place?
Time Analysis: What else happened on the system at same time? Were registry keys modified?
How
How did it originate on the media?
How was it created, transmitted, modified and used?
Does it show how relevant events occurred?
Associated Artifacts and Metadata
Registry entries.
Application/system logs.
Other Connections
Do the above artifacts and metadata suggest links to any other items or events?
What other correlating or corroborating information is there about the item?
What did the user do with the item?
Identify any other information that is relevant to the forensic request.
Use timeline and/or other methods to document findings on "Analysis Results List".
If item or discovered information can generate new "Data Search Leads", document new leads to "Data Search Lead List".
If item or discovered information can generate "New Source of Data", document new lead on "New Source of Data Lead List".
Mark "Relevant Data" item processed on "Relevant Data List".
If new "Data Search Leads" generated, Start "PREPARATION / EXTRACTION".
If "New Source of Data Lead" generated, Start "OBTAINING & IMAGING FORENSIC DATA".
Start "FORENSIC REPORTING" to Document Findings.

Figure 4

ABOUT THE AUTHORS

❑**Ovie L. Carroll** is the Director of the Cybercrime Lab in the CCIPS. He has over twenty years of law enforcement experience. He previously served as the Special Agent in Charge of the Technical Crimes Unit at the Postal Inspector General's Office and as a Special Agent with the Air Force Office of Special Investigations.

❑**Stephen K. Brannon** is a Cybercrime Analyst in the CCIPS's Cybercrime Lab. He has worked at the Criminal Division in the Department of Justice and in information security at the FBI.

❑**Thomas Song** is a Senior Cybercrime Analyst in the CCIPS's Cybercrime Lab. He has over fifteen years in the computer crime and computer security profession. He specializes in computer forensics, computer intrusions, and computer security. He previously served as a Senior Computer Crime Investigator with the Technical Crimes Unit of the Postal Inspector General's Office.

The Cybercrime Lab is a group of technologists in the CCIPS in Washington, DC. The lab serves CCIPS attorneys, Computer Hacking and Intellectual Property (CHIP) units in the U.S. Attorneys' offices, and Assistant U.S. Attorneys, by providing technical and investigative consultations, assisting with computer forensic analysis, teaching, and conducting technical research in support of Department of Justice initiatives.✠

Vista and BitLocker and Forensics! Oh My!

Ovie L. Carroll
Director, Cybercrime Lab
Computer Crime and Intellectual Property Section
Criminal Division

Stephen K. Brannon
Cybercrime Analyst
Cybercrime Lab
Computer Crime and Intellectual Property Section
Criminal Division

Thomas Song
Senior Cybercrime Analyst
Cybercrime Lab
Computer Crime and Intellectual Property Section
Criminal Division

I. Introduction

For almost a year now, many in the forensic community and crime fighting world have been buzzing about Microsoft's new operating system, Vista, its new encryption utility, BitLocker, and the implications it will have on computer forensics and cybercrime fighting. The following information is an attempt to ease the fears of some, the panic of others, and educate many.

Readers may contact the Computer Crime and Intellectual Property Section (CCIPS), or the CCIPS Cybercrime lab, if they have further questions or need assistance. The section and lab can be reached at (202) 514-1026, or via our public website, www.cybercrime.gov. Employees of the Criminal Division and U.S. Attorneys' offices can also access additional resources on our new intranet site, CCIPS Online. From the DOJ Net home page, click the "CCIPS Online" link.

II. Basic, Home, Premium, and Super Duper. What is with all the versions?

The version a consumer chooses is based on the features desired. All but Vista Starter are available in either a 32-bit or 64-bit version.

A. Vista Starter

This version is only available on preloaded, lower-cost systems, through original equipment manufacturers (OEM) and Microsoft OEM distributors in 139 countries considered to be undeveloped technology markets. Vista Starter is a minimally-featured operating system, with its primary features being basic Internet browsing, communications, media player, basic photo editing, and one of the newest features to the operating system, parental controls. The parental controls are a part of every version of Vista and allow a user with administrator privileges to create a different set of controls or restrictions for each user of the system. The areas that can be controlled include web restrictions, time limits, games, and the ability to allow or block specific programs.

B. Vista Home Basic

The Vista version of Windows XP Home version and Home Basic comes loaded with Microsoft's more secure Internet Explorer, Version 7, Windows Defender, and improved firewall capabilities. Starting with Home Basic, all additional Vista versions include Windows Movie Maker, Media Player, Version 11, Windows Mail (the successor of Outlook Express), Windows Contacts, and Windows Calendar.

C. Vista Home Premium

Vista Home Premium is the primary consumer version and is the most likely version law enforcement will encounter outside of a business environment. Home Premium is the first version of Vista that incorporates the new aero glass interface and Windows Media Center. Home Premium also allows users to back up personal files to hard disk, CD/DVD, or a networked drive.

D. Vista Business

This version is the successor to Windows XP Professional. It focuses more on the type of business features available in XP Pro. Vista Business supports connecting to a corporate domain, encrypting files, remote desktop connectivity, roaming user profiles, and the use of Windows shadow copy.

E. Vista Enterprise

This version is available only for volume licensing through Microsoft, and is not anticipated to be available in retail markets. It incorporates all of the features found in Vista Business and includes BitLocker drive encryption.

F. Vista Ultimate

This is the flagship version of Vista and includes everything in Home Basic, Premium, Business, and Enterprise, and then adds several premium products that do not seem significant to forensic examinations.

III. The disk

A. File structure

One of the first changes forensic examiners will notice is the new file structure. Gone are the days of "Documents and Settings" and the myriad of "My Stuff." Microsoft has apparently concluded that the user is intelligent enough to figure out that the files are theirs, so they have dropped the "My" from all user folders. Under the user folder now are Contacts, Desktop, Documents, Downloads, Favorites, Links, Music, Pictures, Saved Games, Searches, and Videos. This is meant to be a flatter file system and easier to navigate. See Figure 1, page 18.

Another significant change to the file system is that Vista no longer tracks last access times. This was done in an attempt to increase system performance. This last access time can be reenabled by modifying the "HKLM\System\CurrentControlSet\Control\FileSystem\NtfsDisableLastAccessUpdate," but in most forensic cases, it will simply not be available. See Figure 2, page 19.

Another feature that should be of interest to law enforcement is the default configuration settings of the defragment program. By default, defrag is scheduled to execute every Wednesday at 3 a.m. Law enforcement should take into consideration that most users do not modify the default settings of the defrag launch. Consequently, it may be advantageous to execute search warrants prior to Wednesday evenings. One last note, of probably little significance to the forensic examiner, is that the first New Technology File System (NTFS) partition starts at sector 2048, rather than sector 63.

B. BitLocker

Of all the new features in Vista, law enforcement personnel are most concerned about BitLocker whole-disk encryption. There are two common fears concerning this software.

- Law enforcement investigators will be unable to forensically image and analyze information from a hard drive with BitLocker enabled.
- Law enforcement will be overwhelmed with the volume of BitLocker encrypted drives.

BitLocker (also known as Full Volume Encryption) is a security mechanism designed to protect data stored on computing devices running Windows Vista, in the event they are stolen, lost, or otherwise physically compromised. This security technology allows an administrator to specify 128 bit or 256 bit Advanced Encryption Standard (AES) for the contents of the volume(s) protected. A system protected by BitLocker will typically require the user to either supply a startup key stored on a Universal Serial Bus (USB) flash drive, or enter a personal identification number (PIN) (up to twenty digits) during the startup process, in order for the system to boot. On newer computers, the key will typically be stored on a hardware device called the Trusted Platform Module (TPM) security hardware, which is a special microchip in the computers that supports advanced security features. The boot process requires the system to unlock a series of keys that are encrypted on the BitLocker protected volume (in the file system metadata), making access to these keys very difficult.

When BitLocker is enabled, and before the volume is encrypted, the BitLocker management interface requires the user to create a recovery password, in the event all other access to the volume fails. This recovery password is a forty-eight digit numeric password that can be stored in a number of ways from the BitLocker interface (printed to paper or a file, saved to a USB flash drive, saved to a folder). When deployed in an enterprise environment (the most typical deployment expected), administrators can require that the implementation of BitLocker call back to the enterprise management infrastructure (Microsoft Windows Active Directory) to store copies of the startup key and/or recovery password. Law enforcement should note that, when dealing with enterprise systems that employ BitLocker, the password recovery key will typically be stored and viewable on the enterprise server.

Microsoft is currently offering an online secure key back up service that allows users of BitLocker to upload their password recovery key, in the event they lose their copy. It is expected that other non-Microsoft affiliated vendors will also offer this service. In order to obtain such keys, law enforcement will obviously have to use the appropriate legal process.

C. BitLocker issues affecting search and seizure

When the computer is started, the TPM chip provides the decryption key for the partition only after comparing a hash of several operating system configuration values. If the drive is removed from the computer it was encrypted on and placed in another computer system, the drive will not decrypt without the password recovery key. Additionally, if changes are detected in the basic input output system (BIOS), or any of the startup files, the TPM will not release the decryption key and the drive will not unlock without the password recovery key, all of which may cause challenges to the forensic examiner if the password recovery key is not available.

If the computer does not have a TPM chip, the encryption and decryption key can be stored on a USB drive. The flash drive would subsequently have to be inserted into the computer every time the computer is booted or comes out of hibernation. One additional challenge that BitLocker can present is its ability to combine the need for a USB storage device and a user-generated four to twenty-digit PIN. Law enforcement must consider including, in the scope of their warrant, the increased authority to search for, and seize, entire computer systems, if BitLocker is suspected or detected. Additionally, at the search scene, investigators must look for USB storage devices of any kind, as well as any written or printed documentation of the BitLocker password recovery key.

BitLocker is capable of encrypting other partitions and removable media, such as external hard drives and thumb drives, among other things. There is no documentation available, at this time, on encrypting external storage media, however, and it is not currently a Microsoft supported feature. The partition that contains the operating system may be encrypted with BitLocker, but it will be some time before external storage devices encrypted with BitLocker are encountered.

While the above information sounds like a formidable challenge for law enforcement, there are many reasons to allay concerns about the ability to image or analyze drives with BitLocker encryption and of the significant increases in the volume of BitLocker encrypted data that will be encountered.

- BitLocker is only available on two versions of Vista Enterprise and Ultimate.
- BitLocker is not presented to the user or administrator at any time during the installation process of the operating system and, therefore, would only be configured and enabled if the administrator knows about it and searches for the configuration application.
- If the administrator wishes to enable BitLocker encryption, setup and configuration is not intuitive. This may be negated by the use of Microsoft's recently released "BitLocker Drive Preparation Tool" which is part of the Ultimate Extra's free download.
- Finally, encryption is still viewed by many computer users as scary because of the potential loss of their own data. Until hardware vendors, such as Dell and Hewlett-Packard, start shipping computers with BitLocker preconfigured, or Microsoft develops an easy-to-follow configuration wizard that is presented to the user during installation, law enforcement will not likely see a dramatic increase in BitLocker encrypted disks.

D. Considerations for changes to incident response procedures

What can be done to determine if a live computer system is encrypted using BitLocker or some other disk or volume? The Department of Homeland Security funded the Software Engineering Institute at Carnegie Mellon University, and the researchers have come up with a very small seventeen-kilobyte tool called "Crypthunter." This file, when executed from the administrative command prompt on a running system, will report the presence of the sixteen different volume-based encryption programs and eight disk-encryption programs, including

BitLocker. Crypthunter will also alert the user if there are indications on the disk that suggest other, possibly unknown, disk or volume encryption is enabled. More information about Crypthunter can be found at http://www.cert.org/forensics.

If the incident responder is aware that disk encryption is active on the computer system, there are several possible paths available to law enforcement. The responder can navigate to the BitLocker key management screen and save a copy of the password recovery key to a USB storage device, or print it if the system is connected to a printer.

For years, some in federal law enforcement, and many in the private sector incident response profession, have been developing incident response procedures to include the collection of volatile data. Thanks to the increased level of awareness BitLocker has brought to the gradual proliferation of whole disk encryption, law enforcement agencies will likely modify their current practices of "pulling the plug," and graduate to a more tactical approach of imaging RAM and collecting other volatile data prior to powering down the computer system. The first step in the collection of volatile data is the capture of RAM. An excellent tool and resource for information on imaging RAM from Vista systems is George Garner's KNTdd site, http://www.gmgsystemsinc.com/knttools/.

Another option to use in the collection of volatile data is to follow these steps:

- Click on the start button, known in Vista as the "pearl."
- Type "BitLocker" in the search bar (clicking on the start pearl by default puts you at the search bar).
- Select "BitLocker Drive Encryption," and select continue when warned this requires administrative privileges.
- Select "Turn off BitLocker."

One additional technique might be to run the below listed cscript command from an administrative command prompt. While these commands will unlock the drives, it leaves them in their encrypted state, and merely stores the Volume Master Key in the clear so that the system can boot without a startup key:

- cscript manage-bde.wsf—unlock c:
- cscript manage-bde.wsf—autounlock—enable c:

E. Can a BitLockered drive be imaged?

Yes. If a BitLockered drive is imaged, the drive will only be able to be read or analyzed after the password recovery key is provided. One technique to consider is to obtain a logical image of the drive while the system is live. A logical image is easily created using either Access Data's Forensic Tool Kit Imager or Guidance Software's EnCase Imager.

IV. Thumbs.db

Since Windows 95, all versions (except Vista) have created a thumbs.db file. The thumbs.db file is a database that contains an image of every thumbnail it displayed. Forensic examiners routinely analyze the thumbs.bd files for evidence of images that were once located in a directory, but have since been deleted. The concept of creating a database of thumbnails to display in the thumbnail view has been completely revamped and improved. Microsoft Vista now creates four files; Thumbcache_32.db, Thumbcache_96.db, Thumbcache_256.db and Thumbcache_1024.db, all of which are stored in a single location, the %userprofile%\AppData\Local\Microsoft\WindowExplorer. The new Thumbcache files now contain thumbnails of every folder the user views. Unlike previous versions of Windows, this includes cameras and external storage devices like USB drives, among other things. This allows the forensic examiner to see all thumbnails users of the computer have viewed, and attribute the viewing to each user's credentials. See Figure 3, page 19.

V. Recycle bin

The Vista recycle bin is in the same location as previous recycle bins, but the name has been changed to "$Recycle.bin." By default, Vista allocates 7 percent of the drive size to the recycle bin. Forensic examiners will quickly find that the familiar "Info2" file is gone. In the Vista Recycle bin, examiners will find "$Ixxxxx" (dollar sign capital I) and "$Rxxxxx" files. An additional feature of the Vista recycle bin is the ability to handle/track the deletion of items on mapped network drives. The files that were deleted can be found in the "$Rxxxxx" files. The actual date and time the file was deleted can be identified by analyzing the eight bytes following hex offset 10 in the "$Ixxxxx" file. The full original path of the file can also be found in this file.

VI. Internet Explorer feature—clearing all evidence with one click

All versions of Vista come with the new, more secure, Internet Explorer 7. Forensic examiners will be happy to know the "Typed URL" registry key can still be found in the "HKU\<GUID>\Software\ \Microsoft\Internet Explorer\TypedURLs" registry key. Additionally, a record of pop-ups authorized by the user from each Web site can now be found in the HKU\<GUID>\Software\Microsoft\Internet Explorer\New Window\Allow" registry key. The location of the temporary Internet files, the directory that caches images and pages previously visited, and favorites or book marked Web sites, has moved, and can now be found in the "%userprofile%\AppData\Local\Microsoft\ Windows\Temporary Internet Files" and "%userprofile%\Favorites" respectively. Another change to Internet Explorer 7 is its redesign for deleting browsing history. As seen in Figure 1, the deletion utility now includes a single "Delete all…" button, which deletes all cookies, history, form data, and saved passwords. Rather than just deleting the temporary Internet files, it zeros out the index.dat file, making it extremely difficult to recover any usable data. See Figure 4, page 20.

VII. Disk clean up utility

The Vista disk clean up utility has been improved. Unfortunately for law enforcement, it now includes the ability to delete the following files:

- Program files
- Temporary Internet files
- Offline Web pages
- Hibernation files
- Setup logs
- Temporary files
- Thumbnails
- Archived Windows error reports
- Empty the recycle bin

By default, the utility deletes downloaded programs, temporary Internet files, and thumbnails. See Figure 5, page 21.

VIII. Event logs

Event logging in Vista has undergone a complete redesign. Like most Microsoft products, event logging has adhered to legacy application program interfaces (APIs) to insure backwards compatibility. There are more than fifty event logs stored in the %SystemRoot%\System32\winevt \Logs directory and they can easily be viewed in XML format through the event viewer interface. Because event logs are stored in .evtx format, examiners attempting to use the Microsoft Log Parser will discover that tool will not work.

IX. Restore points

Windows creates snapshots of the system (beginning with Windows ME), also known as system restore points, at regular intervals, for the user to roll back to, in the event something happens that makes the system unstable or inoperable. Vista continues the tradition of creating restore points at the following intervals:

- Every twenty-four hours of computer uptime

- When Windows Update/Microsoft Update is started
- Before installation of an unsigned driver
- Before installation of applications that call Volume Shadow Service (VSS) API
- Before starting any backup operation
- Before starting the restore process
- When manually created by the user

By default, Windows dedicates 12 percent of the drive for restore points which are saved to the "%SystemRoot%\system volume information" directory and cannot be accessed by the user while the system is running. Included in the restore points are complete copies of the registry, a copy of any unsigned driver or application that is loaded, and select .ini files. As such, restore points are a wealth of information for forensic examiners and can provide ample opportunities to look into the past through the examination of previous versions of the registry archived in the restore points. See Figure 6, page 22.

X. Previous versions

"Previous Versions" is a part of the Volume Shadow Copy Service available in Vista Business, Enterprise, and Ultimate versions. Shadow copies are copies of files that have been modified since the last system restore point was made. Shadow copies are also copies of files on the computer, or shared files on other computers, on a network. This new feature in Vista has great potential to help law enforcement identify and document previous versions of files or folders. It is active by default, and saves the current state of user files when a volume snapshot is made.

While this will not be as granular as saving every version of a saved document, it does provide a lot more potential information than ever before. The presence of previous versions can be identified when in the operating system by right clicking on the file or folder, then selecting "restore previous versions." Vista will present a list of all previous versions and the date of that version. The user has the option to open, copy, or restore, any of the previous versions. With previous versions, it may be possible to restore a shadow copy of a file or complete folder that was deleted, even after the recycle bin has been emptied. The one caveat is that the examiner must know the original location of the file or folder. Initial testing has shown that if previous versions of a file are available, and the file is moved to a new location on the hard drive, the list of previous versions will appear empty. To see the previous versions, return the file to its original location and the list of previous versions will again be displayed to the user. This presents an interesting opportunity for forensic examiners to mount the volume or volume image to their forensic workstation and examine significant files for previous versions. A warning about restoring previous versions: if the user chooses to restore a previous version instead of opening or copying, all other versions will be lost. See Figure 7, page 23.

XI. The registry

The registry is essentially a database of system and application configuration information. It also maintains a great deal of information about events occurring on a computer, such as what files have recently been opened, media files played, and USB storage devices that have been plugged in, among other things. No significant changes have been observed in the Vista registry, although it does appear there are several new data points that are recorded. The registry has only recently become a recognized gold mine of information by law enforcement, and some in the field have made a concerted effort to become experts in registry forensics. One of the "go to" people for registry information and custom tools is Harlan Carvey. His Web site, *available at* http://windowsir.blogspot.com, contains a great deal of valuable forensic information and links to several free tools he has created, usually written in pearl, and an Excel spreadsheet consisting of "keys of interest" useful to forensic examiners and incident responders.

XII. Outlook Express is expunged

Windows Mail has replaced Outlook Express as the default mail client that ships with Microsoft operating systems. Windows Mail stores account information for each mail account created by the user in subdirectories of the %UserProfile%\ AppData\ Local\ Microsoft\Windows Mail directory. Each e-mail or new account will have a unique name with an ".oeaccount" extension. For example, "account{B84DA09C-7482-4144-A71E-D3EB3F65CDD1}.oeaccount" is the unique name of a Gmail account data file. Account settings are easily identified, as shown below. From this file it is possible to identify the mail account, user name, mail servers, and settings such as, if a copy of the mail is to remain on the mail server and for how many days.

```
<?xml version="1.0" encoding="utf-16" ?>

<MessageAccount>

 <Account_Name type="SZ">GMail</Account_Name>

 <Connection_ type="DWORD">00000003</Connection_Type>

 <IMAP_Dirty type= "DWORD">00000001</IMAP_Dirty>

 <POP3_Server type= "SZ">pop.gmail.com</POP3_Server>

 <POP3_User_Name type = "SZ">oviecarroll@gmail.com</POP3_User_Name>

 <POP3_Password2 type="BINARY">encrypted none of your business</POP3_Password2>

 <POP3_Port type= "DWORD">000003e3</POP3_Port>

 <POP3_Secure_Connection type="DWORD">00000000</POP3_Secure_Connection>

 <POP3_Timeout type= "DWORD">0000003c</POP3_Timeout>

 <Leave_Mail_On_Server type="DWORD">00000001</Leave_Mail_On_Server> (a 1 indicates this feature is active, a zero would indicate inactive)

 <Remove_When_Deleted type= "DWORD">00000000</Remove_When_Deleted>

 <Remove_When_Expired type="DWORD">00000001</Remove_When_Expired>

 <Expire_Days type="DWORD">00000005</Expire_Days> (the 5 represents the number of days mail is to remain on the mail server before it is deleted)

 <POP3_Skip_Account type="DWORD">00000000</POP3_Skip_Account>

 <POP3_Prompt_for_Password type="DWORD">00000000</POP3_Prompt_for_Password>

(SMTP mail settings would follow in similar format as above)

</MessageAccount>
```

All e-mail for an account is stored in the "WindowsMail.MSMessageStore" file located in the %userprofile%\AppData\Local\Microsoft\ Windows Mail directory. A review of all e-mail for that account can be accomplished by copying the WindowsMail.MSMessageStore to a Vista examination machine or virtual environment, and placing it in a sterile %userprofile%\AppData\ Local\Microsoft\Windows Mail directory, then simply opening Windows Mail from the examination platform.

As with Outlook Express, examiners may come across corrupt mail store files. Corrupt mail can be repaired and recovered by copying the Extensible Storage Engine Utilities against a copy of the corrupted WindowsMail.MSMessageStore. Simply copy the corrupted WindowsMail. MSMessageStore file to a suitable Vista examination environment and execute the following ESENTUTL.exe commands from an administrative command prompt; esentutl /p <full

path\WindowsMail.MSMessageStore. See Figure 8, page 24.

XIII. Prefetch

Prefetching is the process of loading information from the hard drive into memory, before it is needed. Vista adds six prefetch files. This does not sound significant, however, it is six more chances to identify information that may be important to the investigation. How can this help law enforcement? The %systemRoot%\Windows\Prefetch directory contains a list of up to the last 134 applications that were launched outside of the Windows startup group, their setup instructions/variables, the date and time they were last launched, and the total number of times they have been executed on the system. In an investigation, this enumeration of activity can point an investigator or examiner to recently used programs. Imagine the value of identifying an otherwise covert application used to facilitate a crime that the user took steps to hide. This application can be removed from the program's listing or perhaps even be a stand-alone application carried on a thumb drive. In one investigation, the defendant was identified using a portable Firefox browser on a thumb drive to surf the Internet, without leaving any temporary Internet cache or other evidence on the office computer. Investigators obtained a warrant to search the portable thumb drive, and found that it contained significant evidence of criminal activity, as well as incriminating bookmarks

By examining the application prefetch file, located in the windows\prefetch directory using a hex editor, the name of the application at offset 16(d)/10(h) is visible and the last execution time is an eight-byte value starting at 128(d)/80(h). To find the total number of times the application has been run, look at the four-byte value starting at offset 152(d)/98(h), then subtract five. For some reason, Microsoft starts counting at six. The way to remember this is that Bill Gates' last name is five characters and at Microsoft, everything comes after Gates. See Figure 9, page 25.

One note of caution. On a system that has more than one user account, the prefetch file does not identify which user launched the application. In order to find that, look at the User Assist key in the registry. See Figure 10, page 26.

XIV. Office 2007

Microsoft has completely renovated its Office line of applications. The most significant change to the Office applications is the format in which they are saved. Microsoft Word, Power Point, and Excel, are now saved in a compressed XML format. Examiners will quickly see that the file extensions are four character extensions, docx, pptx, and xlsx. When these files are examined with forensic software such as Access Data's Forensic Tool Kit, examiners will see that the file header is a compressed file, using some form of PKZip. The compressed file structure contains several directories and XML files. In the example, all the written content in the file is under the "word" subdirectory in the file named "documents.xml." Each image in the document is inside the "word\media" directory. Analysis of the new Office 2007 compressed XML files revealed that some small files, such as additional graphics, can be manually hidden inside the compressed file structure without being displayed when the document is viewed with its intended Office application. If the document is opened and any changes are made to it, Office 2007 will audit the contents of the compressed XML file and delete any files that do not belong. When analyzing the file, Forensic Tool Kit will automatically carve all images out of the compressed file and display them under the graphics tab. See Figure 11, page 27.

XV. Conclusion

In conclusion, while law enforcement will need to give consideration to updating their incident response procedures to allow for the collection of volatile data, there does not appear to be any changes in Vista that will negatively affect computer forensics.

Prosecutors interested in these and other computer forensic issues and techniques may register for the Computer Forensics for Prosecutors Course taught by CCIPS at the National Advocacy Center. The CCIPS and the Cybercrime Lab are also available to AUSAs for consultation on computer forensic and other technical investigative matters, by calling (202) 514-1026. Many other resources are available on our section's public Web site, www.cybercrime.gov. In addition, anyone in the Criminal Division or U.S. Attorney's office can find additional resources on our new intranet site, CCIPS Online. Just go to DOJ Net and click on the "CCIPS Online" link.

XVI. References

http://www.microsoft.com/windows/products/windows vista/editions/choose.mspx

http://www.securityfocus.com/print/infocus/1890

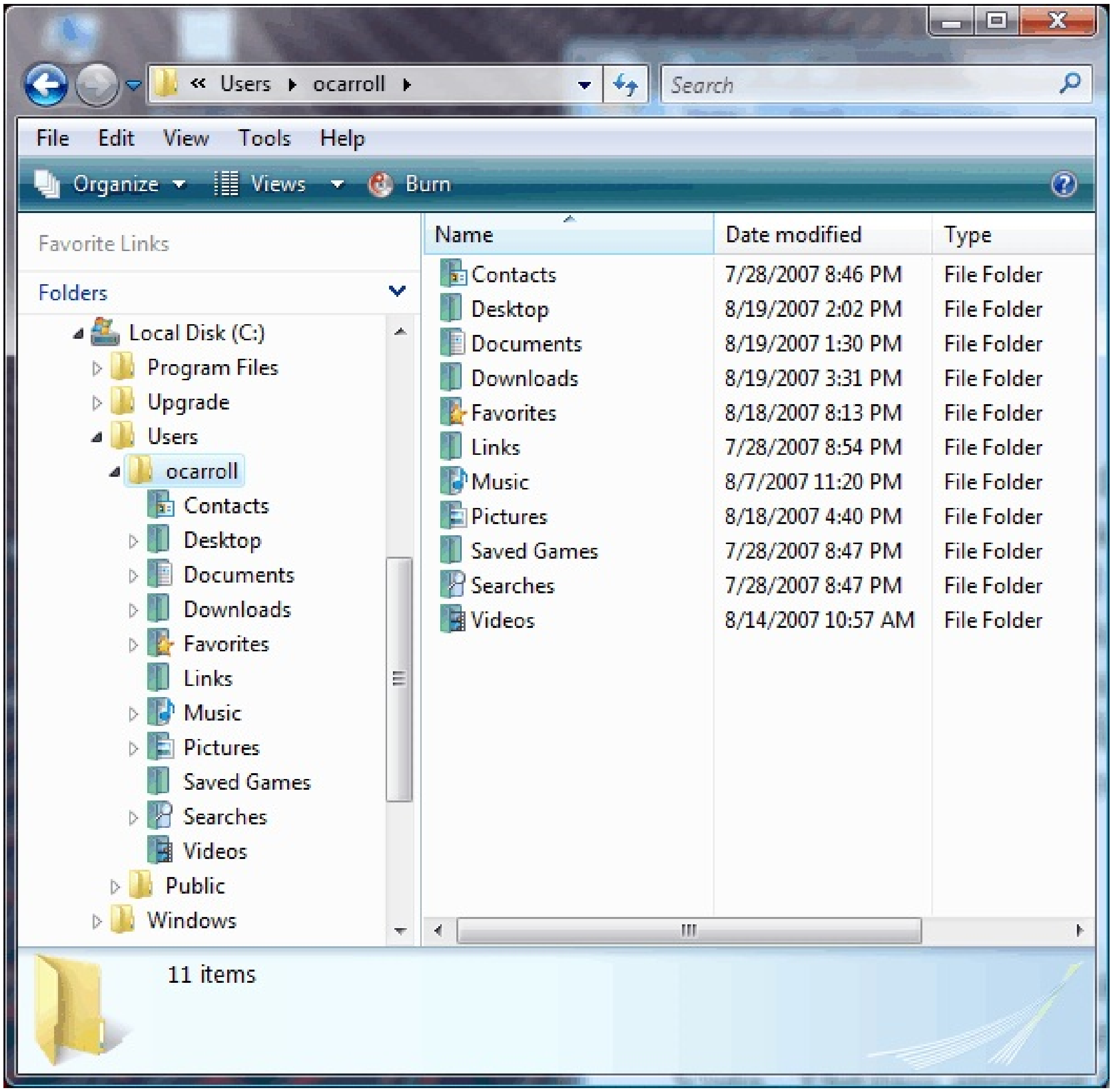
http://www.securityfocus.com/print/infocus/1889

http://technet2.microsoft.com/WindowsVista/en/library/ce4d5a2e-59a5-4742-89cc-ef9f5908b4731033.mspx?mfr=true

http://technet.microsoft.com/en-us/windowsvista/default.aspx

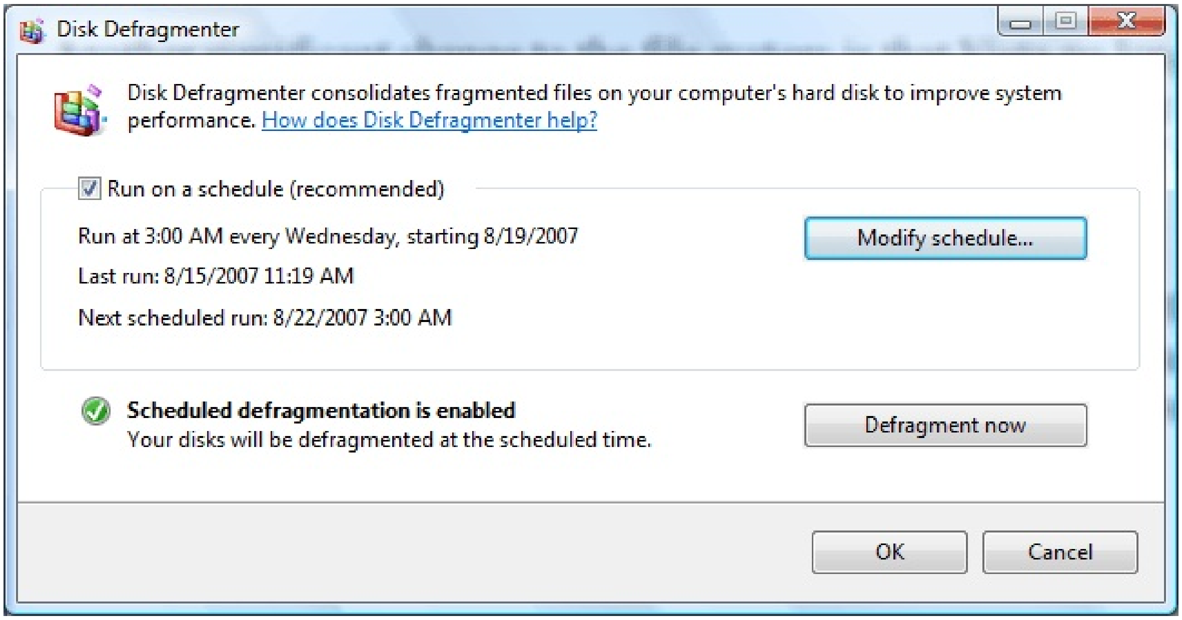
http://msdn2.microsoft.com/en-us/library/ms715237.aspx (Windows Mail)

A special thanks to GW for his contributions and technical support.❖



Name	Date modified	Type
Contacts	7/28/2007 8:46 PM	File Folder
Desktop	8/19/2007 2:02 PM	File Folder
Documents	8/19/2007 1:30 PM	File Folder
Downloads	8/19/2007 3:31 PM	File Folder
Favorites	8/18/2007 8:13 PM	File Folder
Links	7/28/2007 8:54 PM	File Folder
Music	8/7/2007 11:20 PM	File Folder
Pictures	8/18/2007 4:40 PM	File Folder
Saved Games	7/28/2007 8:47 PM	File Folder
Searches	7/28/2007 8:47 PM	File Folder
Videos	8/14/2007 10:57 AM	File Folder

FIGURE 1



Disk Defragmenter

Disk Defragmenter consolidates fragmented files on your computer's hard disk to improve system performance. How does Disk Defragmenter help?

☑ Run on a schedule (recommended)

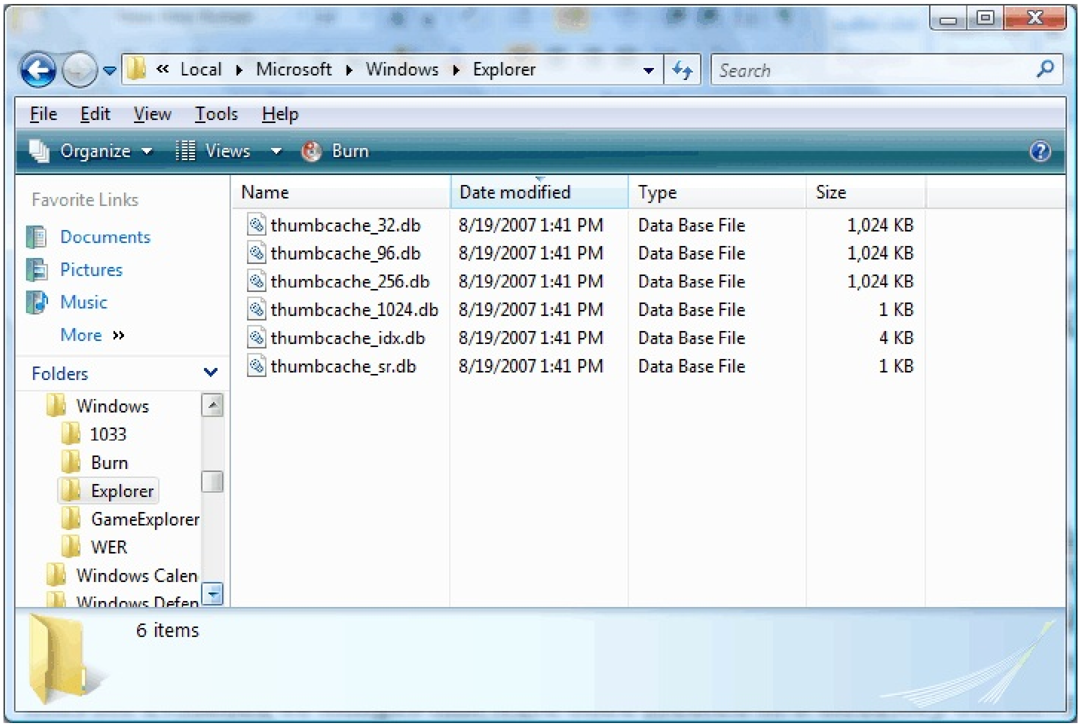
Run at 3:00 AM every Wednesday, starting 8/19/2007

Last run: 8/15/2007 11:19 AM

Next scheduled run: 8/22/2007 3:00 AM

Scheduled defragmentation is enabled
Your disks will be defragmented at the scheduled time.

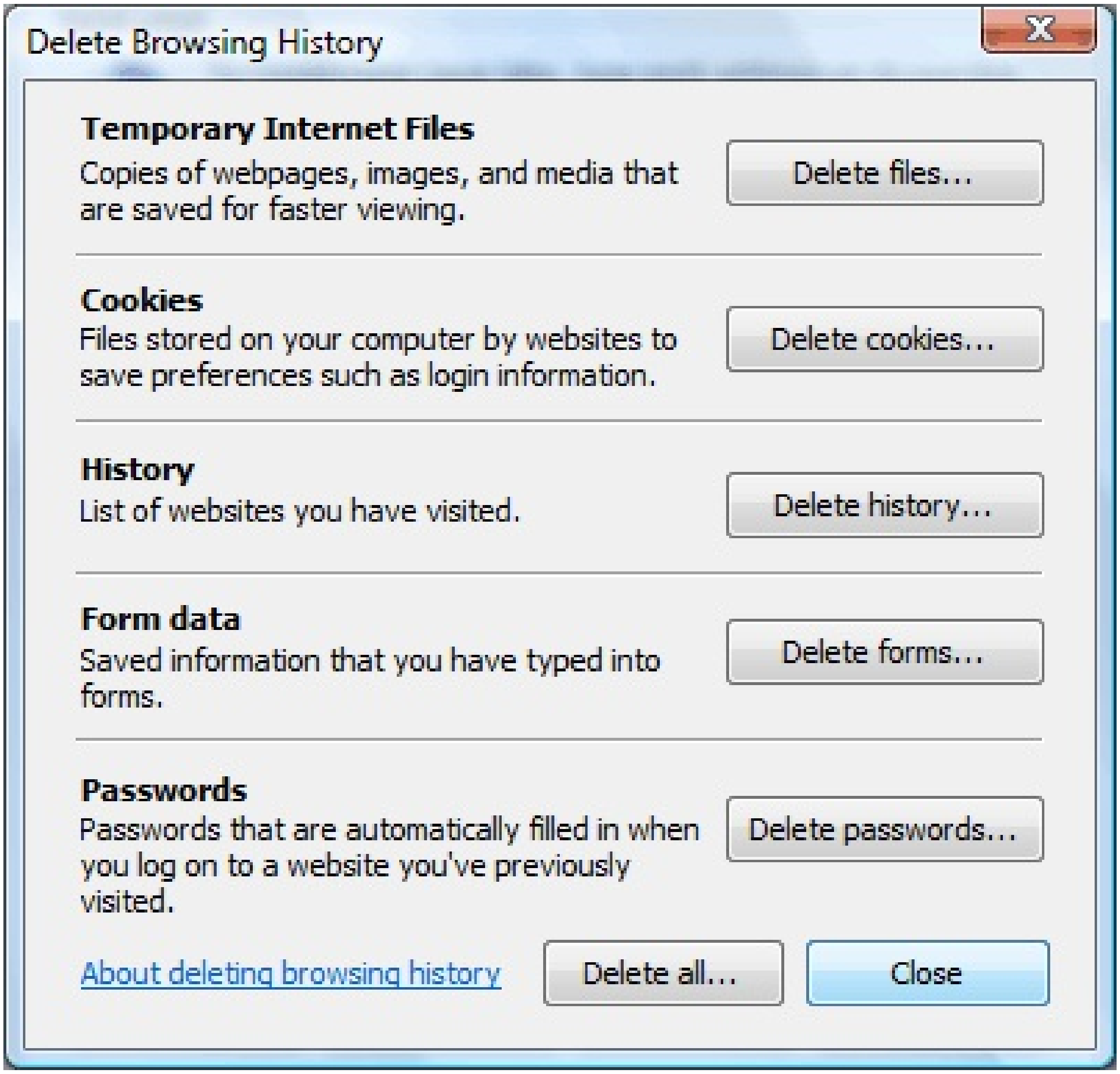
FIGURE 2



Name	Date modified	Type	Size
thumbcache_32.db	8/19/2007 1:41 PM	Data Base File	1,024 KB
thumbcache_96.db	8/19/2007 1:41 PM	Data Base File	1,024 KB
thumbcache_256.db	8/19/2007 1:41 PM	Data Base File	1,024 KB
thumbcache_1024.db	8/19/2007 1:41 PM	Data Base File	1 KB
thumbcache_idx.db	8/19/2007 1:41 PM	Data Base File	4 KB
thumbcache_sr.db	8/19/2007 1:41 PM	Data Base File	1 KB

6 items

FIGURE 3



Delete Browsing History

Temporary Internet Files

Copies of webpages, images, and media that are saved for faster viewing.

Cookies

Files stored on your computer by websites to save preferences such as login information.

History

List of websites you have visited.

Form data

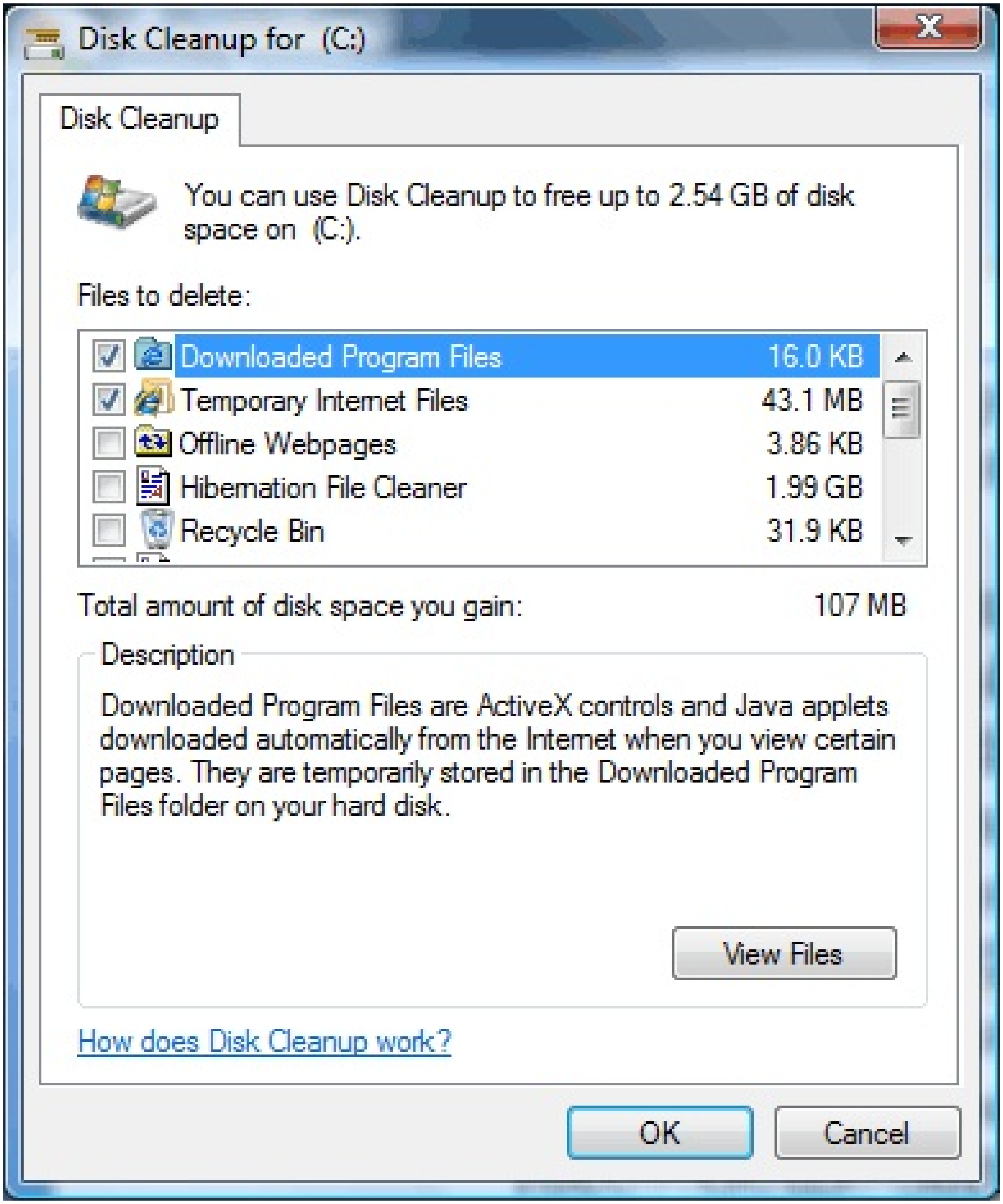
Saved information that you have typed into forms.

Passwords

Passwords that are automatically filled in when you log on to a website you've previously visited.

About deleting browsing history

FIGURE 4



Disk Cleanup for (C:)

Disk Cleanup

You can use Disk Cleanup to free up to 2.54 GB of disk space on (C:).

Files to delete:

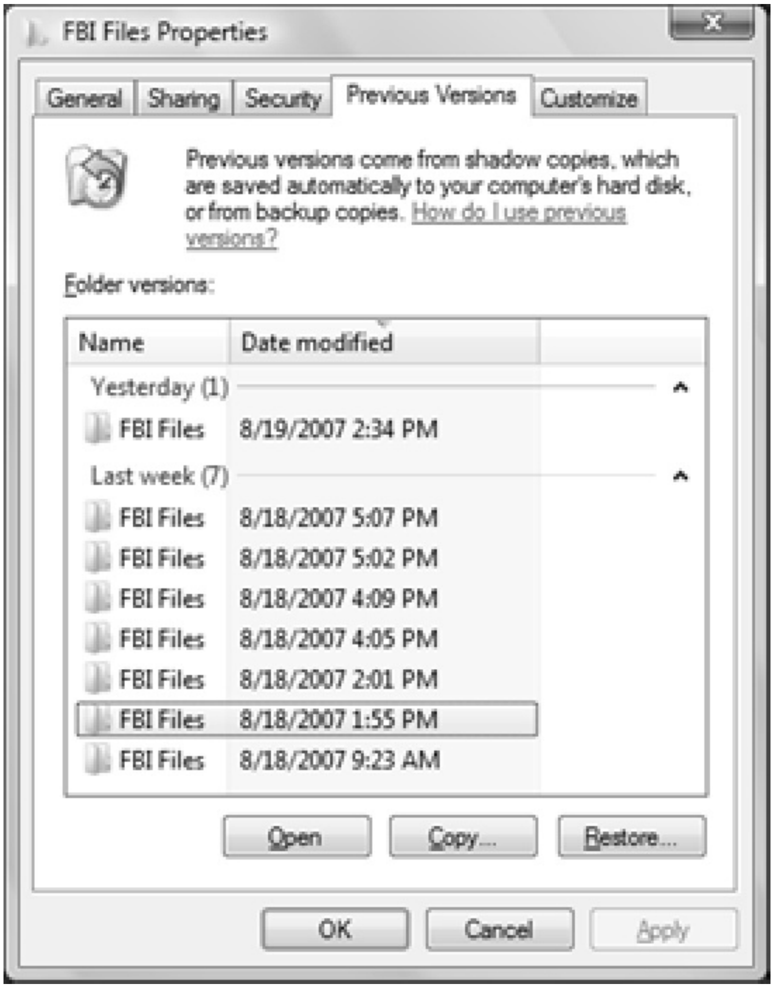
☑	Downloaded Program Files	16.0 KB
☑	Temporary Internet Files	43.1 MB
□	Offline Webpages	3.86 KB
□	Hibernation File Cleaner	1.99 GB
□	Recycle Bin	31.9 KB

Total amount of disk space you gain: 107 MB

Description

Downloaded Program Files are ActiveX controls and Java applets downloaded automatically from the Internet when you view certain pages. They are temporarily stored in the Downloaded Program Files folder on your hard disk.

FIGURE 5

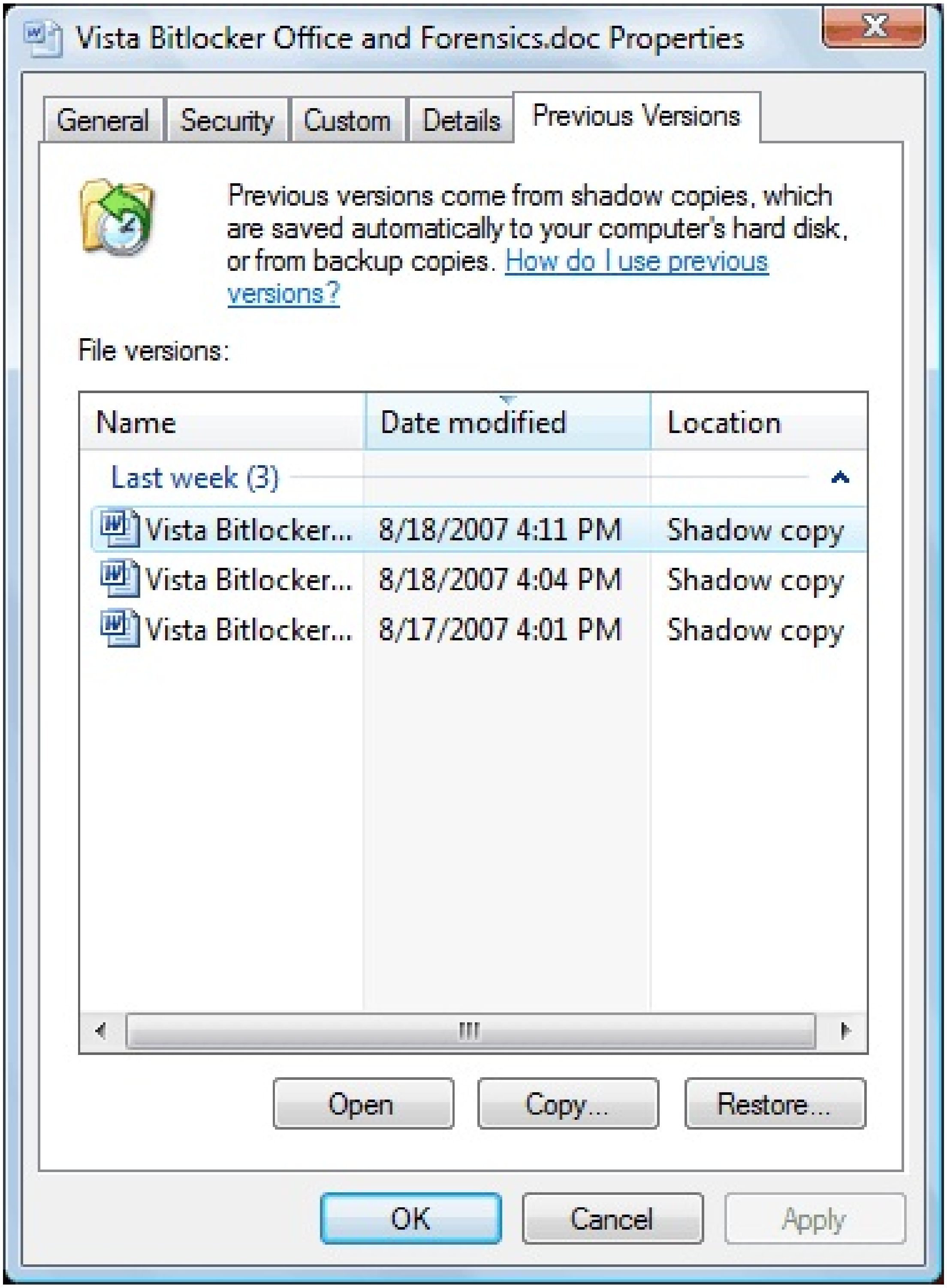


Previous versions come from shadow copies, which are saved automatically to your computer's hard disk, or from backup copies. How do I use previous versions?

Folder versions:

Name	Date modified
Yesterday (1)	
FBI Files	8/19/2007 2:34 PM
Last week (7)	
FBI Files	8/18/2007 5:07 PM
FBI Files	8/18/2007 5:02 PM
FBI Files	8/18/2007 4:09 PM
FBI Files	8/18/2007 4:05 PM
FBI Files	8/18/2007 2:01 PM
FBI Files	8/18/2007 1:55 PM
FBI Files	8/18/2007 9:23 AM

FIGURE 6

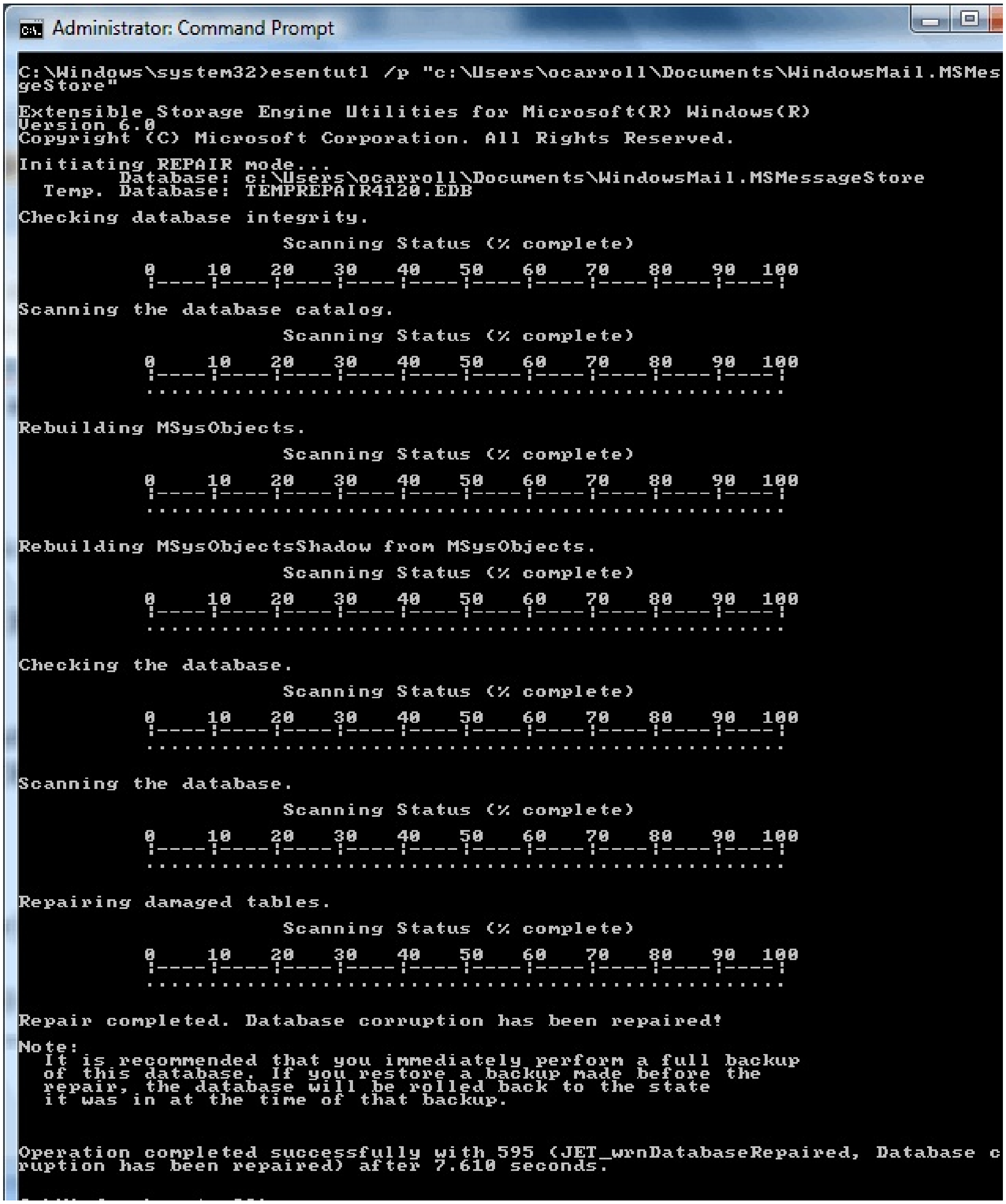


Previous versions come from shadow copies, which are saved automatically to your computer's hard disk, or from backup copies. How do I use previous versions?

File versions:

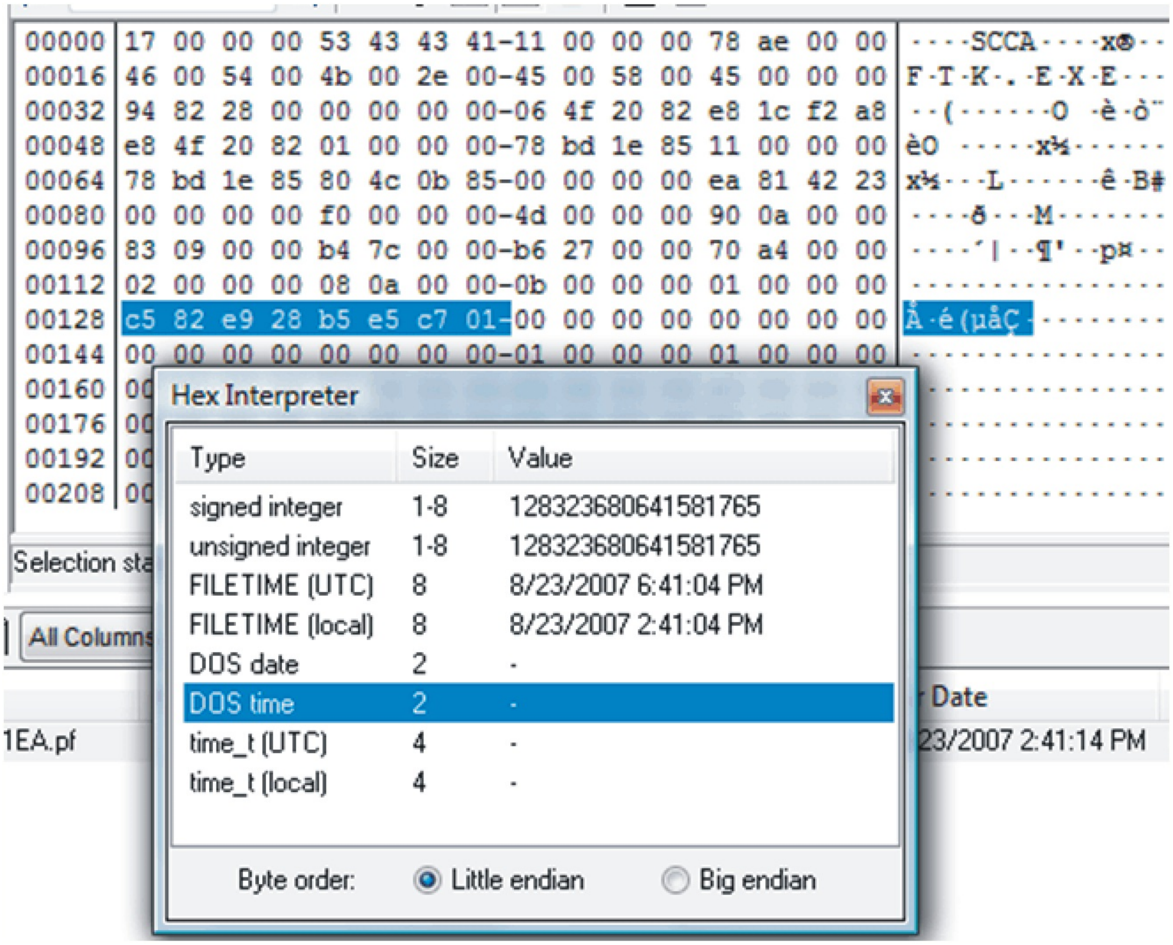
Name	Date modified	Location
Last week (3)		
Vista Bitlocker...	8/18/2007 4:11 PM	Shadow copy
Vista Bitlocker...	8/18/2007 4:04 PM	Shadow copy
Vista Bitlocker...	8/17/2007 4:01 PM	Shadow copy

FIGURE 7



```
C:\Windows\system32>esentutl /p "c:\Users\ocarroll\Documents\WindowsMail.MSMes
geStore"

Extensible Storage Engine Utilities for Microsoft(R) Windows(R)
Version 6.0
Copyright (C) Microsoft Corporation. All Rights Reserved.

Initiating REPAIR mode...
        Database: c:\Users\ocarroll\Documents\WindowsMail.MSMessageStore
  Temp. Database: TEMPREPAIR4120.EDB

Checking database integrity.

                     Scanning Status (% complete)

          0    10   20   30   40   50   60   70   80   90  100
          |----|----|----|----|----|----|----|----|----|----|

Scanning the database catalog.

                     Scanning Status (% complete)

          0    10   20   30   40   50   60   70   80   90  100
          |----|----|----|----|----|----|----|----|----|----|
          ...................................................


Rebuilding MSysObjects.

                     Scanning Status (% complete)

          0    10   20   30   40   50   60   70   80   90  100
          |----|----|----|----|----|----|----|----|----|----|
          ...................................................


Rebuilding MSysObjectsShadow from MSysObjects.

                     Scanning Status (% complete)

          0    10   20   30   40   50   60   70   80   90  100
          |----|----|----|----|----|----|----|----|----|----|
          ...................................................


Checking the database.

                     Scanning Status (% complete)

          0    10   20   30   40   50   60   70   80   90  100
          |----|----|----|----|----|----|----|----|----|----|
          ...................................................


Scanning the database.

                     Scanning Status (% complete)

          0    10   20   30   40   50   60   70   80   90  100
          |----|----|----|----|----|----|----|----|----|----|
          ...................................................


Repairing damaged tables.

                     Scanning Status (% complete)

          0    10   20   30   40   50   60   70   80   90  100
          |----|----|----|----|----|----|----|----|----|----|
          ...................................................


Repair completed. Database corruption has been repaired!

Note:
  It is recommended that you immediately perform a full backup
  of this database. If you restore a backup made before the
  repair, the database will be rolled back to the state
  it was in at the time of that backup.


Operation completed successfully with 595 (JET_wrnDatabaseRepaired, Database c
ruption has been repaired) after 7.610 seconds.
```

FIGURE 8



```
00000 17 00 00 00 53 43 43 41-11 00 00 00 78 ae 00 00  ····SCCA····x®··
00016 46 00 54 00 4b 00 2e 00-45 00 58 00 45 00 00 00  F·T·K·.·E·X·E···
00032 94 82 28 00 00 00 00 00-06 4f 20 82 e8 1c f2 a8  ··(·····O ·è·ò¨
00048 e8 4f 20 82 01 00 00 00-78 bd 1e 85 11 00 00 00  èO ·····x½······
00064 78 bd 1e 85 80 4c 0b 85-00 00 00 00 ea 81 42 23  x½···L······ê·B#
00080 00 00 00 00 f0 00 00 00-4d 00 00 00 90 0a 00 00  ····ð···M·······
00096 83 09 00 00 b4 7c 00 00-b6 27 00 00 70 a4 00 00  ····´|··¶'··p¤··
00112 02 00 00 00 08 0a 00 00-0b 00 00 00 01 00 00 00  ················
00128 c5 82 e9 28 b5 e5 c7 01-00 00 00 00 00 00 00 00  Å·é(µåÇ·········
00144 00 00 00 00 00 00 00 00-01 00 00 00 01 00 00 00  ················
00160 00
00176 00
00192 00
00208 00
```

Selection sta

All Columns

Date

1EA.pf

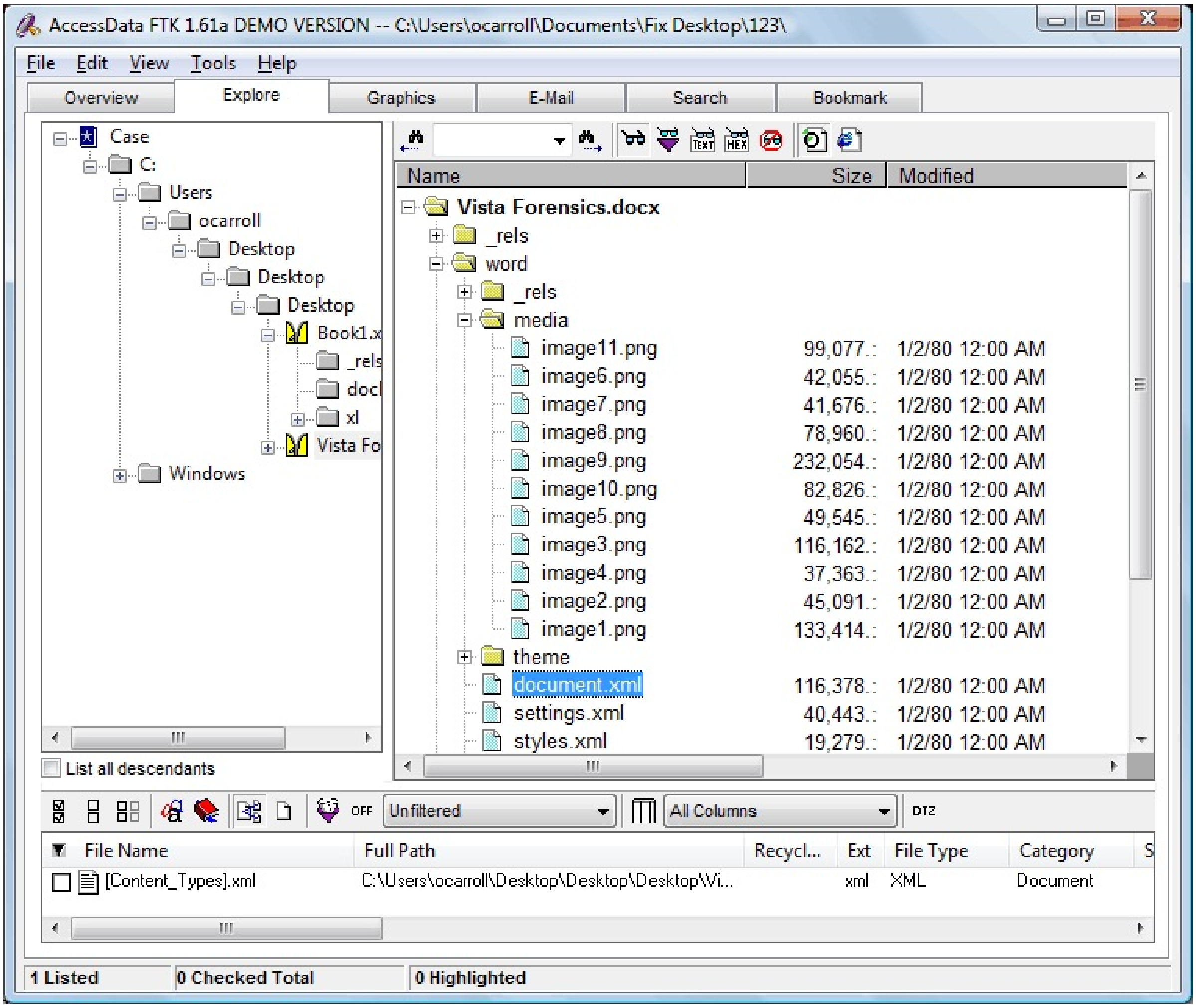
23/2007 2:41:14 PM

Hex Interpreter

Type	Size	Value
signed integer	1-8	128323680641581765
unsigned integer	1-8	128323680641581765
FILETIME (UTC)	8	8/23/2007 6:41:04 PM
FILETIME (local)	8	8/23/2007 2:41:04 PM
DOS date	2	-
DOS time	2	-
time_t (UTC)	4	-
time_t (local)	4	-

Byte order: ● Little endian ○ Big endian

FIGURE 9



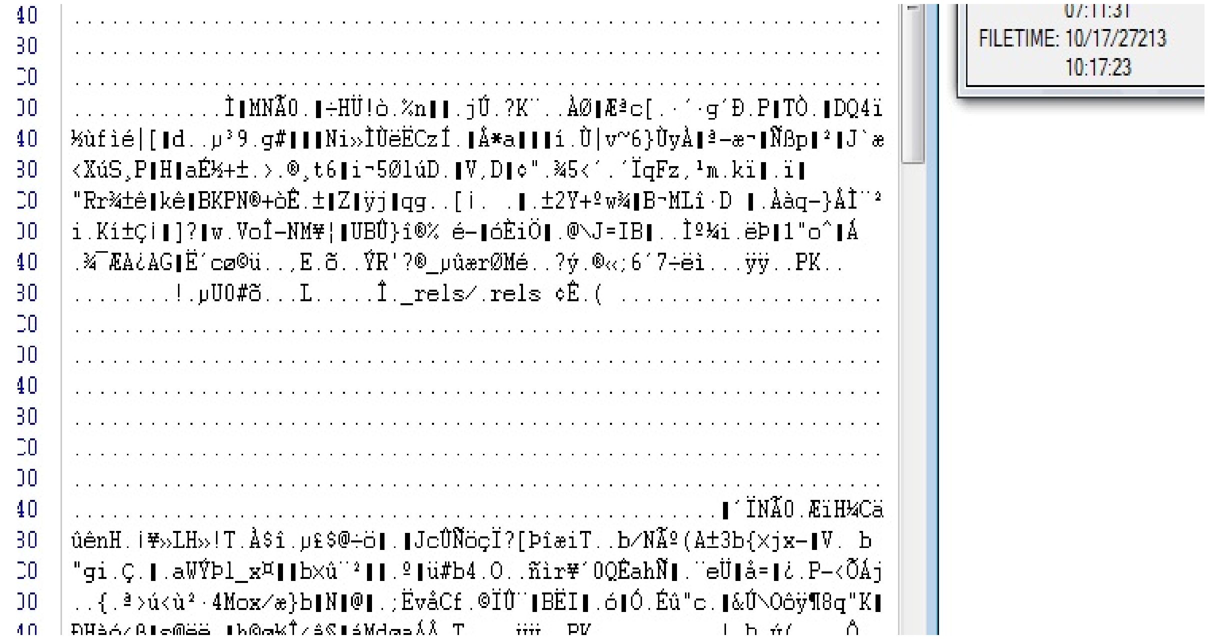
```
Case
  C:
    Users
      ocarroll
        Desktop
          Desktop
            Desktop
              Book1.x
                _rels
                docl
                xl
              Vista Fo
    Windows
```

Name	Size	Modified
Vista Forensics.docx		
_rels		
word		
_rels		
media		
image11.png	99,077.:	1/2/80 12:00 AM
image6.png	42,055.:	1/2/80 12:00 AM
image7.png	41,676.:	1/2/80 12:00 AM
image8.png	78,960.:	1/2/80 12:00 AM
image9.png	232,054.:	1/2/80 12:00 AM
image10.png	82,826.:	1/2/80 12:00 AM
image5.png	49,545.:	1/2/80 12:00 AM
image3.png	116,162.:	1/2/80 12:00 AM
image4.png	37,363.:	1/2/80 12:00 AM
image2.png	45,091.:	1/2/80 12:00 AM
image1.png	133,414.:	1/2/80 12:00 AM
theme		
document.xml	116,378.:	1/2/80 12:00 AM
settings.xml	40,443.:	1/2/80 12:00 AM
styles.xml	19,279.:	1/2/80 12:00 AM

File Name	Full Path	Recycl...	Ext	File Type	Category	S
□ [Content_Types].xml	C:\Users\ocarroll\Desktop\Desktop\Desktop\Vi...		xml	XML	Document	

1 Listed | 0 Checked Total | 0 Highlighted

FIGURE 10



```
................................................................
................................................................
................................................................
...........Î▮MNÃ0.▮÷HÜ!ò.%n▮▮.jÛ.?K¨..ÀØ▮Æªc[.·´·g´Ð.P▮TÒ.▮DQ4ï
½ùfié|[▮d..µ³9.g#▮▮▮Ni»ÌÜëËCzÍ.▮Å*a▮▮▮i.Ù|v~6}ÛyÀ▮ª-æ¬▮ÑßpI²▮J`æ
<XúS¸P▮H▮aÉ½+±.>.®¸t6▮i¬5ØlúD.▮V,D▮¢".¾5<´.´ÏqFz,¹m.kï▮.ï▮
"Rr¾±ê▮kê▮BKPN®+òÊ.±▮Z▮ÿj▮qg..[¡. .▮.±2Y+ºw¾▮B¬MLî·D ▮.Ààq-}ÅÌ¨²
i.Ki±Ç¡▮]?▮w.VoÎ-NM¥¦▮UBÛ}i®% é-▮óÈiÖ▮.@\J=IB▮..Ìº¼i.ëÞ▮1"o^▮Á
.¾¯ÆA¿AG▮Ë´cø©ü..,E.õ..ÝR'?®_µûærØMé..?ý.®«;6´7÷ëi...ÿÿ..PK..
........!.µU0#õ...L.....Î._rels/.rels ¢Ê.(.....................
................................................................
................................................................
................................................................
................................................................
................................................................
................................................................
................................................▮´ÏNÃ0.ÆïH¼Cä
ûênH.¡¥»LH»!T.À$î.µ£$@÷ö▮.▮JcÛÑöçÏ?[ÞîæiT..b/NÃº(A±3b{xjx-▮V. b
"gi.Ç.▮.aWÝÞl_x¤▮▮bxû¨²▮▮.º▮ü#b4.O..ñir¥´0QÊahÑ▮.¨eÜ▮å=▮¿.P-<ÕÁj
..{.ª>ú<ù²·4Mox/æ}b▮N▮@▮.;ËvåCf.©ÏÛ¨▮BËI▮.ó▮Ó.Éû"c.▮&Û\Oôÿ¶8q"K▮
```

07:11:31
FILETIME: 10/17/27213
10:17:23

FIGURE 11

ABOUT THE AUTHORS

❑**Ovie L. Carroll** is the Director of the Cybercrime Lab in the CCIPS. He has over twenty years of law enforcement experience. He previously served as the Special Agent in Charge of the Technical Crimes Unit at the Postal Inspector General's Office and as a special agent with the Air Force Office of Special Investigations.

❑**Stephen K. Brannon** is a Cybercrime Analyst in the CCIPS's Cybercrime Lab. He has worked at the Criminal Division in the Department of Justice and in information security at the FBI.

❑**Thomas Song** is a Senior Cybercrime Analyst in the CCIPS's Cybercrime Lab. He has over fifteen years in the computer crime and computer security profession. He specializes in computer forensics, computer intrusions, and computer security. He previously served as a Senior Computer Crime Investigator with the Technical Crimes Unit of the Postal Inspector General's Office.

The Cybercrime Lab is a group of technologists in the CCIPS in Washington, DC. The lab serves CCIPS attorneys, Computer Hacking and Intellectual Property (CHIP) units in the U.S. Attorneys' offices, and Assistant U.S. Attorneys, by providing technical and investigative consultations, assisting with computer forensic analysis, teaching, and conducting technical research in support of Department of Justice initiatives.✠

Demystifying the Computer Forensic Process for Trial: (Is My Witness Dr. Jekyll or Mr. Hyde?)

Martin J. Littlefield
Senior Litigation Counsel
United States Attorney's Office
Western District New York

I. Experts: Federal Rule of Criminal Procedure 16—Discovery

Rule 16 of the Federal Rules of Criminal Procedure allows a defendant to demand "notice" as to any "expert" the government intends to call at trial.

> [Government] Expert Witnesses. At the defendant's request, the government must give to the defendant a written summary of any testimony that the government intends to use under Rules 702, 703, or 705 of the Federal Rules of Evidence during its case-in-chief at trial.

FED. R. CRIM. P. 16(a)(1)(G).

Thus, the issue of whether a computer forensic examiner should be called as an "expert" might have to be addressed early in the case. This raises a more important question: If you choose to call your computer examiner an "expert," what is the examiner an expert in?

In *United States v. Scott-Emuakpor*, 2000 WL 288443, at *12 (W.D. Mich. Jan. 25, 2000), the trial court denied the defendant's pretrial motion to exclude the expert testimony of government witnesses (law enforcement agents who examined computer equipment and files seized from the defendant). After performing its "gate-keeping" function pursuant to *Daubert* and *Kumho Tire*, the Court ruled:

> [T]here is no reason why either witness may not testify about what they did in examining the computer equipment and the results of their examinations. The question before the Court at this time is not whether these witnesses have the expertise, for example, to develop sophisticated software programs. *The question is whether they have the skill to find out what is on a hard drive or a zip drive*. Apparently, they have this skill because they determined what was on the drives.

Id. (emphasis added).

Calling a computer examiner an expert brings into play a whole panoply of problems, while surely opening an avenue of cross-examination which might otherwise have been avoided.

In the average case, where a letter or graphic is found in a folder on a defendant's computer ("c:\mydocuments\dirtypix\"), it is not necessary for an expert to say it was there. The experience of the average juror generally will include knowing how a file (not a deleted file or a temporary file) sits in a folder. On the other hand, there are cases where seemingly "hidden/unknown" logs, deleted files, or unsaved messages, can be retrieved by the examiner. This would be beyond the computer experience of the average juror.

II. What is an expert? What is a fact witness?

> If scientific, technical, or *other specialized knowledge* will assist the trier of fact to understand the evidence or to determine a fact

in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto *in the form of an opinion or otherwise*, if (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case.

FED. R. EVID. 702 (emphasis added).

If the computer examiner is testifying about locating and/or extracting files, how does one describe his area of "specialized knowledge?" Does the examiner really know—or even have to know—the intricacies of a particular program loaded on the suspect's computer? Does the examiner have to display an in-depth knowledge of the logarithm (MD-5 (Message Digest) hash value)) used to verify that an exact bit-by-bit copy of a drive was successfully executed? As to the latter, he or she might only be able to say that the logarithm for the MD-5 hash value is widely relied upon by examiners and has been rigorously tested to assure that it is completely reliable.

In many cases, the examiner's testimony demonstrates substantial training and "specialized knowledge" in the application of programs and/or tools used to locate files or data such as, EnCase or Forensic Toolkit (FTK). The examiner, however, may not be able to fully explain how or why the program and/or tool works. In this case, the qualifying testimony is focused on how the program and/or tool is widely used, accepted, and relied upon, throughout the world, as a tool which can locate and/or extract from the target drive specified data (*see* Federal Rule of Evidence 702(2), above). Therefore, the response to the FED. R. CR. P. 16 "expert" demand must be carefully worded and the direct testimony carefully circumscribed to limit the cross-examination, or at least to have the judge focused on what is the "true" specialized knowledge of the examiner.

III. Types of testimony—examples

A. The file was found in a directory

Consider whether the computer examiner is really expressing an opinion or merely stating a fact—the file was located on the hard drive in the path *c:\mydocuments\dirtypix*. The examiner's training in the use of programs to more easily locate certain types of files (EnCase to find .jpg files) does not mean that his or her testimony has to invoke that training, or even refer to programs used to more quickly locate certain types of nondeleted files. The fact is that the file was located in a given location.

At trial, the examiner would state that on the computer at *c:\mydocuments\dirtypix*, there is the file called "*lolitta2*." This is not to say that the witness might never have to explain the use of EnCase or that a report listing multiple jpg files (sought to be introduced) is a product of EnCase. The point is that in a given case, if a single file is the evidence relevant to the trial and it resided in a particular folder, going into EnCase's capabilities is a waste of time and might only serve the defense as an opportunity to confuse or divert the evidentiary importance of the file.

B. The file was first saved, last modified, created

If the testimony involves critical metadata, then the examiner's testimony takes on a much more significant role. His or her specialized knowledge might include an understanding and explanation of an operating system's (Windows) logs and/or how a particular program (WordPerfect) maintains information about particular files.

C. The remnant data was recovered from virtual memory

The examiner must be able to explain what virtual memory is, its nature, and how it stores information for a limited purpose and for a limited time (depending on the suspect's usage). The specialized knowledge deals with explaining the types of memory on the computer and how data can be recovered, even if not saved.

IV. Federal Rule of Criminal Procedure 16 response: limiting the examiner's expertise

The purpose of the preceding example is to emphasize that the Federal Rule of Criminal Procedure 16 response must be based on the prosecutor's understanding of what was found and where it was found. To simply say that the government will be calling an expert on computers may give the examiner too much credit. Federal Rule of Criminal Procedure 16(a)(1)(G) also states that the government has to provide even more detail about the expert testimony. "The summary provided under this subparagraph must describe the witness's opinions, the bases and reasons for those opinions, and the witness's qualifications." FED. R. CR. P. 16(a)(1)(G).

Note that the rule refers to the opinions of the witness. More likely than not, the examiner will not be asked an opinion, but merely present the findings from his examination of the computer, that the file "*lolitta2*" was located in *c:\mydocuments\dirtypix*. Thus, as noted below in the sample Federal Rule of Criminal Procedure 16 response, the examiner might be more of a fact witness than what is traditionally thought of as an expert (in my opinion the car was traveling more than 85 mph).

As a result, Federal Rule of Criminal Procedure 16 will force prosecutors to learn more about the examiner and his or her actual work and experience at a relatively early stage of the prosecution. Take the time to do it before penning the response. Even so, a generic and not overly-broad notice can be provided and still meet the requirements of the rule.

V. Sample response to a Federal Rule of Criminal Procedure 16 demand

The government acknowledges receipt of a Notice of Motion and Motion in the above-captioned matter, filed on behalf of the defendant, by his attorney. By this pleading, the government is providing its response to the aforesaid motion.

****

> *Experts*
>
> It is anticipated the government will call a witness who examined the defendant's computer. The government believes that the testimony will involve his technical and specialized knowledge regarding the examination of computers under Federal Rule of Evidence 702. The witness will be asked to explain his training and background (resumé attached), the nature of the examination that he undertook of the computers and related media, and the methods and/or software used to assist him in the examination. Thus, the witnesses will be more in the nature of a fact witness explaining how certain evidence was located. The government is prepared to present the evidence pursuant to Federal Rule of Evidence 702 and the requirements of the Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993) and *Kumho Tire Company, Ltd. v. Carmichael*, 526 U.S. 137 (1999). The government, however, emphasizes that the witness will be offered only for his technical and specialized knowledge in the area of computer examination, not, for example, for expertise in the field of hardware or software development. *See United States v. Scott-Emuakpor*, 2000 WL 288443, *12, (W. D. Mich. Jan. 25, 2000) ("The question before the Court at this time is not whether these witnesses have the expertise, for example, to develop sophisticated software programs. The question is whether they have the skill to find out what is on a hard drive or a zip drive. Apparently, they have this skill because they determined what was on the drives.")
>
> Already provided to the defense is a report and related files and data recovered from the defendant's computer which the government believes adequately summarizes the nature of

the examiner's testimony as required under Federal Rule of Criminal Procedure 16. The defendant should review the aforementioned materials and advise as to whether any further information is needed to adequately prepare for trial.

VI. Observation, issues, and concerns for evidence derived from electronic media

A. Copying the original hard drive

- Did the examiner properly make a bit-by-bit copy?
- How does the examiner know that the copy was exact (MD-5 Hash Value Logarithm)?
- What steps are required to make such a copy?
- Did the examiner follow these steps?

Failure to adequately show that the copy was an "exact duplicate" of the original media would be a major point of attack for the admission of data found on the computer. Clearly an argument could be made that failure to establish that the copy is an exact duplicate would go to the weight of the evidence, however, it may also be the basis for a challenge to the admissibility of the evidence.

B. Locating the data/evidence

Was the data in an easily locatable folder or was it in "unallocated space"? (The answer will impact the degree of "specialized knowledge" the examiner will have to articulate.)

C. Logs, Metadata

Were there logs regarding the data? Are there conflicts in those logs? (Modify vs. creation dates; the time clock is inaccurate in the computer regarding the send/receive data, among other things).

D. Reliability of evidence found

Sometimes the best evidence that the examiner properly located or extracted from the data is the very data produced (a fragment of instant message contains sensible, understandable, and logically progressive communications between the parties).

E. Qualifying the examiner

Utilize the usual set of "qualifying" information for the examiner.

- Prior times as a witness
- Training
- Certifications
- Number of examinations performed
- Peer review of work and accuracy
- Ability to articulate what is usually done in an examination (and, of course, what was actually done)
- Review of literature and currency in evolving technology
- Activity as a trainer to others

Paragraph X is an excerpt from a trial transcript for qualifying an examiner who found instant message fragments and various data from nonfolder areas of a hard drive.

VII. Case law and useful excerpts from the 2002 Advisory Committee Notes on Federal Rule of Evidence 702

A. *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993)

Daubert factors:

- Whether a theory or technique ... can be (and has been) tested
- Whether it has been subjected to peer review and publication
- Whether, in respect to a particular technique, there is a high known or potential rate of error and whether there are standards controlling the technique's operation

- Whether the technique or theory enjoys general acceptance within a relevant scientific community

B. *Kumho Tire Company, Ltd. v. Carmichael*, 526 U.S. 137 (1999)

- Extends *Daubert* gate-keeping responsibility to all expert testimony, not just scientific testimony
- Court may consider *Daubert* factors, when doing so will help determine the testimony's reliability
- *Daubert* test is flexible, the factors do not apply to every expert in every case
- Admissibility based on the relevance and reliability of the testimony

C. *McDowell v. Brown*, 392 F.3d 1283, 1299 (11th Cir. 2004)

> A district court has "considerable leeway in deciding in a particular case how to go about determining whether particular expert testimony is reliable," and we give that discretion a large degree of deference. *Kumho Tire*, 526 U.S. at 152, 119 S.Ct. 1167. The Supreme Court did not intend, however, that the gatekeeper role "supplant the adversary system or the role of the jury: '[v]igorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof are the traditional and appropriate means of attacking shaky but admissible evidence.'" *Allison*, 184 F.3d at 1311-12 (quoting *Daubert*, 509 U.S. at 596, 113 S.Ct. 2786). The judge's role is to see that the jury hears reliable and relevant evidence because of its ability to assist in factual determinations, its potential to clarify issues, and its probative value.

Id.

VIII. Federal Rule of Evidence 702 Advisory Notes

Federal Rule of Evidence 702 has been amended in response to *Daubert v. Merrell Dow Pharmaceuticals, Inc.* and the many cases applying *Daubert*, including *Kumho Tire Co. v. Carmichael.* In *Daubert* the Court charged trial judges with the responsibility of acting as gatekeepers to exclude unreliable expert testimony, and the Court in *Kumho* clarified that this gatekeeper function applies to all expert testimony, not just testimony based in science. Consistent with *Kumho*, Federal Rule of Evidence 702, as amended, provides that all types of expert testimony present questions of admissibility for the trial court in deciding whether the evidence is reliable and helpful. Consequently, the admissibility of all expert testimony is governed by the principles of Federal Rule of Evidence 104(a). Under that Rule, the proponent has the burden of establishing that the pertinent admissibility requirements are met by a preponderance of the evidence. *See Bourjaily v. United States*, 483 U.S. 171 (1987).

A. *Daubert* set forth a nonexclusive checklist; court must use its experience and judgment

Daubert set forth a nonexclusive checklist for trial courts to use in assessing the reliability of scientific expert testimony. The specific factors explicated by the *Daubert* Court follow:

- Whether the expert's technique or theory can be, or has been, tested—that is, whether the expert's theory can be challenged in some objective sense, or whether it is simply a subjective, conclusory approach that cannot reasonably be assessed for reliability
- Whether the technique or theory has been subject to peer review and publication
- The known or potential rate of error of the technique or theory when applied
- The existence and maintenance of standards and controls

- Whether the technique or theory has been generally accepted in the scientific community

The Court in *Kumho* held that these factors might also be applicable in assessing the reliability of nonscientific expert testimony, depending upon "the particular circumstances of the particular case at issue." 526 U.S. at 150.

No attempt has been made to codify these specific factors. *Daubert* emphasized that the factors were neither exclusive nor dispositive. Other cases have recognized that not all of the specific *Daubert* factors can apply to every type of expert testimony. In addition to *Kumho*, 526 U.S. at 150, *see Tyus v. Urban Search Management*, 102 F.3d 256 (7th Cir. 1996) (noting that the factors mentioned by the Court in *Daubert* do not neatly apply to expert testimony from a sociologist). *See also Kannankeril v. Terminix Int'l, Inc.*, 128 F.3d 802, 809 (3d Cir. 1997) (holding that lack of peer review or publication was not dispositive where the expert's opinion was supported by "widely accepted scientific knowledge."). The standards set forth in the amendment are broad enough to require consideration of any or all of the specific *Daubert* factors, where appropriate.

B. No single factor is dispositive for the reliability of a particular expert's testimony

All of the factors mentioned in the previous section remain relevant to the determination of the reliability of expert testimony under the Federal Rule of Evidence 702, as amended. Other factors may also be relevant. *See Kumho*, 526 U.S. 137, 152 (1999) ("[W]e conclude that the trial judge must have considerable leeway in deciding in a particular case how to go about determining whether particular expert testimony is reliable.") Yet no single factor is necessarily dispositive of the reliability of a particular expert's testimony. *See, e.g., Heller v. Shaw Industries, Inc.*, 167 F.3d 146, 155 (3d Cir. 1999) ("not only must each stage of the expert's testimony be reliable, but each stage must be evaluated practically and flexibly without bright-line exclusionary (or inclusionary) rules.").

C. Expert testimony generally is admitted; cross-examination and contrary testimony are best counters

A review of the case law after *Daubert* shows that the rejection of expert testimony is the exception rather than the rule. *Daubert* did not work a "sea change over federal evidence law," and "the trial court's role as gatekeeper is not intended to serve as a replacement for the adversary system." *United States v. 14.38 Acres of Land Situated in Leflore County, Mississippi*, 80 F.3d 1074, 1078 (5th Cir. 1996). As the Court in *Daubert* stated: "Vigorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof are the traditional and appropriate means of attacking shaky but admissible evidence." 509 U.S. at 595.

Likewise, this amendment is not intended to provide an excuse for an automatic challenge to the testimony of every expert. *See Kumho Tire Co. v . Carmichael*, 526 U.S. 137, 152 (1999) (noting that the trial judge has the discretion "both to avoid unnecessary 'reliability' proceedings in ordinary cases where the reliability of an expert's methods is properly taken for granted, and to require appropriate proceedings in the less usual or more complex cases where cause for questioning the expert's reliability arises.").

When a trial court, applying this amendment, rules that an expert's testimony is reliable, this does not necessarily mean that contradictory expert testimony is unreliable. The amendment is broad enough to permit testimony that is the product of competing principles or methods in the same field of expertise. *See, e.g., Heller v. Shaw Industries, Inc.*, 167 F.3d 146, 160 (3d Cir. 1999) (expert testimony cannot be excluded simply because the expert uses one test rather than another, when both tests are accepted in the field and both reach reliable results). As the court stated in *In re Paoli R.R. Yard PCB Litigation*, 35 F.3d 717, 744 (3d Cir. 1994), proponents "do not have to demonstrate to the judge by a preponderance of the evidence that the assessments of their experts are correct, they only have to demonstrate by a preponderance of evidence that their opinions are reliable. . . . The evidentiary requirement of reliability is lower than the merits standard of correctness." (Footnote omitted). *See also Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 43 F.3d 1311, 1319 (9th Cir. 1995) (scientific experts might be permitted to testify if they could show that the methods they used were also employed by "a recognized minority of scientists in their field."); *Ruiz-Troche v. Pepsi Cola*, 161 F.3d 77, 85 (1st Cir. 1998) ("*Daubert* neither requires nor empowers trial courts to determine which of several competing scientific theories has the best provenance.")

D. Federal Rule of Evidence 702 allows for technical or other specialized knowledge—not just science

As stated earlier, the amendment does not distinguish between scientific and other forms of expert testimony. The trial court's gatekeeping function applies to testimony by any expert. An opinion from an expert who is not a scientist should receive the same degree of scrutiny for reliability as an opinion from an expert who purports to be a scientist. Some types of expert testimony will not rely on anything like a scientific method, and so will have to be evaluated by reference to other standard principles attendant to the particular area of expertise. The trial judge, in all cases of proffered expert testimony, must find that it is properly grounded, well-reasoned, and not speculative, before it can be admitted. The expert's testimony must be grounded in an accepted body of learning or experience in the expert's field, and the expert must explain how the conclusion is so grounded.

E. Federal Rule of Evidence 702 allows flexibility for procedure to be followed by court as gatekeeper

The amendment makes no attempt to set forth procedural requirements for exercising the trial court's gatekeeping function over expert testimony. *See* Daniel J. Capra, *The Daubert Puzzle*, 38 GA. L. REV. 699, 766 (1998) ("Trial courts should be allowed substantial discretion in dealing with *Daubert* questions; any attempt to codify procedures will likely give rise to unnecessary changes in practice and create difficult questions for appellate review.")

F. "Expert" is a convenient label but should not necessarily be used with a jury

> The amendment continues the practice of the original Rule in referring to a qualified witness as an "expert." This was done to provide continuity and to minimize change. The use of the term "expert" in the Rule does not, however, mean that a jury should actually be informed that a qualified witness is testifying as an "expert." Indeed, there is much to be said for a practice that prohibits the use of the term "expert" by both the parties and the court at trial. Such a practice "ensures that trial courts do not inadvertently put their stamp of authority" on a witness's opinion, and protects against the jury's being "overwhelmed by the so-called 'experts.'"

FED. R. EVID. 702.

IX. Some questions to be considered when applying Federal Rule of Evidence 702

A. What is the witness going to do?

Is the witness going to express an opinion or merely explain a process and the result of that process so that no "opinion" is given (for example, an explanation of how Bankruptcy Court works)?

B. What is the witness an expert in, or what does he or she have specialized knowledge about?

Be careful of overstating what the witness will be testifying to and what area he or she is qualified in. For example, a computer forensic examiner is not necessarily a computer expert, but rather a person trained to use certain utilities (programs) to extract and/or find certain data or files. At trial, the defense should be limited by the witness's area of specialized knowledge during cross-examination. *See United States v. Scott-Emuakpor*, 2000 WL 288443, at *12 (W.D. Mich. Jan. 25, 2000). "The question before the Court at this time is not whether these witnesses have the expertise, for example, to develop sophisticated software programs. The question is whether they have the skill to find out what is on a hard drive or a zip drive." *Id.*

C. Other witness considerations

Did the witness have all the data or information needed to make an informed, reliable judgment? Did the witness have a reasonable and/or reliable principle (premise), and did he or she use reasonable and/or reliable methodology? Did the witness apply the principles and methods reliably to the facts of the case?

D. Is the witness both a fact witness and an expert?

This type of mixed testimony may generally be presented, but there must be a clear division between the testimony as to the facts he or she saw or heard, and the testimony for which he or she is interpreting or opining as an expert. For the latter, Federal Rule of Evidence 702 and *Daubert* apply just as strenuously, and the prosecutor should be careful to make a clear delineation between a fact and an expert witness and assure that the subject matter is appropriate for expert testimony.

X. Transcript excerpt for qualifying an examiner who found instant message fragments and various data from "non-folder" areas in a hard drive

Testimony of Computer

(Electronic Media)

Examiner/Analyst

John Shumway, RCFL/WDNY

(The name of the child/victim has been changed)

MR. LITTLEFIELD: Stipulation of evidence re: computers. Martin Littlefield and Marie Grisanti, Assistant United States Attorneys, on behalf of the Government, and Kimberly Schechter, attorney on behalf of the defendant, Mark Friedman, hereby enter into the following stipulation dealing with the chain of custody related to two computers.

Individual No. 1's computer, Exhibit 71, and Friedman's computer, Exhibit 80. It is agreed that subject to a showing of relevancy, that in lieu of submitting documentary evidence and testimony, the statements set out below shall be admitted as evidence at trial, and that the chain of custody for both computers is correct as set out below.

Chain of custody, individual No. 1's computer, Exhibit 71. The Government and the defendant agree that individual No. 1's, parens, Mary Doe, close parens, computer, Exhibit 71, was received by the FBI on January 14th, 2002, and that until its presentation at trial, the computer was never out of the control and custody of law enforcement personnel.

Next paragraph header, chain of custody,

Friedman's computer, Exhibit 80. The Government and the defendant further agree that Mark Friedman's computer, Exhibit 80, was seized by the FBI pursuant to a search warrant at the defendant's residence on January 16th, 2002. And that until its presentation at trial, the computer was never out of the control and custody of law enforcement personnel. The above stipulation has been reviewed and agreed to by the defendant, his attorney, and the attorneys for the Government, and is hereby entered into and agreed to as set forth above.

May I proceed with the witness?

THE COURT: Yes.

DIRECT EXAMINATION BY MR. LITTLEFIELD:

Q. Sir, how are you employed?

A. I'm employed as a City of Niagara Falls police officer, currently on assignment at the regional computer forensics lab of Western New York.

Q. I know you're going to have to talk a little slower and more into those microphones, if you would.

And how long have you been a Niagara Falls police officer?

A. Nineteen years and four months.

Q. And how long have you been assigned or had the assignment at the regional computer forensic laboratory of Western New York?

A. Approximately two and a half years.

Q. And, sir, relative to your duties and responsibilities at the lab—What are the duties and responsibilities you have at the lab?

A. My duties and responsibilities involve receiving evidence from law enforcement officials throughout Western New York, making bit stream duplication copies of the storage data media contained therein, and doing a forensic analysis of that.

Q. And just for clarification, when you say making bit stream copies, what do you mean by that?

A. Every single bit, consisting of the hard drive or floppy drive or whatever digital media is brought in, is duplicated, as it—and you get an exact copy of its original.

Q. In other words, one to one?

A. Yes.

Q. And now you said every single bit, that's a phrase that we use all the time. Is bit not a term of art in computerese?

A. Yes, it is.

Q. And what does a bit refer to?

A. A bit refers to one-eighth of a byte. This is going to get confusing. We've all seen computers referred to as possessing megabytes or gigabytes. Those are all just expanded and larger versions of the initial bit. There are eight bits in one byte and so on.

Q. Okay. And, sir, relative to your duties and responsibilities, have you had any training to qualify you as an expert?

A. Yes.

MR. LITTLEFIELD: Not as an expert, I'm sorry, Judge, I withdraw that. I didn't mean that.

Q. Qualify as an examiner.

A. Yes, I have.

Q. And could you tell the jury what that was?

A. To start, I'm A-plus certified repair technician, I am a Microsoft certified professional in four different aspects.

Q. Slow down a little bit.

A. I am a certified Novell administrator. I have had seven certificate training classes on Windows 2000. I am a certified forensic computer examiner, achieved after eighty hours of training at IACIS, which is the International Association of Computer Investigative Services Specialists, I'm sorry. I have had four training sessions at the National White Collar Crime Center, those being basic data recovery and analysis, and three forms of advanced data recovery and analysis, those being Windows 95 and 98 and Me, Internet trace evidence, and Windows NT and XP.

Q. Now, sir, you mentioned IACIS. Could you spell out what those letters are and then tell us what the words are behind those letters?

A. IACIS is I-A-C-I-S, and it represents International Association of Computer Investigative Specialists.

Q. As a forensic examiner, is there a method or methodology by which a person becomes certified to conduct forensic examinations of computers?

A. After achieving the training, you have an extensive one-year period to complete seven exercises and a written exam to become a certified forensic examiner.

Q. Certified by whom though?

A. Certified by IACIS, which is an international company. Or group/organization.

Q. And have you—did you undertake that, sir?

A. Yes, I did.

Q. And have you been so certified?

A. Yes, I have.

Q. Now, sir, in addition to this training that you have, and your two and a half years at the Regional Computer Forensic Lab, have you undertaken any analysis or examination of computers?

A. I've completed approximately a hundred and fifty examinations of computers.

Q. And in that course of time have you had occasion to testify in court?

A. Yes, I have.

Q. About your examination?

A. Yes, I have.

Q. Now, sir, initially you said something about this bit by bit and then an examination. Is there—are they two different, or are they the same thing?

A. They are two different processes.

Q. Explain to the jury as best you can in laymen's terms what exactly the first category is.

A. The first category being the bit stream duplication of the hard drive in this case involves using recognized software, Safe Back and EnCase specifically, that makes a bit stream image to—not image, I'm sorry—bit stream duplicate copy to sterile lab media.

Q. What's a sterile lab media?

A. Well, that would be another hard drive that has been forensically wiped. That is, every bit on the hard drive has been zeroed out so there's absolutely nothing on there.

Once that is done, the duplicate copy can then be analyzed.

Q. Why don't you work on the original computer itself?

A. You would change the data by doing that. You never want to alter the data.

Q. And is that—in the testimony you're about to give, is that what you did in the case for the two computers we're about interest to discuss?

A. Yes, it is.

Q. The second category, could you explain to the jury what duties and responsibilities that involves?

A. That—the analysis end of it involves using certified software, in our case it is EnCase, among other softwares. Used to look at the data contained on the duplicate copy, and extract information, that being data that if the computer were running, you would see, as well as information and data that you wouldn't see, but still exists on the computer.

Q. Now, EnCase, could you explain a little bit to the jury about that program and where it comes from?

A. It comes from Guidance Software, it is what's termed an automated forensic utility. It used to be you did it all manually, where you would look over each bit and copy out portions thereof to do it. This software is built on that technology, and does it automatically for you.

Q. And how do you go about conducting an examination using EnCase?

A. We start by loading the duplicate copy, and make sure that the acquisition hash has verified.

Q. Acquisition hash?

A. Yes.

Q. What's that?

A. When you acquire the copy, it generates what's called an MD five hash value. An MD five hash value is a 32 bit alphanumeric signature generated by a hash program, to verify that the data is as it purports to be. When you load the image into the software, it runs a verify against it, thereby coming up with an identical MD-5 hash value, verifying that your copy is exactly what the original was.

Q. Relative—have you looked at Government's Exhibit 71, 71-X, which purports to be a copy of a hard drive, and Exhibit 80, prior to your testimony today?

A. Yes, I have.

Q. And have you conducted—A, did you make images of them, or a bit-by-bit stream duplication of their hard drives?

A. Yes.

Q. And did you ultimately do an examination of them?

A. Yes.

Q. Are you familiar with a tool, a forensic—by the way, when we say forensic tool, do you know what forensic means?

A. As in the normal meaning of forensics that you hear from watching the television shows, it recovers information, clues, or non relevant data in some cases, left over from computer use.

Q. Okay. Are you familiar with a forensic tool known as FTK?

A. Yes, that is made by a company called Access Data, and it stands for Forensic Toolkit.

Q. And did you use that in the course of this investigation?

A. Only for a comparison, I used the demo version, as we don't have a licensed version at the lab.

Q. Why don't you?

A. Purchasing troubles.

Q. Okay. As between FTK and EnCase, is one better or worse than the other?

A. No, there is no perfect tool. No tool does everything. Some tools work better on one aspect, some tools work better than on others.

Q. Are there versions that you have of any given program?

A. Version 4.15 now of EnCase.

Q. And how many versions before that were there?

A. Numerous.

Q. Okay.

A. I started on version 1.99, two and a half years ago.

Q. Okay. And what does the higher number indicate?

A. That they have made changes to the program to make its functionality better.

Q. Does the tools that you use, whether it's an EnCase program or FTK, is it designed—no. Can it in any way add data or information in the course of using that tool?

A. No, it cannot.

Q. What is its purpose or what does it go about doing then?

A. It simply looks at the physical level of the computer media, in this case hard drives, and allows you to see it, where normally on a running computer you would not.

Q. And if there was something you're looking for, does either or any of these tools always guarantee you're going to be able to find everything that might be of use in the course of the investigation?

A. No software can guarantee all the time everything will be there.

Q. And why not?

A. Things can get overwritten. Certain aspects of computers, data is only in virtual memory instead of actually physically saved to the hard drive.

Q. What's virtual memory; we've heard that term, I think, but why don't you explain it.

A. There are two or three—three, actually, kinds of memory on a computer. There is physical memory as in ram, you've also seen commercials for upgrade your ram and make your computer faster. That's a physical chip about six inches long that goes inside your computer. There is hard drive storage. The actual hard drive is memory for the data you've put in there. And then there is virtual memory. Virtual memory is used by—that created and used by the operating system to make access to certain files faster. It uses free space, and makes it act like physical ram, thereby bringing data in, bringing data out far faster than physical ram can.

Q. Now, sir, when you conduct—use any of these tools, but in particular EnCase, does it—how do you go about looking for things, or how—I mean, there's all this data, you said.

A. Files have specific header information to them, and we can search for those headers. There's also key word searches that we can do to look for words given to us by the investigators to see if they're located on the computer.

Q. And now you said files. What's the information that's contained within a file called?

A. Data.

Q. Okay. And so what is—when I use Word Perfect and I type a letter and save it, I save a file called—if I wrote a letter to Judge Arcara I'd write Judge Arcara letter.WPD. Is that a file?

A. That would be a file.

Q. What is a file defined as, as far as the computer is concerned?

A. A file is anything loaded onto the computer, be it by the user or by the operating system, that contains data.

Q. Could a program also be a file?

A. Yes.

Q. Now, inside the computer it's filled with data then, is that correct?

A. Inside the hard drive, yes.

Q. Is all the data that's inside a hard drive, broken down into files so that you have so much data in this file, so much data in this file, et cetera, et cetera, so that it all adds up to the sum total of data in the hard drive?

A. No.

Q. Why not?

A. There's unallocated space, that is, space that's not visibly being used by the computer, that is not a file, it is just referred to as unallocated clusters. There is virtual memory, which is not a file, but it contains data. There's ram, which holds data for a limited period of time also.

Q. Okay. Now, if a person were to delete a file on their computer, let's say it's a Word Perfect file, my letter that I just referred to, I deleted it. Is it permanently deleted from that computer?

A. No, it is not.

Q. Where does it go?

A. When you first delete a file, assuming you were using the Windows operating system, the first thing that happens is that file, the pointers to that file within the computer's own data base, get changed, and it gets placed in what's called the recycled bin or recycled or trash, can be referred to.

Q. By the way, do both of these computers use the Windows operating system?

A. Yes, they do.

Q. And now they're in trash or recycle; what does that mean?

A. Windows has that incorporated into it in case you accidentally delete a file, it allows you to go into the recycle bin and restore it to its original location. So all it is, is changing the pointer as to where that file is. The actual file, the data in the file is still in its original location.

Q. Suppose—now, what happens, if that data is still out there for that given file, even though it's a deleted file, does it ever go away or is it always sitting there?

A. While it's in the recycle bin, it does not go away, unless your recycle bin has reached its maximum size set by the operating system.

Q. Then what happens?

A. Then it would become overwritten.

Q. What do you mean by that?

A. Another file would use that space.

Q. Could it be possible to get back parts of a file that have been overwritten?

A. Yes, file artifacts can be recovered.

Q. All the time?

A. Not all the time.

Q. What are the prohibitions or what would prohibit you from being able to do that?

A. When a portion of a file or a complete file has been overwritten by another file, it no longer is recoverable.

Q. And how does a person—so my letter to Judge Arcara, I wanted to overwrite that letter with the letter to Judge Skretny. Can I tell my computer I want to you overwrite Judge Arcara's letter and use all that space for my letter to Judge Skretny?

A. Only if you have special utilities that will overwrite a specific area of a hard drive.

Q. Do most computers have that?

A. No.

Q. Then if I write a letter to Judge Skretny after deleting Judge Arcara's letter, does it overwrite other stuff, or what happens?

A. It would be placed on the hard drive in an area that the operating system has deemed is not being used by visible files. And it would go there.

Q. Now, are deleted files, areas that the computer considers to be not being used?

A. Once you have emptied your trash or your recycle bin, the computer deems that area, that data portion, to be usable. And, when necessary, will overwrite that area.

Q. But until it's overwritten, is it still there?

A. Yes, it is.

Q. Okay. Now, you said there are times that things aren't in files, but the data still exists there?

A. That's correct.

Q. Could you explain that a little bit more to the jury?

A. The point of being deleted, and the recycle bin has been emptied, that data area still existed, and if it has not been overwritten, those files can still be recovered. That's one. Physical ram retains specific information—or not specific, but random information—that is placed in it by the operating system and the computer itself. That can sometimes be recovered. And the virtual memory, the swap file or the page file, also holds data that is not intentionally saved on a computer, but is still there.

Q. And when does that—well, can you give us an example of something that would be in virtual memory, and how it would be—how and why it might still be there, and how and when it might not be there?

A. That file you were talking about, you were in the middle of writing it and you changed your mind and closed the program, that would still be in virtual memory. Internet—

Q. Stop right there. You mean if I wrote a letter but I never gave it a name a file name, just stopped writing it?

A. You simply closed the program and chose not to save your changes.

Q. The things that I typed though, would that be somewhere?

A. That would still be on the computer, yes.

Q. Where?

A. That actually would be in a couple of places. You would find it in virtual memory and you would probably find it in what's called a temporary file that retained that data while you had that program open.

Q. Even though I never tried to save it or put my name on it or anything of that nature?

A. That's correct.

Q. Now, you were about to explain there was another kind?

A. Yeah, while you're on the Internet, if you are, for instance, using instant messaging, that will come up in—actually be stored and shown to you in virtual memory until you either choose to save that, or you go on to the next chat conversation and it becomes overwritten.

Virtual memory overwrites itself rapidly, simply because it uses the available space so that it can work faster. It works on a first in, first out basis for data.

Q. With the two operating systems that you have here, did you look and determine whether and how they had a capacity or no capacity for the saving of instant messages?

A. Yes, I did.

Q. And what did you determine?

A. By default, when you load America Online or whatever your chat program is, instant message—in this case it was America Online.

Q. For both?

A. Yes. You have to physically set it to retain your chat logs.

Q. And the review of the computers, did you determine whether that had been specifically set to retain them?

A. It was not set to retain them.

Q. And would it be possible, therefore, for any instant messages that took place and that would—for which these computers might have been used, for you to recover them?

A. Yes.

Q. Any of the instant messages?

A. Yes.

Q. How so?

A. If they were in virtual memory when that computer was shut down, they would still be there and be fully recoverable, to the point that they existed on there. The whole chat or message might not be there, but portions or fragments of it would be.

Q. And how long will that—again, the stuff that's in virtual memory, if somebody sent an instant message from A to B, how long would you, or can you tell us how long that would remain available for recovery or portions of it for recovery?

A. From virtual memory, that would be dependent on the amount of use of the computer. If you get through with a chat session, you immediately turn it off, it would be relatively easy to recover, as long as you didn't turn the computer back on and use it. Continued use of the computer will overwrite virtual memory.

MR. LITTLEFIELD: Your Honor, at this time Government's Exhibit 71, which is the computer, 71- X, which was the hard drive of that—that was part of it, and Exhibit 80—71 is the Mary Doe's computer, 80 is Friedman's computer. We'll move those into evidence at this time.

MS. SCHECHTER: No objection.

THE COURT: All right, they'll be received. (Government Exhibits 71 and 80 received.)

BY MR. LITTLEFIELD:

Q. Now, sir, I want to direct your attention to Government's Exhibit 80, which is the computer that was seized at the residence of Mark Friedman on January 16th. Did you have a chance to conduct a forensic examination of that?

A. Yes, I did.

Q. And did you follow the procedure regarding making a bit stream duplication of it?

A. Yes, that procedure was followed.

Q. And did you work on a copy of it?

A. Yes, I did.

Q. And, sir, in that regard, did you conduct a forensic examination of it?

A. Yes, I did.

Q. And did you receive information from the investigators as to help you focus on that?

A. Yes, I received some key word terms and some more specifics as to what to look for.

Q. And when you say key word terms and some more specifics, explain that to the jury.

A. Names, specific names, user names.

Q. Like what?

A. Hot NJ guy, love hot girls, spells with two Rs and no I. Mary Doe, spelled both M-A-R-I and R-Y.

Q. Is there anything else you looked for in there?

A. I was instructed to look for electronic graphic images.

Q. Now, the one thing is key words. When you say key words, are you looking actually inside text, you know, text like my letter, if I said one of those words?

A. The search does not just search text, it searches text fragments anywhere on the computer.

Q. And these images thing is a separate search?

A. Yes, it is.

Q. What does that look for?

A. Looks for electronic pictures either in bit map format or JIF format or JPEG format.

Q. How does the computer know whether to look for those things?

A. It uses the file header information. When you're looking at say a JPEG image, you see the file extension JPG or JPEG. When I look at it, I look at it in what's called hexadecimal view, and I see the header information, that being hex characters FF D8FF. That's the first three character sets.

Q. Okay. Now, sir, relative to Mr. Friedman's computer, Exhibit 80, were you able to find anything relative to your instructions, that were of evidentiary value?

A. I recovered some key word hits and some stored directional maps, I believe.

Q. Now, sir, relative to that, do you have Exhibit 80.01, .02 and .03 before you?

A. Yes, I do.

Q. Let's start off with exhibit 80.01. Could you describe in general what that purports to be?

A. Starting from the top of the page, this is a print from Exhibit 80, which is the computer of Friedman.

Q. Stop right there. Will—the explanatory language that leads into it, we'll discuss later. But I mean what is the thing?

A. Specifically this is a copy—the portion of an instant message session between hot NJ guy and Mary Doe Mary Doe 27. Mary Doe with a Y in this case, that I recovered from the page file .sys, which is the virtual memory of the operating system on the Friedman computer, which was Windows XP.

Q. Now, Exhibit 80.01 is a single page. How did it come to be printed, if you will?

A. After I got the key word hit on the Mary Doe Mary Doe and hot NJ guy, I highlighted it and was able to extract it from the duplicate hard drive.

Q. Now, is it not evidence from this, that this is only a fragment of a conversation?

A. It is only a fragment, yes.

Q. Where's the rest of the conversation?

A. There again, it was in the virtual memory; it is easily overwritten.

Q. Do you know where the rest of the conversation is?

A. No, I don't.

Q. Did you try to recover it?

A. I did, and I was unable to.

Q. Now, on Exhibit 80.01, on this page there's the top portion of it that has two lines drawn across.

MR. LITTLEFIELD: In fact, if you want, Miss Steblein, for the Court's—blank the jury—and counsel, would you bring up 80.01, please.

Q. There's two lines about a third of the way down from the page. Does that divide something there, and if so, what is it—if the reader's looking at it, what is it designed to divide?

A. It is designed to divide the actual recovered portion of evidence.

Q. Which is on what portion?

A. Which is in the lower portion.

Q. Top—

A. Upper portion gives me the details about that recovered data.

Q. And is the top information, stuff that you typed, that was typed in above that, for purposes of identifying this item?

A. Yes.

Q. And is the information below taken verbatim or exactly from the computer, using your forensic tool?

A. Yes, it is.

Q. And why don't you, relative to this top information, explain to us what information is that designed to give to the reader?

A. At the top it gives me an exhibit number, it's a downloading print of Exhibit 80, that tells me where it came from, the Friedman computer. Then it tells me it's a printout from the Friedman computer, specifically cluster 2657909, which is an area on the hard drive. That it was in a file called the page file .sys, S-Y-S, which begins at cluster 23 2642603. Specifically the page file was in the C drive backslash page file .sys.

Q. Now, all of this is information that if you went into Mr. Friedman's computer to that cluster, you'd find this? Is that pretty much it?

A. If you had forensic tools to go to that specific cluster, yes, you would.

Q. Now, below that there's another line though that says downloaded to. What is that designed to convey to the reader?

A. That explains that on any RCFL evidence disk—

Q. What is that?

A. That's a CD-ROM that I created at the lab after my examination of the evidence that I located, or what I believed to be evidence that I located.

Q. Okay.

A. And this states that I gave this file name text hit Mary Doe dash page file .sys.txt, meaning it's a text file at this point.

Q. Kind of goes back to what we talked before. Is this data that you recovered—when you found it, was it actually in an electronic file?

A. No, it was not.

Q. And yet you're telling us that you put in a file on a disk?

A. I copied it out to a separate file, a text file, so that it was viewable.

Q. Did you provide copies to counsel?

A. Yes.

Q. The disk, I mean.

A. Yes.

Q. And is that what this third line refers to where it says downloaded to?

A. Yes.

Q. The RCFL disk?

A. That's correct.

MR. LITTLEFIELD: At this time, Your Honor, I'd move into Exhibit 80.01.

MS. SCHECHTER: Can I voir dire, Your Honor?

THE COURT: Yes.

VOIR DIRE

BY MS. SCHECHTER:

Q. Is it Officer Shumway?

A. Yes.

Q. Hi. What you've identified as Government Exhibit 80.01, this is not a complete text message instant message, is it?

A. That's correct.

Q. At the top of bottom portion underneath the double lines that counsel was referring to?

A. Yes.

Q. You see some—some—

A. Gibberish.

Q. —characters.

A. Yes.

Q. And that indicates that portion of the file was corrupted, does it not?

A. That indicates that being the virtual memory, other data was stored in that area.

Q. Other data was stored in that area that cut off—

A. That either overwrote part of the instant message, or—that would basically be it, it overwrote part of that message.

Q. And if we go down to the bottom portion of the email above the solid line on the bottom, if we go up, say five lines, there's a plus sign there, do you see that?

A. Yes.

Q. Does that also indicate that part of that file was overwritten in some manner?

A. That would indicate that that is a possibility, yes.

Q. So there's additional conversation that would have been ongoing in that instant message that took place at the beginning of Government's Exhibit 80.01, at the beginning of the first indication of Mary Doe Mary Doe?

THE COURT: All right. This is a good time to break until tomorrow morning. ❖

ABOUT THE AUTHOR

❑**Martin J. Littlefield** is an Assistant United States Attorney and also serves as the Senior Litigation Counsel for the Western District of New York. He is an expert in computer and computer related fraud investigations and prosecutions. Mr. Littlefield serves as instructor for the Office of Legal Education at the National Advocacy Center.✠

Managing Large Amounts of Electronic Evidence

Ovie L. Carroll
Director, Cybercrime Lab
Computer Crime and Intellectual Property Section
Criminal Division

Stephen K. Brannon
Cybercrime Analyst
Cybercrime Lab
Computer Crime and Intellectual Property Section
Criminal Division

Thomas Song
Senior Cybercrime Analyst
Cybercrime Lab
Computer Crime and Intellectual Property Section
Criminal Division

I. Introduction

Investigations usually focus on finding and preserving evidence. A computer-related investigation often generates a particularly large amount of evidence. Managing all this data and using it effectively through the life cycle of an investigation presents special problems. This article explores those problems, and describes general strategies and some specific solutions for managing large amounts of electronic evidence.

II. Concepts and concerns

A. Preliminary concerns

The one cardinal rule for electronic evidence is: always work on a copy. Original evidence, or the single best copy, must be duplicated, kept safe, and a clean chain of custody record maintained. Only use working copies of evidence for review and analysis.

Working directly with original evidence or the best copy is extremely dangerous for the following reasons:

- Interacting with the original or best copy will likely change it.
- There is a greater risk that data will become corrupted or lost due to hard disk failure.
- The integrity of electronic evidence is also important because if it is intact, forensic copies should theoretically be exact.

With some other types of forensic evidence, testing and analysis use up the evidence itself. But with electronic evidence, any number of exact copies can be made, and the defense is often entitled to receive a copy for review.

Pandora's Box is opened if the government cannot produce an exact copy of the evidence that is seized or obtained. If evidence is accidentally modified, and the modification is clearly documented and explained, then the evidence can probably be used. However, the modification may affect the weight of the evidence.

Electronic evidence is much easier to manage if a system of organization is in place before it is collected. As investigators retrieve evidence, it is documented and the original versions are preserved. As copies of the evidence are made, however, it is helpful to have a system of organization and file the copies as they are made. Far too often, investigators let the order in which they retrieve evidence, or its sources, dictate the organization. At the end of the investigation, if the evidence needs to be reorganized, there may not be time.

The idea is to think forward to the analysis in order to guide the initial setup. For example, an investigation of multiple targets using multiple Web sites is started. If most of the questions will be about one target or another, then organize the evidence by target, as it is received. On the other hand, if a coherent picture of the activity on each Web site needs to be shown, organize the evidence by Web site. Planning and setting up an organizational system at the beginning of an investigation may determine whether or not electronic evidence is manageable at the end.

Another preliminary concern for electronic evidence is the use of date and time information. Problems with computer date and time settings can be fatal to an investigation.

- Targets can be misidentified.
- Evidence can show a target did something he or she did not do.
- Evidence from different computers can be inconsistent.

Every computer and server has an internal clock. The date and time—or at least what the computer believes they are—are spread through all the evidence the computer produces, especially any logs. The clock setting may be wrong, or it may be set to a different time zone. Investigators must find and adjust both for inaccuracies of the clock and for discrepancies between different clocks.

Fortunately, it is possible to document and compensate for almost any problem with dates and time, as long as it is both identified and quantified. As an example, an investigator is running an undercover Web site and logging the activity that takes place. He or she discovers that the clock on the computer running the Web site is twenty-three minutes fast, and has been for the past year. With both of these pieces of information, the log evidence can be salvaged and used by subtracting twenty-three minutes from every activity.

Let the computer do the work when dealing with electronic evidence. A human brute-force attack is too slow, and it also introduces the potential for many errors. The most conscientious person cannot avoid making mistakes when he or she is required to repeat an action 1,000 times. On the other hand, when a computer is given accurate instructions, it can easily execute them a million times, error free. A long list of answers without mistakes is often what is needed from electronic evidence.

To use electronic evidence in an investigation, the evidence must be organized and managed so that it is searchable. When searching the electronic evidence, the investigator may often feel like he or she is trying to find a needle in a haystack. For example, millions of lines of logs may have to be searched to find the one record that shows a target used the Web site for illegal activity.

The electronic evidence must be analyzed and interpreted (looking at all the evidence, or a large part of it, and deriving new information). For example, an investigation has many targets using many Web sites. It might be necessary to identify the most egregious user(s) to select targets for prosecution. Each target's conduct, across all the Web sites, must be determined throughout the course of the investigation and then compared.

B. Indexing

Indexing is a technique used to search large bodies of data more quickly. Indexing goes through the entire body of data and creates a map of the location of all information. This map, or index, functions like the index in a book or the card catalog in a library. Building an index can take a long time, however once it is done, searches are accomplished much faster. It is hard to imagine how long it would take to search every word on the Internet, or every word in the Lexis Nexis databases, if they were not indexed. If there is a large amount of data and multiple searches are necessary, it is generally best to index once, and then use that index to search. In the long run, this is faster and more efficient.

Indexed searching can be done within computer forensics programs. It can also be done using stand-alone programs that only index and search data. The computer forensics program, Forensic Toolkit (FTK), can index built-in data and is considered the leader in indexed searching. EnCase, Version 6 (the newest version) also incorporates indexing capabilities. Indexing can be done in previous versions of EnCase by using a third-party add-on, such as Mercury. Once data in a forensics program has been indexed, searches that would have taken minutes or hours are completed almost instantly.

There are also stand-alone programs that just do indexed searching. dtSearch produces a mature suite of programs that use the same indexing engine as FTK. The basic program searches text in multiple formats and highlights results. It also has options to use fuzzy, phonic, wildcard, stemming, and thesaurus search-options (search techniques that finds results similar to, or related to, the term provided). It can find misspelled words, so it is especially useful when searching through anything written by a person. For example, a fuzzy search for "apple" would also find "appple." dtSearch can also display results as Web pages. Another program in the product suite, dtSearch Publish, allows the investigator to publish and distribute evidence in an indexed and searchable package to distribute for review.

C. Visualization

Some results are only useful to a prosecutor or jury when they are presented visually. There are programs available that combine database and visualization features which enable an analyst to find and illustrate connections. These tools are often used in cases with extensive financial data or phone records. They are also particularly useful to show relationships indicated by e-mail exchanges or network traffic. One of the most popular programs is Analyst's Notebook. (http://www.i2.co.uk/Products/Analysts_Notebook/default.asp). It can illustrate relationships as shown in Figure 1, page 53.

Analyst's notebook can also perform and illustrate time line analysis. See Figure 2, page 54.

III. Techniques and tools

A. E-mail

E-mail evidence must be reviewed on a computer that is not connected to the Internet and is dedicated to off-line-evidence review. If e-mail is reviewed on an off-line computer, it may keep track of read receipts. If it is later connected to the Internet, it will send all of the read receipts.

If e-mail is reviewed on a computer connected to the Internet, a read receipt response may be accidentally sent to the addressees on the e-mail, as well as the target. An agent, in a recent case, accidentally double-clicked the "reply-all" button while reviewing an e-mail on his office computer. By doing this, he created and sent an e-mail to all the conspirators, which jeopardized the investigation. Also, some e-mail uses hypertext mark-up language (HTML), the language used in preparing Web pages, to control formatting. Outlook creates messages in this format by default. HTML can have references to images and other files on Web sites, and opening it can cause the computer to connect to those Web sites to retrieve message elements. This can directly or indirectly warn a tech-savvy target that he or she is under investigation.

E-mail provides several searchable sources of valuable information.

- Content
- Elements of the header (sender, recipient, subject, or date sent)
- Number of attachments
- Attachment names, priority, or age

The optimum method of organization and management of e-mail in an investigation is to import it into one e-mail program. Some investigators, however, will review each e-mail, or import groups of e-mail, into different e-mail programs. This technique makes managing and searching the evidence more difficult. It is easier to structure and organize e-mail folders when they are saved into one program. This organizational method allows the investigator to search all e-mail or individual folders.

There are many free and commercial e-mail programs available. Mozilla Thunderbird is one of the best free programs for managing and searching e-mail for most cases. The program is available at http://www.mozilla.com/en-US/thunderbird/. Other free or commercially installed programs include Outlook, Outlook Express, and Eudora. It is best if specialized forensics programs, such as Access's Data Forensic Toolkit and Paraben's E-mail Examiner, are used in larger cases.

In order to operate, most programs need the user to create a profile (name, e-mail address, among others). When a user opens a program for the first time, instructions for completing the account creation process are given. It is fine to use fabricated information for this since the computer is not connected to the Internet.

The most common formats of e-mail can be imported into Mozilla Thunderbird. One common format is the .mbox format (.mbx or .mbox). If a file with e-mail has no file extension, it is likely an .mbox file. The simplest way to import e-mails into Thunderbird in the .mbox format is to copy the file to the directory where Thunderbird stores its files. The next time the program starts, the .mbox file and all its e-mail appears as a folder under "Local Folders." In Windows XP, the directory is C:\Documents and Settings\[UserName]\ApplicationData\ Thunderbird\ Profiles\xxxxxxxx.default\Mail\ Local Folders\ (xxxxxxxx is eight random characters). In Windows Vista, the directory is C: \users\[UserName]\AppData\Roaming\ Thunderbird\Profiles\xxxxxxxx.default\Mail\ Local Folders\ (xxxxxxxx is eight random characters). Copy e-mail evidence files to that directory, then open the program and it is ready to use.

Another e-mail format is the Microsoft .pst (Personal Folders) format. Thunderbird cannot import a .pst file directly, but the file can be imported into Microsoft Outlook and then into Thunderbird. Importing a .pst file into Outlook only takes a few steps. The goal is to get e-mail from a .pst file into the Personal Folders in the Outlook profile. These instructions are specifically for Outlook 2003.

- In Outlook, click the File menu
- Click Data File Management
- Click the Add button
- Click OK
- Find and select the .pst file
- Type in a new name in the "Name" box (for example, "warrant response")
- Click OK
- Click Close

The .pst file should appear as a folder (warrant response) on the bottom of the left pane. The last thing to do in Outlook is move the mail from the new folder to a folder inside the Personal Folders.

- Right-click on Personal Folders
- Select New Folder
- Name the folder (for example, "Bad Guy1")
- Click OK
- Click on the .pst folder at the bottom ("warrant response")
- Select all the e-mails by clicking the Edit menu, then selecting Select All
- Carefully click on any e-mail and drag it into the folder created in the Personal Folders ("Bad Guy1."). This should move all e-mail from the .pst file into the local folder

Close Outlook and open Thunderbird to import the e-mail from Outlook. The following instructions are for Thunderbird 2.0.0.6:

- In Thunderbird, click the Tools menu
- Click Import
- Select Mail

- Click Next
- Select Outlook
- Click Next
- The process imports every folder from Outlook. Note: It may be helpful to delete empty or unrelated folders.

There are two main advantages of bringing all e-mail evidence into one program. The first is the ability to organize and manage the e-mail in a way that works best for the case. The second is that all e-mail evidence is searchable simultaneously, or individual sections may be searched.

To open Thunderbird's search interface, follow these instructions:

- Click Edit
- Find
- Search Messages

The following screen shot depicts the search interface. See Figure 3, page 55.

The box at the top selects which folder or folders to search. Select "Local Folders" and leave the "Search Subfolders" box checked to search in all e-mail. The two radio buttons and the middle pane specify search conditions. The radio buttons determine whether *all* the conditions must be met for a result to be included, or if it will be included when *any* of the conditions are met. Each search condition specifies where in the e-mail to look, the condition to meet (contains, does not contain, begins with), and the search term. The investigator can easily add or remove any number of conditions by clicking the + or – buttons.

Click the Search button and search results are displayed in the bottom pane. The list can be sorted by any field. A search hit is easy to file into a folder. For example, create a folder named "key e-mails." When an e-mail is selected in the search results list, click the File button on the bottom and select the folder in which to move it. A useful search may also be saved. Click the Save as Search Folder button and it will create a search results folder. The results will be viewable as if they were a folder, but the original e-mail will not be moved.

B. Chat logs

Many computer-related investigations involve records of online chat, or instant messaging. Instant messaging lets two or more people have a real-time, text-based conversation, over the Internet. Each user types messages on his or her computer, and every user who is party to the conversation sees all the typed messages in real time. For example, in a one-on-one chat, two people send text messages back and forth. Both people see the conversation scroll by in a window. In a chat room, or channel, with several people communicating, each person sees a window representing the room and everything everybody types is visible. Most instant messaging programs allow users to log their chats and some programs even log by default. Chat logs may be obtained from a target's computer, from a victim, or recorded using a cooperator or undercover agent.

The program mIRC is the most common program used for Internet Relay Chat (IRC). IRC is a type of chat popular in many tech-savvy crime circles, such as hacking, identity theft, and high-level copyright infringement. The mIRC program has an option to log its communications. The person using the program only needs to check the right boxes and the program produces its own logs and organizes them in folders. Even as it creates the logs, mIRC can introduce a level of organization. See Figure 4, page 56.

The program also has an interface for viewing and searching its logs. If the chat logs are reorganized into a different folder structure that is easier to manage, this interface can still see and interact with them. It looks like Figure 5, page 57.

This interface allows basic searching, sorting, and analysis for small collections of chat logs. The controls at the top search and filter which chat logs are listed. The bottom of the interface lists log files that meet the criteria in the top half. It provides ways to view and manage them. Any chat can be opened by double-clicking on it. By default, it will open in a text editor, such as

Notepad, where specific terms can be searched for within that chat. It is also possible to merge related logs into a single file. Select multiple files from the list (or all of them), and click the Merge button. This combines the selected files into one new file.

Several examples of the use this interface to search and analyze chat logs follow.

- To view who the target has logs of conversation with, sort by name.
- To view the subject of the chats, search for the target's user name.

The resulting list of log files would be the chats in which he spoke. These may be merged into one file for further analysis.

- To find out who was talking about a particular topic and when, search for a term linked to the topic.

Sorting the results by name (first) and date (second) will reveal who was involved in conversations about the topic, and when the conversations took place.

C. Logs

Commercial, off-the-shelf programs are sometimes best for managing electronic evidence of the type they are designed to manipulate. A program can often manage its own logs. Some types of evidence, however, have no readily available management program, or if a program is available, it may not do everything needed. This is often the case with raw log evidence. Log evidence is a file generated by a computer that usually records events sequentially. These files can be logs of system events (each time a user logs on) or activities (a file server may log every file transfer). Network elements, such as fire walls, can also generate logs that record activity on a network. These log files can easily be millions of records, and tools and techniques for managing them quickly become insufficient.

Microsoft Excel can open small log files, but there are several limits to its usefulness.

- First, in versions up to Excel 2003, a worksheet could not have more than 65,536 rows.
- Excel 2007 can now have up to 1,048,576 rows, so it can, at least theoretically, open most typical log files.
- When Excel opens a file, it attempts to load the entire file's data into memory at the same time. For large files this can be impractically slow.
- Excel's final limitation is that its search and analysis capabilities are far inferior to those of databases.

The Cybercrime Lab has had great success using custom Microsoft Access Databases to manage log evidence. Using Access has several benefits.

- It is already installed on most computers.
- It is reasonably easy for people with other technical experience to learn and use.
- It is powerful enough to handle all but the most voluminous log evidence (the Lab almost never needs to move to a more robust database with bigger capacity).

Not everyone is comfortable working with databases. It is likely, however, that someone in the organization has an aptitude for basic database work and can assist in the investigation.

The same tools and techniques can be used for any kind of log evidence but, for the sake of clarity, one type of log is used as an example. The Cybercrime Lab manages log evidence for many "warez," or online piracy cases. Targets in these cases often use file servers where each file transfer is logged. Each time a file is transferred, a line is written to a log file with information (date and time, the file name, the direction (upload or download), and the user's name). These log files can easily grow to be millions of lines. Managing, searching, or making sense of them as text files quickly becomes difficult, and oftentimes impossible.

The logs are easy to manage, once they are imported into a table in an Access database. The essential step is to split each line of the log into pieces, and put each piece into a separate column in the table. In the transfer log databases, a row in the table represents one line of the log. Every row has a column for each piece of information. For example, there is a date/time column, a filename column, and a direction column, among others. Splitting each log line into its parts is essential. It allows use of the full power of the database. Depending on the format of a log file, Access may be able to import it and split each line into separate fields simultaneously. Otherwise, a simple Visual Basic module can parse the log files into pieces and perform any additional logic necessary, while it imports them. A sample code for importing lines of a log file into a table in Access follows:

```
Public Function import (path As String)
    Dim rs As Object 'destination table
    Set rs =
CurrentDb.OpenRecordset("tablename")
    Dim pcs() As String 'pieces
    Dim inp As String 'line read from input file
    Open path For Input As #1 'open file for input
        'import file
    Do While Not EOF(1) 'check for end of file
        Line Input #1, inp 'read line of data
        If inp <> "" Then
        'split line
        pcs = Split(inp)
        'put in record
        With rs
            .AddNew
            .field1 = pcs(0)
            .field2 = pcs(1)
            .field3 = pcs(2)
            'etc.
        End With
    End If
    Loop
    Close #1 'close file
    rs.Close
End Function
```

Once the logs are in a table in a database, any searching or sorting is accomplished by creating queries. A query is a structured method of retrieving data from a database to answer a question. Access's interface guides the user through the process of setting up a query. The operator selects the fields needed to answer the question, the fields to sort, and any needed conditions. Save the queries created because Access can rerun it. If data is added or changed, and if the query is needed again, the answer will include the updated information.

Examples of queries follow.

- The investigator wants to know who was using a particular file server. A query was constructed that told the database, "show just the user column, sort it in alphabetical order, and do not show duplicates." When the query is run, the database quickly generates a new set of data (a mini-table) that answers the exact question described. It gives an alphabetical list of unique user names. Databases easily run through millions of records and give an answer in a few seconds.
- A list of a target's transfers is needed. This query consists of the file transferred, date of transfer, and user's name. The query is sorted by the date and limited to one user (a condition for the user column). The answer to this query will also be obtained in a matter of seconds.
- Who are the most active users on a server? The answer to this question is a table with three columns: user, a count of his uploads, and a count of his downloads. Counting

> something for each user requires a crosstab query. Fortunately, Microsoft knows it is a little more complicated, so the designers provide a special wizard that walks you through creating this type query. It is not necessary to understand how it works. The wizard is used to describe what is wanted.

In the Cybercrime Lab, there are numerous related cases like this and many people need to use the evidence. A user interface was programmed to make the functions described above appear as a friendly program. A screen shot of the program's main window is below. It uses tabs to group tasks (Database Management) and questions (General Lists, Analysis—All Users, Analysis—One user, for example). On each tab there is a button for each query. If the mouse is pointed at a button, it shows a brief description of what the query does. A database application with a user interface like this certainly is not necessary for every case. It may be appropriate, however, when the power of a database needs to be harnessed and the data retrieved needs to be made available to a large number of nontechnical users. See Figure 6, page 58.

IV. Conclusion

The staff of the Cybercrime Lab hopes that the strategies and examples in this article will help in the management of electronic evidence. The Computer Crime and Intellectual Property Section, and the Lab personnel, are available to AUSAs for consultation on these issues as well as computer forensic and other technical investigative matters. The staff can be reached at (202) 514-1026. Many other resources are available on the section's public Web site, www.cybercrime.gov. In addition, anyone in the Criminal Division or U.S. Attorneys' offices can find additional resources on the new intranet site, CCIPS Online. Go to DOJ Net and click on the "CCIPS Online" link. AUSAs are also encouraged to take advantage of the many courses we present at the National Advocacy Center throughout the year.❖

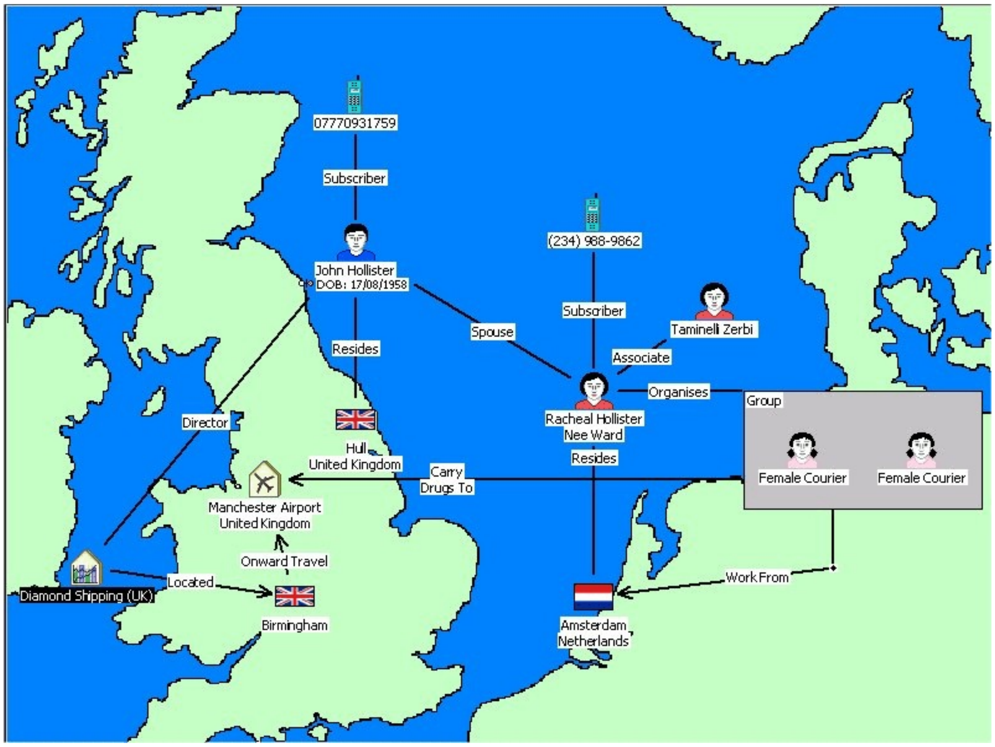
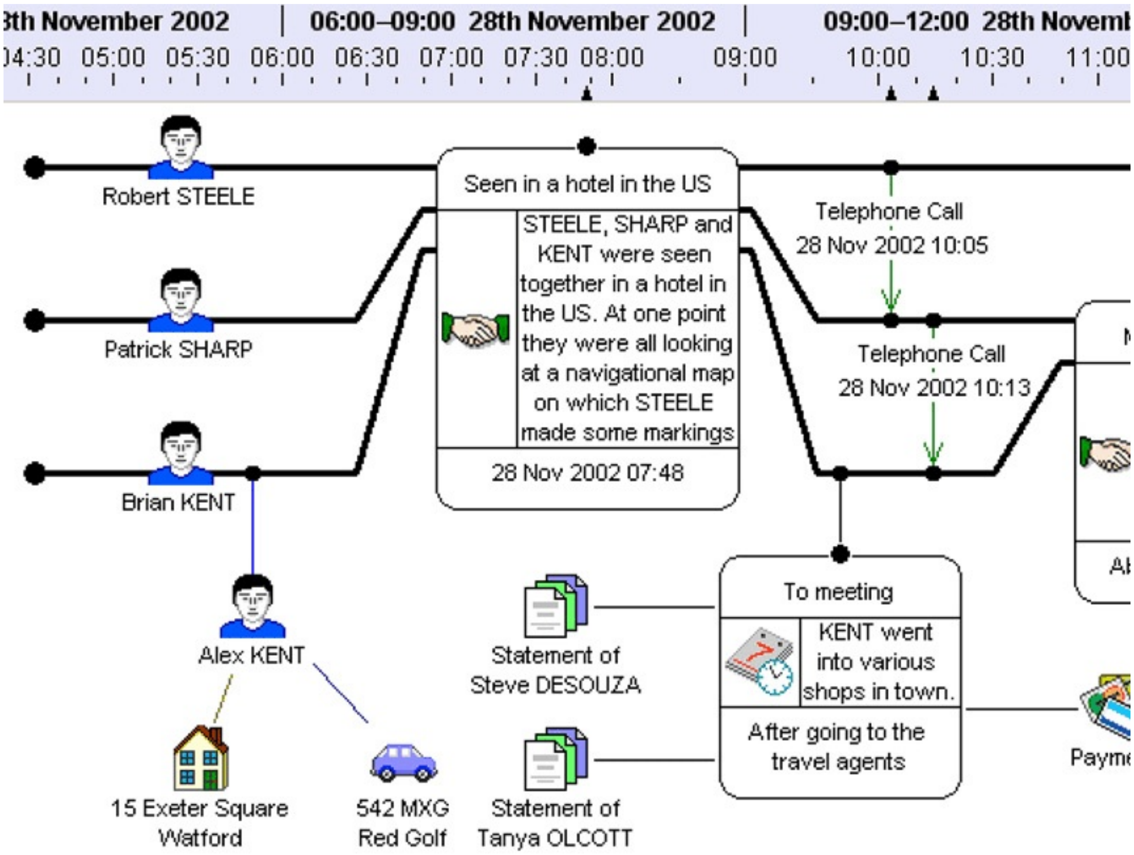

07770931759
Subscriber
John Hollister
DOB: 17/08/1958
Resides
Hull
United Kingdom
Director
Spouse
(234) 988-9862
Subscriber
Taminelli Zerbi
Associate
Racheal Hollister
Nee Ward
Organises
Group
Female Courier
Female Courier
Resides
Carry
Drugs To
Manchester Airport
United Kingdom
Onward Travel
Diamond Shipping (UK)
Located
Birmingham
Amsterdam
Netherlands
Work From

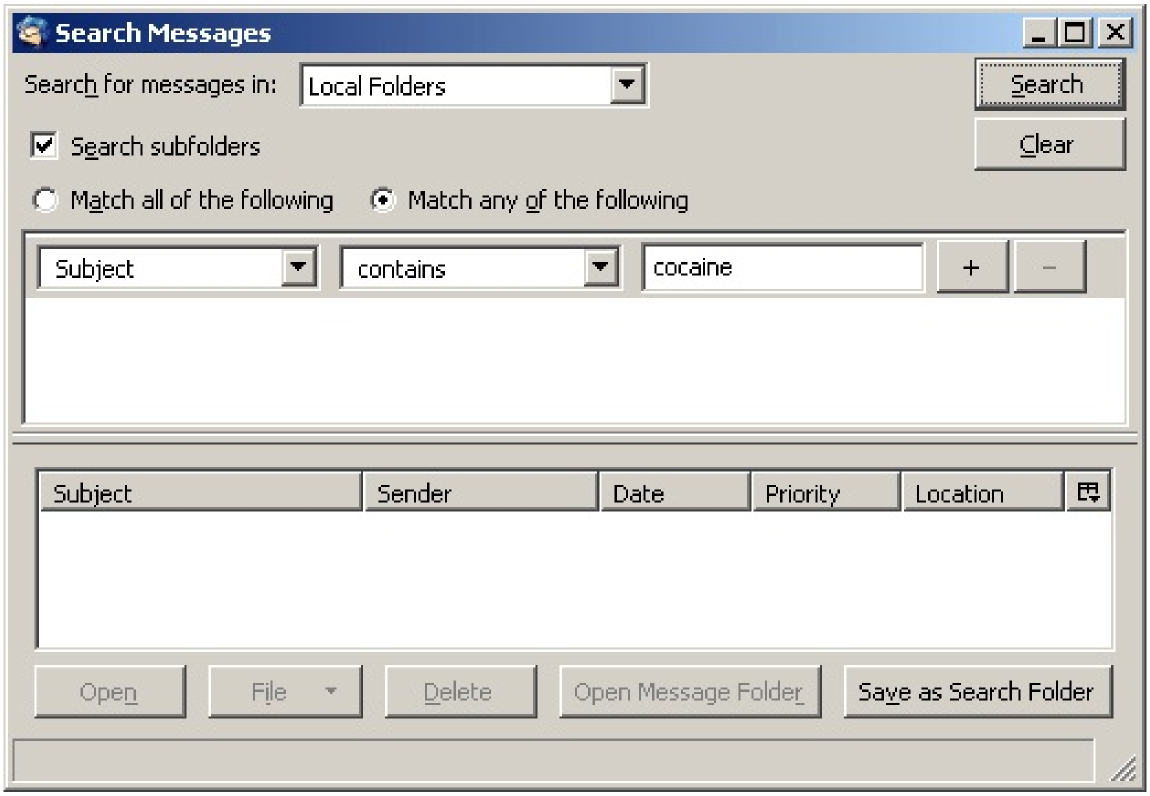
FIGURE 1



STEELE, SHARP and KENT were seen together in a hotel in the US. At one point they were all looking at a navigational map on which STEELE made some markings

KENT went into various shops in town.

FIGURE 2



Search Messages

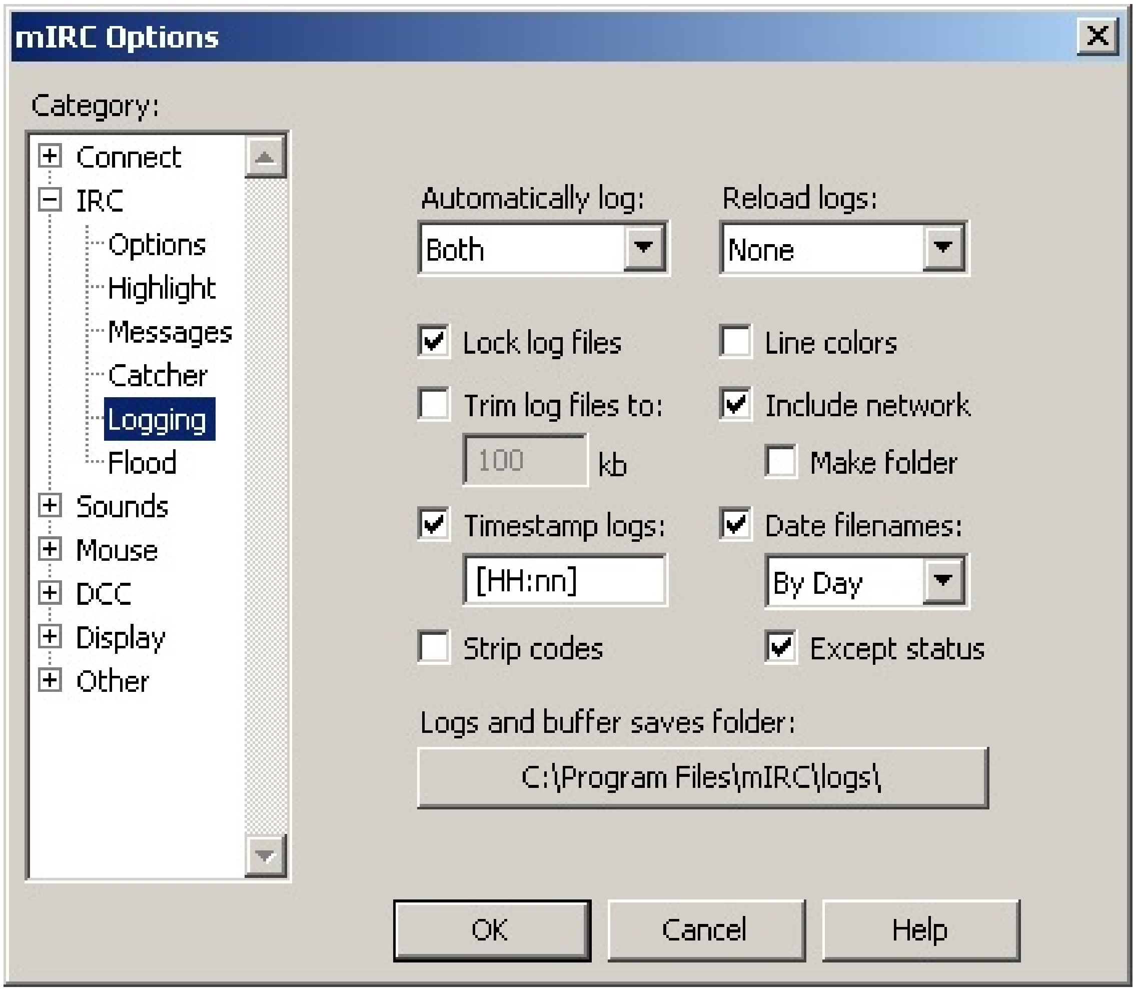
Search for messages in: Local Folders

☑ Search subfolders

○ Match all of the following ● Match any of the following

Subject contains cocaine

FIGURE 3



Automatically log: Both

Reload logs: None

☑ Lock log files

□ Line colors

□ Trim log files to: 100 kb

☑ Include network

□ Make folder

☑ Timestamp logs: [HH:nn]

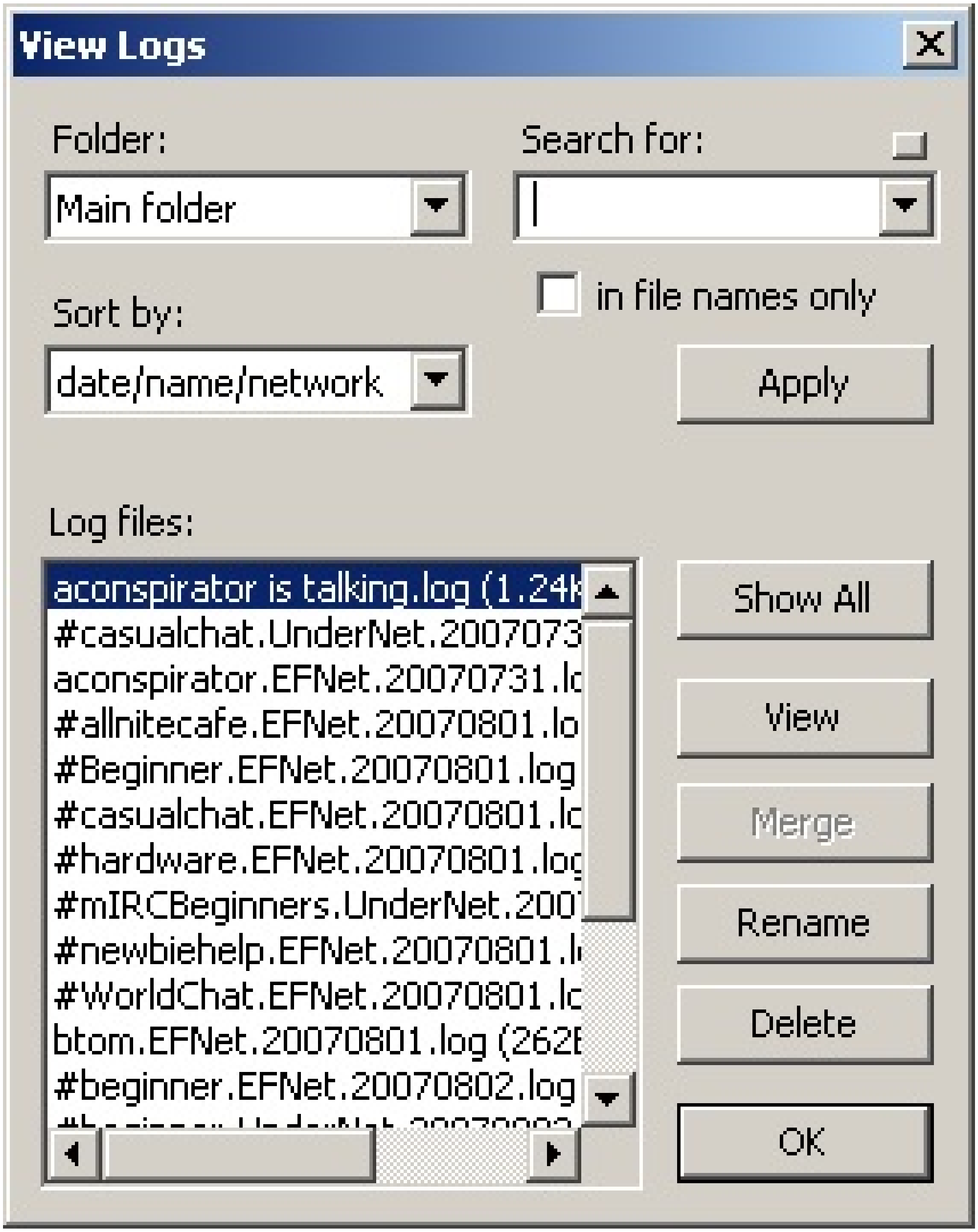
☑ Date filenames: By Day

□ Strip codes

☑ Except status

Logs and buffer saves folder: C:\Program Files\mIRC\logs\

FIGURE 4

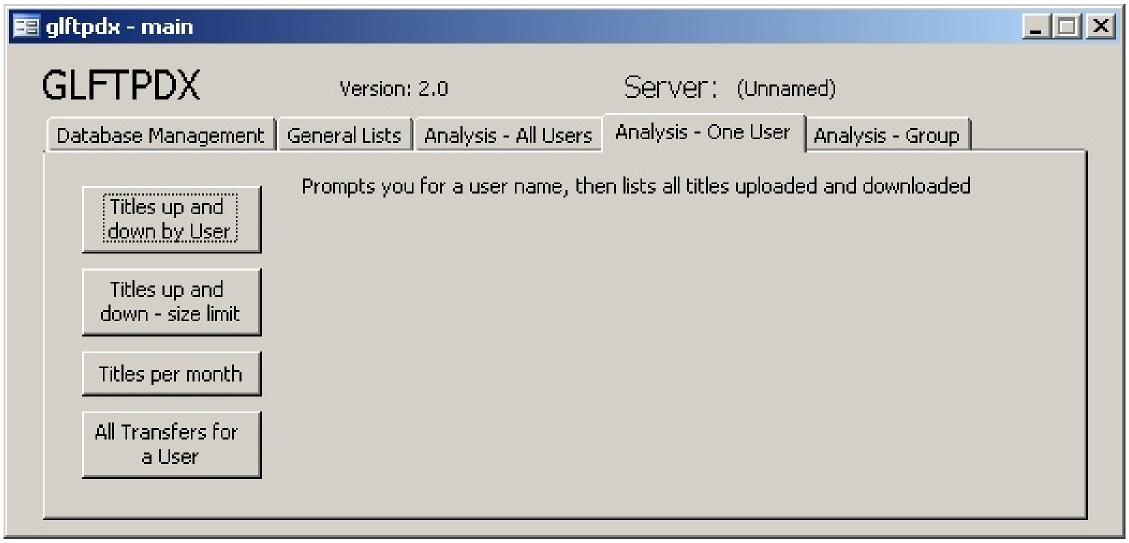


Folder: Main folder

Sort by: date/name/network

```
aconspirator is talking.log (1.24k
#casualchat.UnderNet.2007073
aconspirator.EFNet.20070731.lc
#allnitecafe.EFNet.20070801.lo
#Beginner.EFNet.20070801.log
#casualchat.EFNet.20070801.lc
#hardware.EFNet.20070801.log
#mIRCBeginners.UnderNet.2007
#newbiehelp.EFNet.20070801.l
#WorldChat.EFNet.20070801.lc
btom.EFNet.20070801.log (262E
#beginner.EFNet.20070802.log
```

FIGURE 5



Prompts you for a user name, then lists all titles uploaded and downloaded

FIGURE 6

ABOUT THE AUTHORS

❑**Ovie L. Carroll** is the Director of the Cybercrime Lab in the CCIPS. He has over twenty years of law enforcement experience. He previously served as the Special Agent in Charge of the Technical Crimes Unit at the Postal Inspector General's Office and as a special agent with the Air Force Office of Special Investigations.

❑**Stephen K. Brannon** is a Cybercrime Analyst in the CCIPS's Cybercrime Lab. He has worked at the Criminal Division in the Department of Justice and in information security at the FBI.

❑**Thomas Song** is a Senior Cybercrime Analyst in the CCIPS's Cybercrime Lab. He has over fifteen years in the computer crime and computer security profession. He specializes in computer forensics, computer intrusions, and computer security. He previously served as a Senior Computer Crime Investigator with the Technical Crimes Unit of the Postal Inspector General's Office.

The Cybercrime Lab is a group of technologists in the CCIPS in Washington, DC. The lab serves CCIPS attorneys, Computer Hacking and Intellectual Property (CHIP) units in the U.S. Attorneys' offices, and Assistant U.S. Attorneys, by providing technical and investigative consultations, assisting with computer forensic analysis, teaching, and conducting technical research in support of Department of Justice initiatives.✠

Rethinking the Storage of Computer Evidence

Tyler Newby
Trial Attorney
Computer Crime and Intellectual Property Section
Criminal Division

Ovie L. Carroll
Director, Cybercrime Lab
Computer Crime and Intellectual Property Section
Criminal Division

I. Introduction

When a federal criminal investigation involves computer evidence, prosecutors and investigators often rely on specially trained and accredited personnel in the field of computer forensics. These forensic examiners are typically responsible for the collection, processing, and analysis of digital evidence acquired during an investigation. Primary among the computer forensic examiner's duties is to ensure that the data seized during an investigation remains unaltered through the trial.

The foundation of electronic evidence collection and analysis, and the subsequent admissibility and use of that evidence at trial, is the creation of a forensic image. Once a forensic image of the original data is created, it is typically copied to a hard disk drive, which is then stored in a locked evidence room. Chain of custody logs are maintained when anyone accesses the hard drive image.

In complex cases, such as intrusion cases, a prosecutor or case agent may request full forensic analysis of an image to search for evidence to be used at trial. In less complex cases, the investigative team may want to conduct a triage review of the image to search for easily identifiable evidence of a crime, such as pirated software and movies, chat logs and e-mails discussing the crimes, digital photographs, and the like. In that situation, the case agent may want to review a working copy of the forensic image. If so, he or she must request that a working copy image be made.

In either instance, case agents and prosecutors are likely to encounter a long wait when they ask for assistance from computer forensic specialists. As electronic storage of data has become increasingly common, the demands placed on a limited pool of computer forensic examiners have increased. For example, the Federal Bureau of Investigation's (FBI) *FY2008 Authorization and Budget Request to Congress*, noted that its Computer Analysis and Response Team's (CART) case backlog increased 58% from FY2004 to FY2005 (1,258 cases to 1,991cases), and is likely to continue to increase in the future. As electronic communication devices, home networks, and increasingly capacious hard drives become more prevalent, already thinly-stretched investigative resources are likely to be in even more demand. Thus, it is possible that a hard drive containing evidence that a prosecutor needs to prepare and try a case will sit on a shelf for several months, if not years.

This reality raises the basic question of whether storing an increasing number of hard drives—which, like all things mechanical, can break—for years on shelves in evidence rooms is the best way to store digital evidence. An alternative evidence storage method for forensic images is to store them on secured Redundant Array of Independent (or Inexpensive) Disks (RAID) systems. A RAID is a category of disk drives that employ two or more drives in combination, for fault tolerance and performance. This method may save space in evidence rooms and will better protect sensitive evidence from inadvertent destruction. Furthermore, storing images on a RAID, if done properly, will not affect authentication of the image as a duplicate of the original electronic media at trial.

This storage technique most clearly applies to cases in which investigators make an image copy of the electronic media at the scene. When computers containing electronic evidence are removed from the crime scene, the use of RAID storage is also appropriate. Prosecutors should consider the possibility of defense challenges, however, before wiping the original computer hard drive or returning it to its owner. Of course, if the computer hardware is seized because it is contraband, the fruit of a crime, or an instrumentality, it should be retained pending disposition of the case or forfeiture proceeding.

II. The basics of forensic imaging

Forensic imaging is the process used to obtain a bit-for-bit copy of the data residing on the original electronic media obtained by law enforcement. The media may be a single hard disk drive, flash memory card, digital versatile/video disk (DVD), compact disc, or mobile phone subscriber identity module (SIM) card. The imaging process entails copying all of the data present on the original storage media device, including system files, hidden and deleted data from allocated (partitioned), unallocated (unpartitioned), and free space (unused space on a formatted partition).

The image of the hard drive contains all logical files, erased files, and unused space, which are available to the original hard disk drive. The investigator can examine the image for relevant evidence, without accessing the original, seized hard drive. This process allows investigators to review a duplicate of the original evidence while preserving that evidence in exactly the form it existed at the time of seizure.

III. Evidentiary issues raised by forensic imaging

Prosecutors and investigators must be mindful that the ultimate goal of any investigation is to acquire evidence that will be admissible at trial. The creation of a copy of original electronic evidence raises authentication, best evidence, and reliability concerns. How can one be sure the forensic imaging process produced a true copy of the original evidence? Could the forensic image have been altered or corrupted in the time between its creation and offering it into evidence at trial?

A. Best evidence issues

Federal Rule of Evidence 1002 requires the use of an original writing, recording, or photograph, to prove the contents of those items, unless provided otherwise by federal statute or the Federal Rules of Evidence. FED. R. EVID. 1002. The exception that proves the rule for forensic images is Federal Rule of Evidence 1003, which provides that a "duplicate" is admissible, to the same extent as an original, unless a genuine challenge is made to the authenticity of the original, or it would be unfair to admit the duplicate instead of the original. FED. R. EVID. 1003. Federal Rule of Evidence 1001(4) defines a duplicate as a copy of the original made by, among other things, "mechanical or electronic re-recording . . . or by other equivalent techniques which accurately reproduces the original." FED. R. EVID. 1001(4). Thus, the focus must be on whether the image is an accurate and authentic reproduction of the original evidence.

B. Authentication of forensic images

Authentication is a predicate to the admissibility of any physical evidence. *See*

FED. R. EVID. 901(a). To satisfy Federal Rule of Evidence 901, the proponent must produce "evidence sufficient to support a finding that the matter in question is what its proponent claims." *Id.*; *see, e.g., United States v. Simpson*, 152 F.3d 1241, 1250 (10th Cir. 1998). This requirement is typically easy to satisfy when the evidence is a single document and a cooperating witness, such as a recipient, author, or custodian, is available to authenticate it.

While the authentication requirements for computer data are no different than for other forms of evidence, authentication can appear more daunting when the data was extracted from a copy of the media that was made outside the defendant's presence. It is likely that the seized media has been in an evidence room for an extended period of time. These factors, combined with the ease (perceived or real) of altering computer data without notice, may tempt a particularly aggressive defense counsel to challenge the authenticity of the proffered data.

Courts have generally looked askance at authenticity challenges to electronic evidence that are unsupported by anything other than speculation that the original data was altered by an unseen hand. *See, e.g., United States v. Whitaker*, 127 F.3d 595, 602 (7th Cir. 1997) (affirming admission of computer records where allegation of tampering was "almost wild-eyed speculation . . . [without] evidence to support such a scenario."); *United States v. Glasser*, 773 F.2d 1553, 1559 (11th Cir. 1985) ("The existence of an air-tight security system [to prevent tampering] is not, however, a prerequisite to the admissibility of computer printouts. If such a prerequisite did exist, it would become virtually impossible to admit computer-generated records.") In *Whitaker*, the Seventh Circuit upheld a district court's admission of printouts of spreadsheets from the original computer seized, where the FBI agent involved in the seizure and the printing testified as to their authenticity. *Whitaker*, 127 F.3d at 602. Despite the permissive standard applied in *Whitaker,* good trial strategy is to foreclose potential authenticity challenges before they are raised.

IV. Hash algorithms—an answer to evidentiary issues

To blunt potential authentication challenges to data extracted from a forensic image, it is useful to have a procedure to verify that the data on the image is an exact match of the original media. Computer forensic specialists have developed a procedure that guarantees just that. This process uses "hash" algorithms, which verify that the acquired image is an exact copy of the original media. The most commonly used hash algorithms—the Message Digest 5 (MD5) and Secure Hash Algorithm-1 (SHA-1)—take as input a message of arbitrary length, and produce as output an n-bit "fingerprint" or "message digest" of the input. The algorithm then produces a digital signature which can be used to identify a uniquely given file, and therefore establish that the image is an authentic copy of the original evidence.

Verification using hash algorithms is highly reliable. The odds of two random files having the same hash are astronomically small—estimated to be approximately a 1 in 10^{38} chance. Moreover, the use of the hashing algorithm is a one-way function. This means that it is easy to create a hash from a file, but almost impossible to create a file matching a particular hash.

Hash validation, when combined with evidence of a chain of custody between the time the original computer media was seized and the image was created, is strong authenticating evidence that the forensic image is an exact duplicate of the original. Hash algorithms fit the examples listed in Federal Rule of Evidence 901(b)(4) of "distinctive characteristics" that can be used to authenticate evidence. FED. R. EVID.

901(b)(4). What are hashes if not indicators of "internal patterns, or other distinctive characteristics" of data?

Although published decisions addressing the use of hashing algorithms to authenticate forensic images are few, they are uniform in recognizing hashes as a proper means of establishing authenticity. *See, e.g., Williams v. Sprint/United Mgmt. Co.,* 230 F.R.D. 640, 655 (D. Kan. 2005) (recognizing that hashing "allows a large amount of data to be self-authenticating with a rather small hash mark, efficiently assuring that the original image has not been manipulated."). In *Williams*, the district court rejected a civil litigant's purported concerns about producing electronic evidence in its native format by noting that the parties could detect any alteration by comparing hash values. The court found that a hash value is a "'digital fingerprint' akin to a tamper-evident seal . . . the file cannot be altered without a change also occurring in the hash mark." *Id.* at 655; *see also Ohio v. Morris*, 2005 WL 356801, No. 04CA0036 (Ohio App. Feb. 16, 2005) (admitting forensic image even where testimony established that imaging software had validated the MD5 hashes of the original and image matched before forensic examiner erased the original hard drive); *Krause v. State*, 2007 WL 2004940, No. 01-05-01136-CR (Tex. App. July 12, 2007) (forensic analyst's methodology was sufficiently reliable for purposes of expert testimony, where analyst used forensic software that compared hashes on the image and the original media). Similarly, the Federal Judicial Center has identified MD5 and SHA hashes as commonly used algorithms to establish the authenticity of a forensic image. *See* FEDERAL JUDICIAL CENTER, MANAGING DISCOVERY OF ELECTRONIC INFORMATION: A POCKET GUIDE FOR JUDGES, FEDERAL JUDICIAL CENTER 24 (2007), *available at* http://www.fjc.gov/public/pdf.nsf/lookup/eldscpkt.pdf/$file/eldscpkt.pdf, quoted with approval in *Lorraine v. Markel American Ins. Co.*, 241 F.R.D. 534, 536-37 (D. Md. 2007).

V. Storing forensic images—an alternative to the shelf

Provided that proper chain of custody is established between the times the original computer media are seized and forensic images are created, the hash verification process eliminates any concerns that the forensic image was altered prior to trial. The practical concern of how and where to store the forensic images remains.

While the prevailing method of storing forensic images is certainly adequate and relatively simple, it has shortcomings. First, as anyone who has dealt with electronic evidence knows, hard disk drives fail. A recent study of 100,000 different types of hard disk drives, conducted by researchers at Carnegie Mellon University, found that the actual reported failure rate of hard disk drives is much higher than stated in manufacturers' data sheets. Bianca Schroeder and Garth A. Gibson, *Disk Failures in the Real World: What Does an MTTF of 1,000,000 Hours Mean to You?*, FAST07, 5TH USENIX CONFERENCE ON FILE AND STORAGE TECHNOLOGIES (2007), *available at* http://www.cs.cmu.edu/~bianca/fast07.pdf. Although the observed real world failure rates were approximately 2%-4% (with some as high as 13%), which are relatively low, a prosecutor does not want to request a continuance of trial because the hard disk drive on which the forensic image was stored failed. Moreover, frequent handling and transportation of hard disk drives inevitably jostles the sensitive mechanical parts in the drives and can only increase the potential for drive failure.

A more advanced and safer method of maintaining forensic images is to upload, or copy, the forensic image and hash to a fault tolerant RAID. The entire purpose of RAID storage is redundancy—if one disc in the array fails, the data remains secure on one of the other redundant discs. Also, unlike a powered-down hard disk drive, a running RAID system can be configured to conduct routine backups to tape archives, which can be stored off-site. This is a useful data recovery backstop in the event of a disaster, such as a flood or fire at an evidence storage location. Indeed, the implementation of secure RAID evidence storage appears to adhere to the National Institute of Justice, Office of Justice Programs recommendation that investigators preserve evidence "in a manner designed to diminish degradation or loss." DEPARTMENT OF JUSTICE, OFFICE OF JUSTICE PROGRAMS, NATIONAL INSTITUTE OF JUSTICE, CRIME SCENE INVESTIGATION: A GUIDE FOR LAW ENFORCEMENT (2000), *available at* http://www.ncjrs.gov/pdffiles1/nij/178280.pdf.

Moreover, a RAID storage system would simplify the process of locating evidence when requested. Forensic images could be stored in folders corresponding to investigation name and number, subject name, and search location, making it easier to locate desired images when they are requested by prosecutors.

When the time comes to use the image at trial, forensic examiners copy the image back to a hard drive and verify that the hash is unchanged. Hash validation after the image is transferred onto the RAID will ensure that the image stored on, and ultimately recovered from, the RAID is no different from the original data that was seized. A RAID-based storage system should not undermine the authenticity or reliability of the forensic image that is eventually offered into evidence at trial because it relies on the already approved hash validation process.

Care needs to be taken to keep the RAID in a secure setting, such as in a locked, limited access server room, with no Internet connections. Logging software could be added to the RAID to keep track of access to the virtual evidence lockers stored on it, and forensic images could be stored in password protected virtual lockers on the RAID. Testing should be performed before a RAID–based evidence storage system is put into use.

VI. Conclusion

Prosecutors interested in these and other computer forensic issues and techniques may register for the Computer Forensics for Prosecutors Course taught by CCIPS at the National Advocacy Center. The Computer Crime and Intellectual Property Section and the Cybercrime lab is also available to AUSAs for consultation on computer forensic and other technical investigative matters by calling (202) 514-1026. Many other resources are available on our section's public Web site, www.cybercrime.gov. In addition, anyone in the Criminal Division or U.S. Attorneys' offices can find additional resources on our new intranet site, CCIPS Online. Just go to DOJ Net and click on the "CCIPS Online" link.❖

ABOUT THE AUTHORS

❑**Tyler Newby** is a Trial Attorney with the Computer Crime and Intellectual Property Section of the United States Department of Justice's Criminal Division. Prior to joining the Justice Department, Mr. Newby practiced civil intellectual property litigation in Silicon Valley and San Francisco. While in private practice, Mr. Newby worked on multiple cases with computer forensic specialists who stored forensic images on secure RAID systems.

❑**Ovie L. Carroll** is the Director of the Cybercrime Lab in the Computer Crime and Intellectual Property Section. He has over twenty years of law enforcement experience. He previously served as the Special Agent in Charge of the Technical Crimes Unit at the Postal Inspector General's office and as a special agent with the Air Force Office of Special Investigations.✠

Request for Subscription Update

In an effort to provide the UNITED STATES ATTORNEYS' BULLETIN to all federal law enforcement personnel who wish to receive it, we are requesting that you e-mail Nancy Bowman (nancy.bowman@usdoj.gov) with the following information: Name, title, complete shipping address, telephone number, number of copies desired, and e-mail address. If there is more than one person in your office receiving the BULLETIN, we ask that you have one receiving contact and make distribution within your organization. If you do not have access to e-mail, please call 803-705-5659. Your cooperation is appreciated.