

Field Guide Part Eight

Timothy E. Wright 2001-03-21

The Field Guide for Investigating Computer Crime, Part Eight: Information Discovery - Searching and Processing

by *Timothy E. Wright*

last updated March 21, 2001

Last Time...

This is the eighth and final article in Field Guide for Investigating Computer Crime. In our last installment, [Information Discovery - Basics and Planning](#), we briefly compared the physical search and seizure with its logical (i.e. data-oriented) counterpart, information discovery. We introduced the basics for the information discovery process, noting how establishing and protecting the chain of custody for logical evidence was delightfully straight forward! We then discussed three basic rules of thumb that should act as guides for any information discovery, and mentioning along the way how each rule has a parallel in the world of physical search and seizure. We are now ready to bring things to a close by examining the final two stages - searching for and processing data evidence. So! Without further ado, let us tackle the remaining stages of information discovery...

The Remaining Stages of Information Discovery

Unlike search and seizure, information discovery is a more compact process in terms of tools and effort. You'll recall from the last installment in this series, that in search and seizure of information, we are concerned only with locating, retrieving, and then processing relevant data in a secure fashion. Hence, the remaining stages required to carry out an information discovery are limited to two: searching for, and processing informational evidence.

Searching for Information

(A)	(B)	(C)
Formulate plan	Search For Evidence	Process Evidence

Figure 1: Searching - Stage B of an Information Discovery

There is no specific predetermined procedure that is used to search for information, as the steps required will depend entirely upon the case being investigated. For example, depending on the case in question, an investigator may need to create mainframe batch jobs to sift through hundreds of megabytes of log data. On the other hand, the situation may create a need to comb the various accounting files maintained by a UNIX server. Or circumstances may demand that the investigator search a router's log files for certain signs of network activity.

The possibilities here are clearly endless. Because of this, the steps taken to gather data during an information discovery are really left up to the circumstances and the investigator's practical know-how and imagination. Once data are located and are ready to be analyzed, however, the whole process returns to a more structured format, with strict guidelines around checking information files into the evidence preservation lab, and handling them afterwards.

The only step that is definitely required to be taken during information discovery is for the investigator to keep a detailed log of everything he is doing. In essence, this log is the counterpart of the one maintained during the search and seizure at a computer crime scene (see Part Six of this series, [Search and Seizure Evidence Retrieval and Processing](#)), and is aptly referred to as the information discovery evidence log. Log entries should begin as soon as the investigator starts information discovery, and should include:

- date and time of the investigation;
- investigator's name;
- description of what the investigator is attempting to discover;
- description of how the investigator is proceeding; and,
- description of results.

Processing Information Discovery Files

(A)	(B)	(C)
Formulate plan	Search For Evidence	Process Evidence

Figure 2: Processing Evidence - Stage C of an information discovery

After the investigator has succeeded in gathering data from an information discovery, he should commence the task of bringing these data into the evidence preservation lab, at which point he can begin his forensic examination. As with physical computer evidence (such as computer systems, hard drives, media, etc.,) data must also go through an accounting procedure to ensure that the chain of custody is maintained, and be introduced into the lab environment. This procedure consists of the following steps:

1. backup the information discovery file or files;
2. start the lab evidence log;
3. mathematically authenticate the information discovery file or files; and,
4. Proceed with the forensic examination.

Step 1: Backup the Information Discovery File or Files

Note that backups of information discovery files do not need to be of the bit stream variety. This is because the information discovery process does not take place on seized media, and so does not involve preserving file systems as evidence. The [search and seizure](#) process detailed earlier in the field guide deals with preserving file systems.

Under the search and seizure steps given earlier in the field guide, backups do not take place until crime scene computers, devices and media are brought to an evidence preservation lab. Hence, as a prelude to backing up the data stored on these items, a check-in step is carried out. In the information discovery process, however, the evidence being scrutinized is not physical, and often is not high capacity (say, more than 500 megabytes). For these reasons, things can be handled in a slightly different order than with search and seizure. In particular, it is convenient that the first step of processing information discovery files be to back up these files.

As expected, forensic work on discovered information should always take place on backup copies. It is most helpful to back such data up to some form of random access media (i.e., ZIP or JAZZ cartridge, read/write CD, floppy, or hard drive) simply because it expedites their storage and retrieval. Whenever possible, backups should be implemented in a raw, uncompressed format, creating duplicates that better mirror their originals. Also, this removes a layer of complexity in accessing the backups. It is a good idea to use a revision control system to maintain and manage the backups of information discovery files. If the investigator elects to use such a system, copies of the files to be inspected should be placed on an evidence preservation lab machine running the revision control software.

Step 2: Start the Lab Evidence Log

A lab evidence log that details discovered information must be kept within the lab. This log is identical in form and purpose to the log discussed in Search and Seizure Evidence Retrieval and Processing, in fact the same log can be shared for search and seizure, and information discovery. The following data should be entered into this log:

- current date and time;
- description of each information discovery file, including the file's format (e.g., ASCII, EBCDIC, binary, Postscript, etc.); and,
- name of the investigator checking in the file(s).

A revision control system, such as Revision Control System (RCS) or Source Code Control System (SCCS)², provides an excellent way to ensure the integrity of any information discovery file (binary or text). With such systems, backup copies of original files are automatically managed: only one user at a time can have any given file checked out, and there is a simple audit trail maintained to show who has what files checked out at any point in time. Of course, some sort of strict locking should be implemented to keep accidental changes to a given file from being checked in (although, such an accident will not harm the file's original version).

After any information discovery files have been checked into the lab, anyone wishing to interact with the files, must check copies of them out, and then back in when finished. Checking files in or out should include logging the following in the lab evidence log:

- current date and time;
- identification of the file(s); and,
- name of the investigator checking the file(s) in or out.

Step 3: Mathematically Authenticate the Information Discovery File or Files

At some point the investigator needs to authenticate any data located during an information discovery. Such authentication is necessary to confirm that no alteration of electronic evidence has taken place since the evidence has been in the investigator's care. The md5sum utility is useful for generating 128-bit hashes (e.g., digital fingerprints) of files. Because the MD5 algorithm is computationally secure (which is to say, it would take an impractical amount of time to generate a file that matches a pre-determined MD5 hash value), it provides an excellent

way to prove the authenticity of forensic evidence. To authenticate files:

1. find the MD5 message digest for the original information discovery file or files;
2. log all message digest values in the lab evidence log; and,
3. When forensic work is complete, regenerate the message digest values using the backups on which work was performed; log these new values along side the hashes that were originally generated. If the new values match the originals, it's reasonable to conclude that no evidence tampering took place during the forensic examination of the information file(s)

Step 4: Proceed with the Forensic Examination

Using the backups from Step 1, the investigator may safely perform forensic examinations of the information discovery files. As always, all forensic work should be carefully noted in the lab evidence log, with each entry including:

- current date and time;
- description; and,
- name of the investigator.

Fin!

With this installment of the Field Guide for Investigating Computer Crime, we have concluded our discussion of the information discovery process. In particular, we described the two remaining stages of information discovery in detail, including searching for and processing logical or data-based evidence.

The notable difference between searching for physical evidence and searching for logical evidence, is that in the latter there is much less structure. Because the format and location of information varies tremendously from case to case, how information is discovered really depends on the circumstances of the case and the imagination of the investigator. Once information is found, however, rigorous methods are applied to its handling and processing.

The methodology presented in this field guide addresses the two broad areas within which computer forensics may be applied: search and seizure and information discovery. Although different in their implementations, both of these areas share a few prominent common principals. These include the important concept that evidence should always be backed up and digitally authenticated prior to forensic work. As well, both approaches require that everything

the investigator does should be carefully documented. In addition, for both areas, the evidence preservation lab plays an important role as a secure, controlled environment for computer forensics work and evidence storage. Without such a facility, the investigator will have a difficult (if not impossible) time maintaining the chain of custody while examining and holding evidence. Finally, the use of secure case-management software is highly desired, since it lends structure, efficiency, and safety to the gathering and management of case notes and data.

One closing comment about how our methodology relates to a case that may potentially end up in litigation. In a venue where law enforcement authorities are investigating a computer crime, there is a measurable chance that a case could find its way to court. Within a corporation or other organization, however, things are vastly different. Companies loathe being involved in litigation - even in situations where it appears the law is on their side!

It's no surprise that legal fees and bad publicity can take a mighty toll on the "bottom line." For this reason, much of what the corporate computer fraud and abuse investigator does is for naught. It's easy for a corporate investigator to become frustrated and even disillusioned with his work when he sees good cases ending up on the wayside due to fears of bad P.R. Such feelings must be contained, as they will quickly result in laziness and incomplete work on the part of the investigator.

Most of the computer crime cases handled by the corporate investigator won't end up in litigation; however, this does not apply to all cases! Even a seemingly low-profile case can take a sudden twist and end up garnering the attention of the CEO. Since practically any case can turn into a matter for litigation, the corporate investigator needs to treat all cases with a proper and reasonable amount of attention.

For the past several years, Timothy Wright has been investigating computer fraud and abuse as a Senior Technology Investigator at one of the country's largest financial corporations. Before then, he worked as a lead developer within the financial industry, designing and building web-based home banking software. He holds an M.S. in Computer Science, and a B.A. in Philosophy.

Relevant Links

[Subscribe to the FOCUS-Incident Handling Mailing List](#)

SecurityFocus.com

[An Introduction to the Field Guide for Investigating Computer Crime](#)

Timothy E. Wright, SecurityFocus.com

[The Field Guide for Investigating Computer Crime Part 2: Overview of a Methodology for the Application of Computer Forensics](#)

Timothy E. Wright, SecurityFocus.com

[The Field Guide for Investigating Computer Crime Part 3: Search and Seizure Basics](#)

Timothy E. Wright, SecurityFocus.com

[The Field Guide for Investigating Computer Crime Part 4: Search and Seizure - Planning](#)

Timothy E. Wright, SecurityFocus.com

[The Field Guide for Investigating Computer Crime Part 5: Search and Seizure - Approach, Documentation, and Location](#)

Timothy E. Wright, SecurityFocus.com

[The Field Guide for Investigating Computer Crime, Part 6: Search and Seizure - Evidence Retrieval and Processing](#)

Timothy E. Wright, SecurityFocus.com

[The Field Guide for Investigating Computer Crime, Part 7: Information Discovery - Basics and Planning](#)

Timothy E. Wright, SecurityFocus.com

[Privacy Statement](#)

Copyright 2006, SecurityFocus