Computer forensics and litigation support have always seemed to be separate entities. Lately, these two fields are starting to come closer and closer together. Will there be a time when there is little to none discernable difference between the two? Possibly, but that is still a ways away. There might be a time where everything in relation to litigation support and electronic discovery will need to be done in a forensic sound manner.

It is the opinion of the author that most service providers are unaware of copying data properly and to do it in a forensic sound way. Even some law firms that have a litigation support department as a cost center are also unfamiliar with copying data in a forensic sound manner.

Pivotal Guidance a computer forensics company has some tools to help in relation to litigation support and electronic discovery that should make your life a lot easier.

Founders Jon Rowe (Formerly of Image-Capture Engineering Inc. Software) and James Beasley are CCE's (Certified Computer Examiners) and members of the ISFCE (International Society of Forensic Computer Examiners). Jon and James started Pinpoint Labs to address the digital forensic needs of corporations, legal departments and law enforcement.

Pinpoint Labs, a division of Pivotal Guidance provides computer forensic services that include custodial data acquisition, metadata analysis and computer investigations. Pinpoint Labs also designed several free applications to assist in computer forensic exams and electronic evidence analysis. Those will also be discussed in this document. Thousands of computer investigators, litigation support professionals and corporate IT personnel in more than 25 countries rely on Pinpoint Labs software every day.

PG PinPoint-

History on PG PinPoint from Pivotal Guidance.

"PG Pinpoint was created after we heard from several service bureaus that they needed a way to filter files based on NIST hash list. After trying to accomplish this using current forensic software the service bureaus discovered it was going to require sometimes 8-10 hours to complete. After extended research we were able architect a way to store the hash values and perform a high speed lookup which significantly reduced this process."

PG Pinpoint (referred from now on as PinPoint) is the parent product from PinPoint Labs. Its focus is to provide a quicker turn around on electronic discovery and forensic projects. Its main purpose is to eliminate “known” file types identified by the more than 35 million hash values in the NIST (National Institute of Science and Technology) list. To eliminate system and application files that cannot be converted. In essence to deNIST these files. PinPoint is also used to compare file lists and/or locate suspect files on a target system or network. This utility works off either a MD5 or a SHA-1 hash code.

What is a MD5 hash code? It stands for Message- Digest algorithm 5 and is a widely used cryptographic hash function with a 128 bit hash value. It has been used in a wide variety of security applications, and is commonly used to check the integrity of files. An MD5 hash is typically a 32 character hexadecimal number. Two documents with the same hash code will undoubtedly be duplicates. MD5 hash is still the standard in most electronic discovery applications. It is also known in this industry as an electronic file fingerprint. A few have also integrated SHA-1.

What is SHA-1? It stands for Secure Hash Algorithm and is one of 5 different cryptographic hash functions designed by the National Security Agency (NSA) and published by the NIST as a U.S. government standard. The other 4 algorithms are SHA-224, SHA-256, SHA-384, and SHA-512. They are sometimes collectively referred to as SHA-2. In some circles SHA-1 is considered to be the successor to MD5.

PinPoint uses hash lists (MD5 or SHA-1) from NIST and those added through the utility to identify commonly used operating system and application files. When performing electronic discovery or even forensic analysis on a custodian hard drive. A significant percentage of the files can be filtered out prior to processing.

It is much faster using PinPoint and filtering these files out than loading all the contents in an electronic discovery processing program. It is more efficient this way.

This utility is simple to use. The user needs to select to compare against which hash list. PinPoint writes the data for the files that are not referenced in the hash lists. This allows the user to also specify which delimiters to use. Be default, PinPoint will also write the path to files that are not included in the hash list to a log file. The log file can be used as a project list for other applications. PinPoint will begin a process of tallying the number of folders, files and bytes in the source directory. Once the tally is completed it will then begin comparing the files and displaying the status. Once it is finished the user can view the log file by opening it from the location the user specified. I would use something other than Notepad or WordPad. The log file can be quite large especially if it is from a hard drive. Those two programs tend to hang on large files. TextPad or UltraEdit are more powerful programs at opening large text or log files.

PinPoint can even create a hash list of all files in a directory.

PinPoint also is effortless in terms of getting the files you want ready to load into an electronic discovery processing program. The user will select which files they want hashed. There are two ways to copy these files to another location. One way is to check the option marked Robocopy files to a new location. The other option is to open up the file viewer in PinPoint and select the files in the list you want to copy out.

Robocopy is a utility from Microsoft that will copy files to a new location without changing the metadata. It is the preferred method of copying. Robocopy XP or newer must be accessible from the system shell in order to work with Pinpoint. Once the files are indeed copied over to a new location it is now ready to be processed in an electronic discovery program.

Copying files to DVD after filtering will alter file system metadata (the burning software uses the current date/time). The internal (OLE) metadata of the files will be the same but the file system metadata gets altered. An option may be copying files to a new hard drive or include a warning about the burning process to those that copy the files.

Another great feature of PinPoint is the ability to check for duplicates.

Another feature of PinPoint is to be able to continue to add to the hash list. For example say you as a law firm hashed a hard drive using MD5 algorithm. A week later you received a DVD and decided to hash it against the list that was created for the hard drive. Pinpoint found 12 matches. Those are duplicate files and can be excluded from the files you will send to a vendor.
Exclude those with Pinpoint and use Robocopy to copy to a location and burn a new DVD. (Remember by using RoboCopy with the correct switches the metadata does not get altered.)
The same can hold true for a vendor. This way the de-duplication process is handled outside of the electronic discovery processing program. The hash list and file names can be sent to your client. PinPoint besides being able to filter out system files also works on multiple levels as a filter tool. This way a user can continue to build a master (live) hash list.

PinPoint works great as either a tool for electronic discovery and or Forensics. In addition this utility also works with locating matching files.

Electronic Discovery- If a user needs to create a list of files from separate productions. By creating a hash list for each production PinPoint will provide a list of matching files.

Forensics- If a company believes certain files have been stolen or used by another company. This utility can verify if that is true by simply creating hash lists and then identifying if there is indeed a match.

Limitations of RoboCopy from Microsoft:

"In certain instances, Robocopy will not be able to copy all of the file system information that was available for the original file. These instances concern the complexities of the source and destination file systems. If the source file system is NTFS and the destination is FAT32, the auditing information and the last few least significant digits of the timestamps will be lost because there is no corresponding data point in FAT32 to which this information can be copied. Similar instances have been noted with Joliet, FAT16, FAT12, and a variety of other disparate file systems."

What need would a service provider have for this tool?

For this example let us say that law firm XYZ gave to 123 service provider an external hard drive that contained 36 gigs of data. The client has no idea what is on it but they do know that all of it needs to be processed. As a vendor here is what may happen.

Take hard drive and copy it to a location on your server. (Hopefully service provider has some type of chain of custody procedure in place.) Some service providers are unaware that by copying using Windows the metadata is being altered. The last accessed or last modified dates are changed to the copy date. If you are using XCopy you are still modifying the metadata. (The metadata is already compromised) Your next step could be taking these 36 gigs of data and creating a new project in an electronic discovery processing program. It will take quite a long time to load, extract metadata and possibly full text (there are some programs that will skip binary files for full text extraction). But another question is do you as a user really want records in your SQL or Access database of files which will be exceptions anyway? What if this hard drive has 20% or more of files that cannot be processed?

Another option which is also recommended by Randall Consulting. Take the hard drive and either use Robocopy or use a forensic write blocker to copy the data to a location on your network. Once that is done and verified then open PG PinPoint. Run a hash list against the master. Create a log of all files and their corresponding MD5 or SHA-1 hashes. This can then be given to your client of files that cannot be processed. Of course this does not include password protected files which you will not find out about until bringing it into an electronic discovery program. (PinPoint is unable to create hashes for files inside e-mail stores. It will give one single hash for the e-mail store. Theoretically there could be system and application files as attachments.) Once you have decided what files are left for processing then you can use RoboCopy to copy those files to another location. By having RoboCopy as a feature in PinPoint it assures that the metadata will not be altered. It also helps service providers and law firms copy data the right way. With the filtering complete, the data is now ready to be loaded and processed. You have probably saved hours in discovering the data.

A week later you receive another DVD for the same case. The client tells you that it is a continuation of the same project. You can then open up PinPoint and add the hashes for the new data against the old and see if you receive any hits. This list can only be a project hash list. This will take care of duplicate binary files on the front end. In this author's opinion the pitfalls of electronic discovery processing go away significantly if the possible issues are addressed on the front end. PinPoint is a utility to help you with that.

Before delivering this DVD project to your client you decide to run a compare file list against the hard drive data you delivered to your client last week. At 123 service provider you decide to think outside the box and make a file list of every single production thru PinPoint before sending it out the door. You now ask PinPoint to compare the two lists. To your horror you see that it found three items. Somewhere along the line your EDD program did not do as good of a job at removing duplicates as you had hoped. The good news is you found this before it went out. The bad news is you are going to have to redo the export and eliminate those three duplicate files from this DVD collection. This is another avenue where PinPoint comes in handy.

What need would a law firm have for this tool?

At a minimum, this is a great tool to have from a computer forensics point of view. Case in point would be if Law Firm ZZZ had a client that believed certain files have been stolen. This utility would be great to hash the computer that the ex-employee in question used to work on. Once this hash list is created then you can use PinPoint to scan systems and directories at his new company for files that match. (Randall Consulting highly recommends that you hire a professional who is certified in computer forensics to

handle this. If that is not possible then have the most experienced person in forensics at your firm handle this task.)

Say for instance ZZZ is a pretty tech savvy law firm with experienced litigation support personnel on staff. They want to have more control over the data that goes out to a service provider. By eliminating known file types on the front end you have a better idea of what has to be converted. You also have a better idea of the size of the project before it goes out the door. This allows the service provider to only produce what can in fact be tiffed or produced as native. It turns into a win win situation for both parties. (It is imperative to have a great understanding of this program before doing this) (If it is an e-mail store it is not possible to look for binary files inside the e-mail store unless it has been ripped using an electronic discovery processing program.)

Another area where PinPoint would be an invaluable tool is to give law firm ZZZ the ability to look for a single or a set of files from separate productions. Let us say that this law firm has a list of hashes for all privileged documents in a case they are working on. By running PinPoint either before giving the data to a service provider or after you can isolate those privileged documents. This is just one way of being able to manage your productions with this tool. (If they are email stores then the hashing of files would have to come after the data is back.)

If law firm ZZZ also does small scale electronic processing in-house this tool becomes even more valuable. In this paper are but a few ways to be able to use PinPoint. Companies can get creative with how it manages its productions.

PinPoint has many different uses from a standalone de-duplication utility to filtering different native production sets. This straightforward utility benefits those in computer forensics, litigation support and electronic discovery. It truly is a tool that service providers, law firms, government agencies and law firms can use and make their life simpler.

Randall Consulting highly recommends this program. Especially with computer forensics and litigation support coming closer together there will be more needs for this program for law firms and service providers. PinPoint is not meant as a replacement for any electronic discovery processing program. It is in this author's opinion another option in managing electronic discovery.

This is also the first software program that has been reviewed that will be included in the 2008 Electronic Discovery ToolKit book.

Overall Grade **9.9**

Here are a few features PinPoint Labs are working on for the next version:

1) **Implement PG SafeCopy - Eliminates the need for Robocopy and 25% faster copy speed**
2) **Electronic Chain of Custody**
3) **Increased performance – Double the deNISTing speed and 10x speed improvement for file compare/matching**
4) **Runs from a USB drive or CD (no installation required)**

NAME	DESCRIPTION
PG Pinpoint	PG Pinpoint is a software program for 'DeNISTing' custodial files. PG Pinpoint reduces the amount of time required to process electronic files from a custodians hard drive by identifying and allowing users to 'ignore' or filter 'known' file types. The NIST (National Institute of Standards and Technology) maintains a hash list of more than 34 million known files that are associated with different operating systems and applications. It is safe and saves time to eliminate these files from electronic discovery projects based on known files in the NIST list. PG Pinpoint can scan a custodian drive in approximately 20-30 minutes and identify all files from the NIST list. A DeNISTed file list is created and the resulting files can be copied for processing. There is an average of 30-50% reduction in the overall file collection which saves several hours in electronic discovery processing per custodian hard drive.
Pinpoint MetaViewer **Free Utility**	Pinpoint MetaViewer allows users to quickly view file system metadata, OLE metadata and hash values for Microsoft Office Files. Pinpoint Metaviewer is a right-click send-to utility that places the power of viewing metadata and hash values inside Windows Explorer. Once the information is retrieved users can paste all or selected fields into any application.
Pinpoint Hash **Free Utility**	Forensic examiners need to quickly obtain the hash values for potential evidence files for reports or to verify their results. Pinpoint Hash was created to allow users to quickly obtain the hash values for CRC-32, MD5, SHA-1 and SHA-256 for any file and quickly copy the results to the clipboard for easy transport to any other

	application.
Pinpoint FileMatch **Free Utility**	FileMatch scans for duplicates of a specified file in ultra-rapid fashion. In a recent test, FileMatch was able to locate two copies of a file in 58Gb of allocated space in just 28 seconds! For litigation support professionals, law enforcement and computer forensic examiners this means that locating suspect files from custodians is only moments away!
Pinpoint SafeCopy **Free Utility**	Safecopy is a much-anticipated graphic user interface (GUI) that sits atop Microsoft's popular Robocopy utility. Safecopy is easy to use for everyone. SafeCopy is commonly used in place of Windows copy command to safely copy custodial files for electronic production requests. Whether you need to keep a file's timestamps intact during copy, you need full-scale backups, or you need to remove the security settings from files as they're copied so that they can be moved into a production environment, Safecopy is your solution.
Pinpoint MetaScrubber **(Released June, 2007)**	MetaScrubber is an advanced Microsoft Office metadata viewer, editor and scrubber. MetaScrubber displays 50 metadata fields, 13 hash values, 4 checksums and last 10 authors and locations. MetaScrubber also provides fast metadata scrubbing for several common fields and last 10 authors and locations. MetaScrubber can scrub 500 Microsoft Office files in about 5 minutes. As an option MetaScrubber will create a safe copy of the original file before scrubbing and maintain and electronic chain of custody through the entire process.

PinPoint labs from Pivotal Guidance has these free tools available via their website.
http://www.pinpointlabs.com/freetools

Randall Consulting highly recommends that any service providers offering electronic discovery as a service and unfamiliar with Robocopy should have a copy of Safecopy. The reason is Robocopy is not an

easy program to use or understand. Safecopy makes using Robocopy so much easier. To get Safecopy to work a user would need to download Microsoft's Robocopy utility and then installed the 2003 Resource kit. Copy the robocopy.exe executable to your /WINDOWS/System32 directory so it can be used running from the shell of SafeCopy.

For more information on Pivotal Guidance please visit their website: http://www.pivotalguidance.com

For more information on PinPoint labs please visit their website: http://www.pinpointlabs.com

Randall Consulting is not affiliated with PinPoint labs or Pivotal Guidance in any way nor has any compensation been given for this article. It has been written to help the litigation support community have a better understanding of software programs and utilities that cover the electronic discovery lifecycle.

This article was written by John Randall who has over 7 years experience working in the litigation support community. He has written numerous articles on electronic discovery. The Randall Report which he wrote in December 2006 comparing and contrasting third party electronic discovery programs is still available for purchase.

http://www.randallconsulting.net

Any questions about this paper can be sent to jrandall@randallconsulting.net