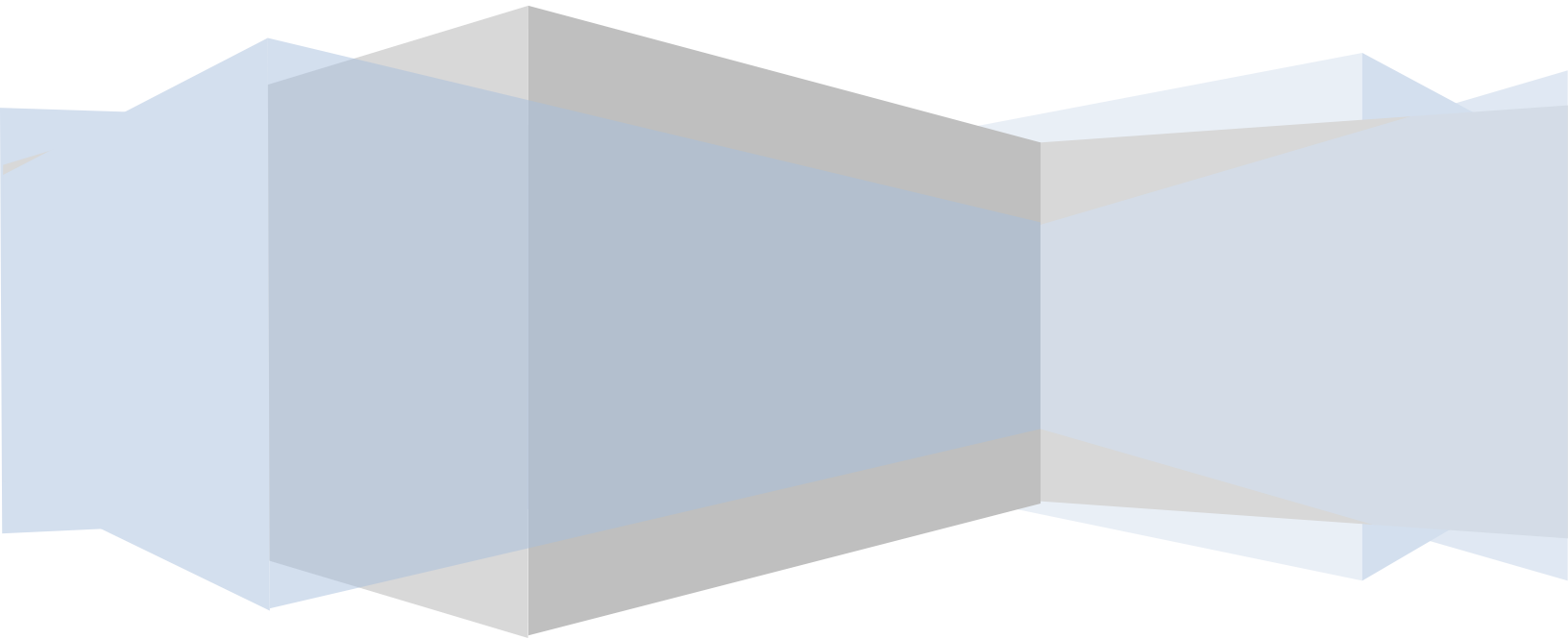
DS Forensics, Inc.

Validation Testing of Titan Electronic Discovery Tool

A Review of the Import and Export Functions and the Preservation of Metadata

David P. Stenhouse



Executive Summary

MicroForensics' "Titan" will help computer forensics examiners and corporate IT departments worldwide collect data quickly and accurately. Collections of electronic documents are paramount in today's electronic discovery process, and the tools to perform this important task are becoming too numerous to allow examiners to keep track of the best for their particular project.

It is paramount to preserve metadata associated with electronic documents in today's e-discovery arena. Clients' request of dates, times, authorship, and company information, and the court's expectation of accurate information provided by computer forensics examiners, requires proper tools and techniques to be used during processing and examinations.

Internal testing of a tool by the forensics firm utilizing it should be a common occurrence, as software tools are well-known to have "bugs" and "fixes" and do not always perform as advertised. Titan was tested using techniques a computer forensics examiner would use to export out data and use for further production to a client. I decided to perform the following tests as if Titan accuracy was to be challenged, and I had to defend the use of the tool in an examination and production of electronic evidence.

The testing of this tool proved to me Titan's ability to collect and preserve quickly and efficiently. This paper is the result of testing Titan's ability to identify, collect, and export files in the heat of an electronic discovery collection, while maintaining proper documentation and verification of each file.

Introduction

Titan, produced by MicroForensics, is a tool for use in the electronic discovery arena, but can be utilized by forensic examiners in a number of different environments. The tool's primary purpose is the location and capture of specified documents and document types, then extraction of those located files to a number of different options, including removable media in full native directory structure, a file "container", much like a logical image file[1], and/or movement to Symantec's Enterprise Vault[2].

Electronic discovery vendors, pure computer forensics firms, traditional consulting companies, and corporate IT departments worldwide are jumping on the bandwagon of mass collections of electronic documents for further review by their clients. The collections can be a daunting task for the firms that do not have the proper tools and procedures in place to capture data from individual hard drives and network shares. In corporate environments where looming litigation requires an IT staff to quickly collect electronic documents from custodians' shares on the network, processes such as using copy methods are common in the collection of live data. These processes, although allowing a staff to quickly respond to requests by their counsel, can alter information about the documents that were collected. This metadata, relied upon by attorneys in the accurate accounting of documents, if changed, may be changed forever.

Titan is designed to allow the personnel performing the collection to locate and select documents based on a number of user-selected criteria. The tool then scans the selected node for the specified file types, displaying the files in the Titan user interface by native directory structure, dates, or file extension.

I focused on the environment in which a forensics examiner would use Titan: the on-site capture of data from a mapped drive, such as a network drive from a client machine, and the examination of a logical image file, such as one created by an IT department and transmitted to the examiner for forensics review.

I created Logical Image of data from a server containing thousands of user-created files and imported the image into Titan for testing purposes. I used the same logical image for each test.

[1] A file created by the forensic software that gathers specified files from their original directory structure and "seals" the contents in a file for further review. Logical evidence files are useful to forensics examiners who do not need to capture the entire physical disk, but just specific files and/or directories.

[2] http://www.symantec.com/business/products/overview.jsp?pcid=2242&pvid=322_1.

Testing Methodology

I decided to use two (2) forensics suites of tools for verification of Titan's file reports and extracted data:

1. **EnCase** – a forensics suite of tools created by Guidance Software (www.guidancesoftware.com), widely used worldwide by law enforcement agencies and private computer forensics examiners. I have used EnCase since v. 1.99, and have found the tool extremely powerful in sorting files and providing numbers for verification purposes;

2. **Forensics Tool Kit (FTK)** – created by AccessData Corporation (www.accessdata.com), FTK is a forensics suite of tools also used by law enforcement agencies and private computer forensics examiners worldwide. FTK is extremely helpful in identifying files not easily located, such as those within compressed files, such as .zip, and password-protected files. FTK also allows ease in the creation of MD5 hash sets with a given list of MD5 hash values.

A Note About "File Safe" Files

Titan can produce a logical image of selected files appropriately named a "File Safe". This type of file is similar to the function of a logical image file created by comparable forensic tools, and allows the computer forensics examiner to store a specified set of files for preservation and further review, without altering the metadata associated with the files within the file safe. The file can be opened with Titan at a later date, with all data preserved as it was when collected.

When I first wanted to test Titan and see its forensics capabilities and accuracies, I concentrated on getting answers to the following questions:

Does the software "see" what it is supposed to see? – A forensics examiner wants to capture information from a computer system for further analysis, or to provide that data to his/her client for further review. When the software tool locates files specified to be found, by the examiner, the examiner wants to be assured after extracting those files, no further copies of files remain to be located. I mounted a logical evidence file, and then a local folder, with Titan to review the listing of files and their associated file extensions. I then used EnCase and FTK to verify the two forensics suites located the same files.

During evidence import, does the software import specified evidence correctly? – The import of evidence through a logical file or mapped drive is a common procedure with forensics collections in the E-discovery arena. Titan has a feature that allows the examiner to filter the inbound evidence, either from a logical image or mapped drive, by file extension and/or created, modified, and last accessed dates (see Figure 1). This feature also has pre-installed file type "selection sets" that allows the examiner to pick word processing, email, and archive files, etc. to filter into the examination. Each selection set can be modified and saved, or new selection sets can be created according to the examiner's preferences.

After locating a set of files I sought for collection, with the use of EnCase and FTK, I imported those same files into Titan for further review. I then checked the metadata for individual files to verify nothing had been changed during the import.

On file extractions for eventual review by the client, does the software extract everything it is supposed to extract? – Once the forensics examiner locates the files sought and executes a file extraction to a separate location, such as a removable hard drive, does Titan capture all the selected files and copy them to the new location? This process is completed by the examiner using Titan to send the selected files out to a FileSafe or native export[3]. The examiner expects files selected for further review, to be exported in full, without corruption. I used Titan to make a selection in the "cabinet" of the software tool, and then created a folder on

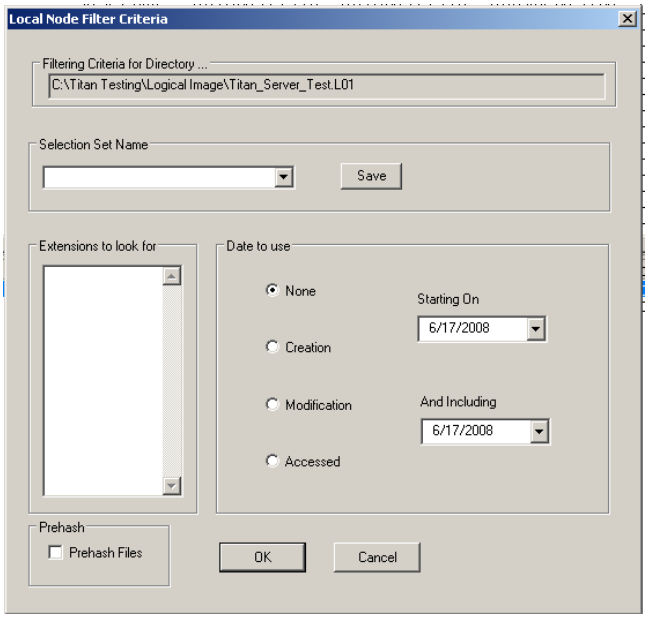
[3] The same directory structures as the files originally were located. For example, a file originally located in the user's My Documents folder would be exported by Titan to a My Documents folder for review.

a separate hard drive as a target location for the "native" export. Upon execution of the exportation of files, I used EnCase and FTK for verification purposes of the number of files exported, by extension.

For a FileSafe file, I performed the export directly out to the FileSafe option, then loaded the file back into Titan to verify file counts did verify, and then exported all files out of the file safe file to a native export. The resulting directory structure was reviewed with EnCase and FTK to verify the correct numbers of files.

During the extraction process, does the software alter any files or file metadata? – Full confidence in a software tool's integrity in the handling of files, and to not alter the data associated with the files, is primary in the examiner's mind when extracting data for further review. Integrity of each file extracted can be assured by creating an MD5 Hash of each from the original location, and then checking each file's MD5 hash at the new location. Native exports out of Titan will definitely alter the last accessed date of the exported file; however does not alter the contents of the file itself, preserving the data sought by the client for review. This can be seen through the use of MD5 analysis and verifying the hash value is the same once the files are exported to the new location.

Figure 1 – Local Node Filtering Options



Titan also includes the ability to track each file exported and records of any errors that occurred during the export by creating a listing of each file exported. This listing includes the MD5 hash value for each file exported for verification purposes and the original file path. This is a great tool to use to provide clients with a listing of files to review, and for the forensics examiner to verify his/her results.

I created a logical evidence file from an EnCase image file and imported as evidence into the Titan user interface. I then selected 50 files for export to a native export. Each file was extracted to the selected formats, and then the hash values read through the file listing provided by Titan, and then verification through EnCase and FTK.

Testing Results

Does the software "see" what it is supposed to see?

Using an EnCase logical image file I created from random data, and then importing the image into Titan, I performed file counts by extension to verify Titan's ability to see each file record that EnCase captured in the creation of the logical image of data.

Table 1 gives results in the number of file records, by a sample of extensions I chose for testing purposes, located in the logical image as viewed by EnCase, and then viewed by Titan. FTK was not used for verification purposes, as the FTK software will not mount an EnCase logical image file.

Table 1 – File Count by Extension in Comparison with EnCase (Logical Image)

File Type By Extension	EnCase View	Titan View
.doc	3,123	3,123
.ppt	20	20
.xls	508	508
.txt	390	395
.htm	5,940	5,940
.html	1,391	1,391

Using a local folder[4] on a DSF machine that I created by random data collected, I used EnCase and FTK to review the contents of the folder, and to make file counts according with the same file extensions. I then imported the same folder into Titan to record file counts. Table 2 displays the results of the review of the local folder.

Table 2 – File Counts by Extension in Comparison with EnCase and FTK (Local Folder)

File Type By Extension	EnCase View	FTK View	Titan View
.doc	119	119	119
.ppt	132	132	132
.xls	49	49	49
.txt	45	45	45
.htm	268	268	268
.html	17	17	17

[4] A folder of subfolders and files located on a computer, and not a logical image file. This would simulate the capture and extraction of data on a computer network or user's local machine.

It should be noted that contents of compound files[5] were not included in these results, but individual compound files were included. Titan, by default, is not designed to "dig" into compound files and locate sought-after data, but is designed as a "first-pass" capture of selected file types.

During evidence import, does the software import specified evidence correctly?

Using a local folder created by random data as a sample, I viewed and selected 50 individual files, recording each file's metadata. This was done using EnCase first and then FTK to verify each software tool's ability to record identical metadata in the same files. I then imported the local folder into Titan and reviewed the same selected files within the Titan user interface. Each of the 50 individual files was reviewed for creation and modified dates, and file size, to verify the metadata associated with each file imported. The results are listed in Table 3.

Table 3 – Metadata Changes upon Import into Titan

File Type By Extension	Number of Files Viewed in EnCase	Changes in Metadata Viewed in FTK	Changes in Metadata Viewed in Titan
.doc	10	0	0
.ppt	10	0	0
.xls	10	0	0
.txt	10	0	0
.htm	10	0	0

On file extractions for eventual review by client, does the software extract everything it is supposed to extract?

Taking a sample of random data from an EnCase logical file, I imported that data into Titan, and extracted the same data out to a native export. I then checked the numbers of files exported to the new location with EnCase and FTK for verification of exact export as selected. The result of the files selected and the numbers exported is displayed in Table 4.

Table 4 – File Counts upon Export from Titan

File Type By Extension	Files Selected for Export from Titan	EnCase View of Exported Files	FTK View of Exported Files
.doc	1,281	1,281	1,279*
.ppt	1	1	1
.xls	72	72	72
.txt	24	24	24
.htm	384	384	384
.html	8	8	8
**FTK would not import the two (2) missing documents (.doc) into its database. This was not a Titan issue, but an FTK issue.*			

[5] A file constructed of multiple components, such as MS Office documents, MS Outlook .pst files, and WinZip compressed files. A compound file may contain multiple files within, such as displayed graphics, or attachments to specific email messages.

During the extraction process, does the software alter any files or file metadata?

I imported a logical image into the Titan software and then selected 50 files for export out to a native format. The hash values for the native export were checked in accordance with the values recorded in EnCase prior to the creation of the Logical image file. The results of this analysis are displayed in Table 5.

Table 5 – Hash Value Changes in Extraction from Titan

File Type By Extension	Hash Value Changes in Titan Export Log	Hash Value Changes in Native Export
.doc	0	0
.ppt	0	0
.xls	0	0
.txt	0	0
.htm	0	0

The exported files were also reviewed by EnCase and FTK to verify no changes were made to the creation, modified, or last accessed dates of each file. The results of this analysis are displayed in Table 6.

Table 6 – Metadata Changes in Extraction from Titan

File Type By Extension	Metadata Changes in Titan Export Log	Metadata Changes in Native Export
.doc	0	0
.ppt	0	0
.xls	0	0
.txt	0	0
.htm	0	0

Conclusions

Titan uses a sound forensic process to locate, verify, and extract files for the forensic examiner. I found Titan's ability to locate files based on pre-installed and user-input criteria to be accurate. Files located by Titan, collected, then exported suffered no change in metadata, by verification of spot checks and hash value analysis.

My testing was limited to how a forensic examiner would use Titan in small, highly-intensive analysis cases, as opposed to the collection of mass amounts of files. However, the results should be the same as the collection of evidence and chain of custody will not change based on amounts of data retrieved. Titan does an excellent job in the sound, forensic collection and extraction of data.

About the Author

David Stenhouse is the President of DS Forensics, Inc. (DSF). Mr. Stenhouse brings 10 years electronic discovery and computer forensics experience in Federal Law Enforcement computer crimes investigation and civil litigation. Prior to founding DSF, he provided consulting and expert witness services to law firms and corporate clients with a mid-sized consulting company, and also a small computer forensics company, both based in Seattle, Washington.

Mr. Stenhouse is also a forensic examiner, and has performed hundreds of forensic examinations on multiple types of hardware and operating systems, in criminal cases and civil litigation. He routinely provides expert guidance and training to attorneys and corporate clients faced with the task of electronic discovery. He also provides training to law enforcement in forensic procedures and evidence handling.

Mr. Stenhouse is a former Special Agent in the United States Secret Service, where he conducted investigations involving the use of electronic data in crimes and he executed federal search warrants to perform seizures of numerous computer systems, and completed in excess of 80 individual forensic examinations of the contents of those systems.

Mr. Stenhouse received his computer forensics training from the Federal Law Enforcement Training Center (FLETC) in Glynco, GA as a part of the Treasury Department's CIS 2000 program. He has performed computer examinations for local, state and federal agencies including the ATF, Royal Canadian Mounted Police and the Vancouver, B.C. Police Department Organized Crime Unit. He has extensive experience in the criminal and civil court process and evidentiary chain-of-custody issues. Mr. Stenhouse has testified in state and federal court in numerous criminal and civil cases, and has testified in federal court as an expert witness in computer-generated evidence. Prior to his employment with the Secret Service, Mr. Stenhouse was a State Trooper with the Washington State Patrol.