

UNIVERSITY OF CENTRAL OKLAHOMA

Edmond, Oklahoma

Jackson College of Graduate Studies

**THE UTILIZATION OF FORENSIC CORPORA IN VALIDATION OF
DATA CARVING ON SATA DRIVES**

Submitted by

Caitlin G. Willimon

Forensic Science Institute

In partial fulfillment of the requirements

For the Degree of Master of Science in Forensic Science

2019

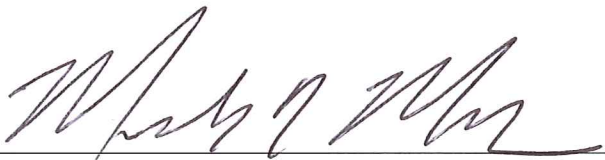
**THE UTILIZATION OF FORENSIC CORPORA IN VALIDATION OF
DATA CARVING ON SATA DRIVES**

By: Caitlin Greer Willimon


A THESIS

APPROVED FOR THE W. ROGER WEBB FORENSIC SCIENCE INSTITUTE

2019

By 
Dr. Mark R. McCoy Committee Chair

By 
Dr. John Mabry Committee Member

By 
Rachael Elliott Committee Member

Acknowledgements

I would like to start out thanking my advisor, Dr. Mark McCoy, for believing in me and being there when I had any questions or concerns. Thank you for checking in when I became absent and reaffirming my belief that I had it in me to see this through to the end.

Without the support of my advisors, Dr. John Mabry and Rachael Elliott, as well as the staff of UCO's Forensic Science Institute this study would not have come to fruition. A large thanks to Professor Craig Gravel and Dr. James Creecy for taking time to stop and talk to me while I was working on my research and help ease my mind when I was stressing about it.

Finally, the completion of this study would not have been possible without the support I received from my friends and family members. My parents, John Ed and Kirie Willimon helped me manage my stress and refused to give up on me. Without my friends, Caitlyn and Andreas constantly checking in on my progress and believing in me when I didn't believe in myself, I would have lost sight of the light at the end of the tunnel.

Abstract

The field of digital forensics has become more prevalent in the court of law due to the increase of availability of technology. With digital evidence coming up in court consistently, digital forensics and its tools are coming under scrutiny and being held against disciplines that are more standardized. Wildson and Slay went so far as to call digital forensics the “neglected family member of the forensic sciences” when discussing the lack of standards in the discipline (Wildson & Slay, 2005, p. 2). In order to begin addressing this issue, we must start looking at the source - the tools. Validation and Verification of tools is vital to maintaining the integrity of the evidence received by them. Utilizing standardized data sets, or forensic corpora, as a part of validation and verification techniques has shown to be effective. This study will replicate Hegstrom (2016). This study will focus on validating the file carving function of Access Data’s Forensic Tool Kit (FTK) using forensic corpora on SATA drives instead of the Solid-State Drives (SSDs) that Hegstrom used. The goal of the study is to assess the use of forensic corpora in the validation and verification of one of the most commonly used digital tools.

**THE UTILIZATION OF FORENSIC CORPORA IN VALIDATION OF DATA
CARVING ON SATA DRIVES**

Acknowledgements.....3
Abstract.....4
Table of Contents.....5
List of Tables.....7
List of Figures.....8
List of Appendices.....9
Chapter 1: Introduction.....10
 Current Issues.....12
 Significance to the Field.....13
Chapter 2: Literature Review.....15
 Validation & Verification.....15
 Validation & Verification Approaches.....15
 Forensic Corpora.....16
 Issues With Corpora.....17
 Definition of Terms.....17
 Limitations.....17
Chapter 3: Methodology.....20
 Setting.....23
 Sample.....23
 Materials.....23
 Expected Results.....24
Chapter 4: Results.....25
 Results.....25
 Level.....26
 Type.....27
Chapter 5: Discussion.....33
 Discussion.....33
 Limitations.....34

Recommendations for Future Research.....	34
Conclusion.....	35
References.....	37

List of Tables

Table 1: <i>CFTT Test Image Data Levels and Descriptions</i>	22
Table 2: <i>CFTT Test Image Types and File Types</i>	22

List of Figures

Figure 1: <i>Percentages of files carved correctly by file level.</i>	26
Figure 2: <i>File Type v. Percentage of Files Carved</i>	27
Figure 3: <i>Document File Type v. Percentages of Files Carved</i>	28
Figure 4: <i>Graphic Files v. Percentages of Files Carved</i>	29
Figure 5 <i>Video Files v. Percentages of Files Carved</i>	29
Figure 6: <i>Archive Files v. Percentages of Files Carved</i>	30
Figure 7: <i>Audio Files v. Percentages of Files Carved</i>	31
Figure 8: <i>Hegstrom v. Willimon Correctly Carved Files</i>	32

List of Appendices

Appendix A: File Carving Test Images	37
Appendix B: Recorded File Carving Results	39

CHAPTER 1

Introduction

The field of Forensic Science is under constant scrutiny and the sub-field of Digital Forensics is no different. As the newest addition to the Forensic Science family, it is the farthest behind when it comes to validation testing. With the constant technological advances, it is important for digital analysts to have an accurate validation technique to ensure their programs are looking for deleted files in the correct place. In September of 2016, the President's Council of Advisors on Science and Technology (PCAST) gave recommendations that would not only strengthen the field but would also aide examiners in the courtroom. PCAST's report titled, "Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods" discussed scientific validity in the courtroom and touched upon the topic of certain disciplines no longer being allowed in court. The main recommendation was increasing validation techniques of other disciplines. This has been a common subject within Forensic Science. Similar reports recommended reform for the scientific standards utilized in the Forensic Science community.

Digital Forensics was recognized as an established discipline in the 1980s after the personal computer became accessible. This makes Digital Forensics the youngest of the disciplines and brings a special set of problems not faced by other disciplines. The field began solely focusing on computer related crimes but has evolved alongside technology to include a wide variety of digital devices. As technology has become an integral part of our everyday lives, it is fair to say that the complexity and number of digital devices has evolved with time.

Many Forensic disciplines utilize comparative analysis with access to a "known" sample to use as a comparison during their investigation. For example, drug analysis has standardized

sets of data that give information on different types of drugs. This rarely applies to a Digital Forensic examiner. This is mainly due to the amount of time that it would take to create data sets on each type of file in existence. Another challenge is the ever-evolving aspect of technology. Digital Forensic examiners typically rely on forensic tools during their evidence recovery. Digital devices have the capacity to hold various file types. This leaves the examiners with copious amounts of files to sort through in order to find ones relevant to the case.

Examiners in this discipline utilize Digital Forensic tools in order to locate certain types of data based on characteristics such as file type and size. These tools also assist examiners in the task of recovering deleted and corrupted files. Understanding the role of these tools is important for examiners as well as attorneys, judges and jurors. Digital Forensic tools are used throughout the entire analysis process. The tools are used to generate forensic images which will appear in the examiner's report and could be utilized in the courtroom. Guo and Slay stated a challenge of the discipline "is to ensure that the digital evidence acquired and analyzed by investigative tools is forensically sound" (Guo & Slay, 2010, p 297). Due to the necessity of these tools, their functionality needs to be confirmed. The tools are tested by their manufacturer, however, the data sets they use to validate the programs are not available to the public.

To ensure the functionality and reliability of these tools, research on validation and verification techniques is required. The tools that are most common must be tested to confirm that they are performing a true forensic analysis. Examiners need to be sure that the tools they are utilizing are accurate and performing the task they are designed to do. Studies, such as these have proven difficult due to lack of monetary resources as well as time. The constant development in technology in general also adds to the difficulty due to the Digital Forensic community constantly having to play catch up. Due to the fluid nature of technology as well as

the accessibility, research in Digital Forensic is more than necessary. This research must focus on tool performance on different types of technology, new and old, to ensure a reliable examination process during forensic analysis. It also much focus on different file types and levels of corruption to fully test the lengths of the tool.

Current Issues

As previously stated, tools are a major part of the digital analysis process. Garfinkel stated that “many in leadership positions now rely on these tools on a regular basis – frequently without realizing it” (2010, p S64). While these tools are a helpful aide in the analysis process, the reliance on the tools and their accuracy is cause for worry. Many companies do their own validation and verification testing of their products. This can take some of the timely burden from those using the tool, however; the data sets used in verification are not available. This means that the purchaser cannot use them to verify the program on their computer. This puts a large amount of trust in the manufacturer however, leaves the examiner with little to show when it comes to verifying the results that the tool produces. This increases the already tough scrutiny of digital forensics in courtroom proceedings.

An issue that arises with the constant change in technology is updates. Once an update is performed on the tool software or the computer system, the tool must be re-validated. If the tool is not re-validated prior to being used in an analysis, the results may not be usable in a case. According to SWGDE, not using proper validation techniques can have “detrimental effects” (p. 4). The analyst would not be able to prove that the program is as accurate as it previously was. This can be a costly and timely task, especially if standardized data sets are not readily available to be tested by the purchasers.

Significance

Due to the amount of time in the field, the area of digital forensics does not have the amount of research that its fellow forensic disciplines have. With technology becoming a major part of day-to-day life, the importance of research in this area is paramount. It has been widely agreed upon by researchers and examiners in the field that the lack of standardization in digital forensics is one of the main issues holding the field back from advancing. The limited availability of standardized data sets, formally known as digital corpora, makes the research and standardization process more difficult. If digital corpora were widely available, validation techniques could be used to ensure the tools used by examiners were performing the tasks effectively. Garfinkel et al states that “the development of representative standardized corpora for digital forensics research is essential for the long-term health and legal standing of the field” (p. S10). If this validation technique is found to be successful, it could be tested on other programs to test its efficiency on multiple platforms.

With the amount of technology that is present within crimes in this day and age, it is extremely important that we ensure that the programs that we rely upon are doing their job properly. Without proper validation, the test results can be seen as less trustworthy or valid especially when compared to other forensic disciplines. With Digital Forensics under a microscope, it is critical that analysts are able to produce viable and accurate examinations of evidence.

Previous studies have used this methodology on solid-state drive (SSDs) in which the program was not 100% successful. Technology is leaning towards solid-state drives due to their efficiency and speed. Although technology is shifting towards solid-state drives (SSDs), SATA

drives are not considered obsolete. Many users will still be utilizing the older models of hard drives so, testing using this method on SATA drives is needed.

CHAPTER 2

Literature Review

Validation and Verification

Validation is defined by Wilsdon and Slay (2005) as “the process of evaluating a system or component during or at the end of the development process to determine whether it satisfies requirements” (p. 2). In short, validation is a process used to determine whether a tool is performing the tasks assigned properly. The Scientific Working Group on Digital Evidence (SWGDE) states that validation “is required to demonstrate that examination tools (hardware and software), techniques and procedures are suitable for their intended purpose” (p. 4). While validation deals with the performance of the tool, verification evaluates whether the tool was properly built.

Validation and Verification Approaches

The Scientific Working Group on Digital Evidence (SWGDE) defines validation testing as “an evaluation to determine if a tool, technique or procedure functions correctly and as intended” (p. 4). These techniques are used to keep the integrity of the evidence found by the tools. Guo et al. (2009) reports that much of the research done on validation and verification (VV) shows the results produced by the tools, not the performance of the tools themselves. They discussed two techniques for VV including a tool-oriented approach and the functionality-oriented VV. The tool-oriented approach tests each individual aspect of a tool instead of the tool in its entirety. While this is an important part, the tools tend to operate in a package format and it is important to test how they operate as a whole. The functionality approach looks at the tool as a whole. This seems to be the most cost-effective approach.

Forensic Corpora

Corpora refers to a standardized data set with known characteristics. Forensic corpora introduces this into testing within the forensic spectrum of disciplines, including digital forensics. While forms of corpora are present in many other disciplines, it is rarely used in the digital realm. The implementation of such corpora could be extremely helpful in the task of validation and verification of forensic tools. Access to digital corpora would provide “a basis for comparing methodologies and tools so that the advantages and shortcomings can be identified” (Yannikos, et al, p. 310). By giving researchers access to data sets that they can use to validate their tool, it increases the validity of the results produced. In order to do this, forensic (digital) corpora is a must.

An issue with the lack of forensic (digital) corpora is reproducibility. This means that researchers in different locations or with different computer systems currently do not have the ability to conduct the same research in their respective locations to compare results. Tools have been researched by different researchers, but results are varied due to the different data sets used between them which renders them incomparable (Garfinkel, et al, p. 4). There are few sets of data corpora that have been made available for public use. These include: Real Data Corpus, DARPA Intrusion Detection Data Sets, Global Intelligence Files, Computer Forensic Reference Data Sets, MemCorp Corpus, MORPH Corpus, and Enron Corpus (Yannikos et al, p. 310-313). They all offer valuable data, but the majority of the sets are either focused on one type of file, produced for training purposes, or require permission. The Computer Forensic Reference Data Sets seem to be the best choice due to its variety of file types and accessibility although it is small compared to the standardized corpora available for other forensic disciplines.

Issues with Corpora

The main issue that is present with forensic corpora is building the data sets. This is a very time-consuming task and brings with it several issues. In order to have a “complete” set of corpora, there would be an immense amount of file types to be assessed. Compiling such file types brings forth its own problems such as, where are these files coming from? This task will have to be an effort spread amongst the digital forensics community. Access to standardized data sets would help “increase the efficiency and the quality of forensic examinations while allowing objective evaluations by third parties” (Yannikos, et al, p.310).

Limitations

Due to this study focusing on one of many tools utilized in digital forensics, the results cannot be generalized for other tools. While this study is focused on utilizing corpora in the validation and verification process of digital tools, the results will likely vary from tool to tool. The procedures could be replicated on other tools to test their validity in future studies.

The next limitation is the amount of data samples that will be used. Due to the limited amount of corpora that is available, it is difficult to predict how this process will adapt to various file types. Even if we had access to copious amounts of corpora, predicting how the process would work with various conditions is difficult.

Finally, due to the specific conditions that this study will be performed in, it should be easily replicable in another lab with similar conditions. However, other factors including software and hardware conditions/configurations could lead to various results.

Definition of Terms

- Computer Forensics Tool Testing Program (CFTT)- A joint project of the Department of Homeland Security (DHS), the National Institute of Justice (NIJ), and the National

Institute of Standards and Technology (NIST) Law Enforcement Standards Office (OLES) and Information Technology Laboratory (ITL), whose objective is to provide measurable assurance to practitioners, researchers, and other applicable users that the tools used in computer forensics investigations provide accurate results.

- Contiguous File- A file on disk that is not broken apart. All sectors are adjacent to each other.
- FAT32- (File Allocation Table32) The 32-bit version of the FAT file system, widely used for USB drives, flash memory cards, and external hard drives for compatibility on all platforms.
- File Carving- The process of extracting a collection of data from a larger data set. Data carving techniques frequently occur during a digital investigation when the unallocated file system space is analyzed to extract files. The files are "carved" from the unallocated space using file type-specific header and footer values.
- File Extension- A file type that is appended to the end of a file name (ex. .DOC, .JPEG).
- Forensic Corpora/digital corpora- A standardized, representative reference data set; can contain various file types, file sizes, and file systems.
- Forensic Image- A bit stream copy of the available data. The result may be encapsulated in a proprietary format (e.g., E01, 001, etc.).
- Forensic Wipe- Completely erasing the data in disk sectors.
- Fragmented File- Storing data in non-contiguous areas on a disk.
- Hash or Hash Value- Numerical values, generated by hashing functions, used to substantiate the integrity of digital evidence and/or for inclusion /exclusion comparisons against known value sets.
- Hashing Function- An established mathematical calculation that generates a numerical value based on input data. This numerical value is referred to as the hash or hash value.
- National Institute of Standards and Technology (NIST)- A measurement standards laboratory, and a non-regulatory agency of the United States Department of Commerce. Its mission is to promote innovation and industrial competitiveness.
- Offset- The distance from a starting point, either the start of a file or the start of a memory address. An offset into a file is simply the character location within that file, usually starting with 0.

- Sector- The smallest unit of data that is written to and read from a storage drive.

CHAPTER 3

Methodology

This study focused on the data carving function of Access Data's Forensic Tool Kit (FTK) on SATA drives using forensic corpora on SATA drives. The following research questions were addressed in this study:

1. Does FTK's data carving function perform as intended and recover all possible data when used on SATA drives?
2. Does FTK's data carving function ability vary depending on file type?
3. What is the error rate of FTK's data carving function when used on SATA drives?

In this study, the methodology from Hegstrom's, *The Use of Forensic Corpora in Validation of Data Carving on Solid State Drives* (2016), was utilized and adapted for SATA drives. Her methodology followed the process section from the Scientific Working Group on Digital Evidence (SWGDE) Recommended Guidelines for Validation Testing, Version 2.0 (2014). NIST's CFTT project's forensic corpora consisting of graphic, document, archive, audio, and video files were used for the data sets in this study. The corpora consisted of four levels ranging from Level 0 (L0) to Level 3 (L3). Each level indicated a different status of the files (i.e. complete, partial, fragmented, etc.).

From there, the remaining steps of Hegstrom's methodology were followed, replacing SSDs (solid-state) drives for SATA Drives:

“Prior to the corpora being placed on the drives, each solid-state drive was forensically wiped, ensuring no residual data remained on the drive. Each data set was then restored to each

of three SSDs in the study and subsequently carved using FTK. FTK's carving function was directed to carve for the specific file types and recover each item within the respective image. In measuring the recovery rate, the tool was graded on a pass/fail scale, in which the recovery of the data in full from the correct location warranted a pass, and all other results warranted a fail. The occurrence of false negatives and positives was also noted during the course of the study." (p. 33).

Setting

This study was performed at two locations, using the same device in both. The research was split between the University of Central Oklahoma Digital Forensics classroom and the Digital Forensics Research Lab at the same University. A freestanding forensic duplicator was used to wipe the SATA drives. After being wiped, the same device was used to run a "blank check" to make sure the device was truly clear of any data. The drives were then plugged into a Forensic Recovery of Evidence Device (FRED) and a hash algorithm, checksum32, was ran to guarantee they were clean. The program WinHex was used to perform this algorithm. Once the checksum32 returned all zeros, confirmed there was no data on the drive, the corpora was restored to each drive using the same program. After the corpora was written to the drive, it was then imaged using a disk-imaging program and saved to be used for carving later.

Sample

Convenience sampling was used as the sampling procedure in this study. Due to the limited amount of digital corpora available, the study was limited to the CFTT project's already developed corpora. The project had developed 30 test images including five file types and various levels of structure that are available online for research purposes. For this study, 20

images including all five file types and levels zero through three. These were chosen due to type of data that examiners will encounter in digital forensic examinations. The four levels are described in Table 1:

Table 1

CFTT Test Image Data Levels and Descriptions

Level	Description	Content
Level 0	Cluster padded	contiguous files with assorted levels of content ranging in size from 1, 2, 4, 8, 16 ... 128 sectors
Level 1	Fragmented in order	contiguous and sequential fragmented files with content separating the files
Level 2	Fragmented out of order	contiguous and disordered fragmented files separated by other content
Level 3	Incomplete	contiguous and partial (i.e., only a portion of the file is present) files

A spreadsheet was provided for the individual test images to provide details including the file size, starting sector, and ending sector. The test images were assigned names with the template of FILESTYTEM_LEVEL_FILETYPE.dd.bz2. For reference, FAT_L1_Archive.dd.bz2 identifies a Level 1, Archive test image with a FAT 32 file system. This example test image would include a file allocation table (FAT) file system, files that are fragmented in order, and would contain 7Z, BZ2, GZ, TAR, WIM, RAR, and ZIP file types. Table 2 displays the image and file types for the corpora utilized in this study.

Table 2

CFTT Test Image Types and File Types

Image Type	Images	File Types
Graphic	FAT_L0_Graphic.dd.bz2 FAT_L1_Graphic.dd.bz2 FAT_L2_Graphic.dd.bz2 FAT_L3_Graphic.dd.bz2	JPG, PNG, BMP, GIF, TIF, PCX
Document	FAT_L0_Document.dd.bz2	DOC, XLS, PPT, PDF

	FAT_L1_Document.dd.bz2 FAT_L2_Document.dd.bz2 FAT_L3_Document.dd.bz2	
Archive	FAT_L0_Archive.dd.bz2 FAT_L1_Archive.dd.bz2 FAT_L2_Archive.dd.bz2 FAT_L3_Archive.dd.bz2	7Z, BZ2, GZ, TAR, WIM, RAR, ZIP
Video	FAT_L0_Video.dd.bz2 FAT_L1_Video.dd.bz2 FAT_L2_Video.dd.bz2 FAT_L3_video.dd.bz2	MP4, AVI, MOV, FLV, MPG, WMV
Audio	FAT_L0_Audio.dd.bz2 FAT_L1_Audio.dd.bz2 FAT_L2_Audio.dd.bz2 FAT_L3_Audio.dd.bz2	MP3, WAV, AU, WMA

Materials

Three 40 GB SATA hard drives were used including two Hitachi and one Western Digital. The independent variables measured in this study are level and file type. The test images included various states that a file can be in from contiguous and complete to incomplete and fragmented. They represent files that are commonly encountered in digital forensic investigations. Contiguous and complete files are very commonly found in digital examinations. Fragmented files represent the data that is commonly found when the files were intentionally deleted/destroyed. Both are fundamental in digital forensic examinations. Including various file types permits the study to compare carving results between files categories commonly encountered in the field.

FTK’s carving function as well as whether the program carved the data in question are the dependent variables measured in this study.

Expected Results

With Access Data's Forensic Toolkit (FTK) being one of the highly utilized tools in digital forensics, the carving function's reliability is assumed to be quite high. However, little testing has been conducted to back up this assumption. The expected recovery rate for contiguous files is close to 100% while for fragmented files, it is possible that the rate is significantly lower. The expectation is that this research will back up those proposed claims. Also, this research could provide useful information about the use of corpora in the verification of tools in digital forensics, leading to more studies testing various tools.

CHAPTER 4

Results

Three identical SATA drives (40GB), four levels of files (cluster padded, fragmented in order, fragmented out of order, and incomplete), and five types of images (archive, audio, document, graphic, and video) were analyzed in this experiment. A total of 684 files were processed using the data carving function of FTK. The three possible outcomes for each file were if the program carved the data correctly, carved the data but not the correct file, or did not carve the data.

Results

The analysis of the test image files focused on whether or not the files were carved correctly. Chi-square tests were performed to determine the differences in percentages of files carved correctly for the various levels and types. P-values less than 0.05 were considered significant. Those chi-square tests that were significant were then analyzed using pairwise chi-square tests to determine whether the level(s), and/or type(s) were significantly different.

Each file type was then broken down by file extension. Chi-square tests were performed to identify differences between the files correctly carved for the various file extensions.

If the sample size proved to be too small for a valid chi-square test, Fisher's exact test was performed instead. P-values which utilized Fisher's exact test instead of the chi-squared test were marked with a (F). All statistical tests were performed using SAS 9.3.

Level

This study showed that when analyzing different levels of fragmentation, FTK was more likely to correctly carve files that are either cluster padded or fragmented in order. It was more than twice as likely to carve either of those levels than it was to carve files that were fragmented out of order or incomplete. The level of significance was $p < 0.05$. Level pairwise chi-squared tests were ran, showing that there were significant differences between all levels other than 0-1 ($p = 0.7834$).

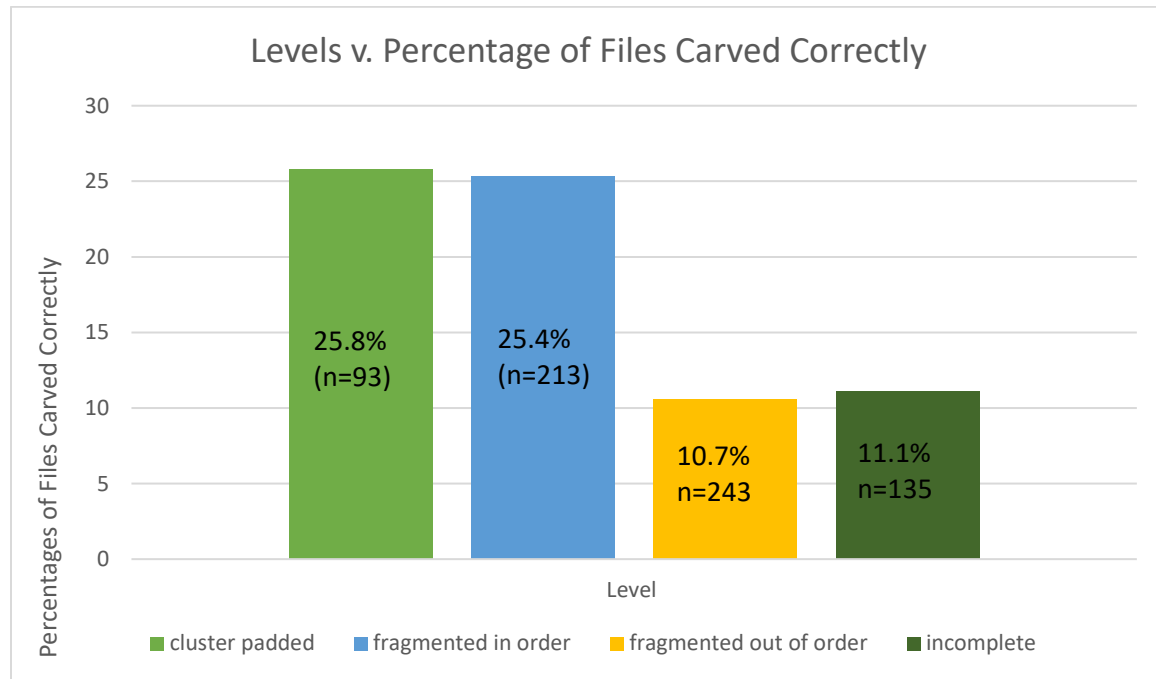


Figure 1
Percentages of files carved correctly by file level.

Type

Graphic file types had the highest percentage of correctly carved files, followed by Document file types. Archive file types had the highest percentage of files that were carved incorrectly.

Audio and Video file types were not carved. After running type pairwise chi-square or Fisher's exact tests, no significant differences were found between file types.

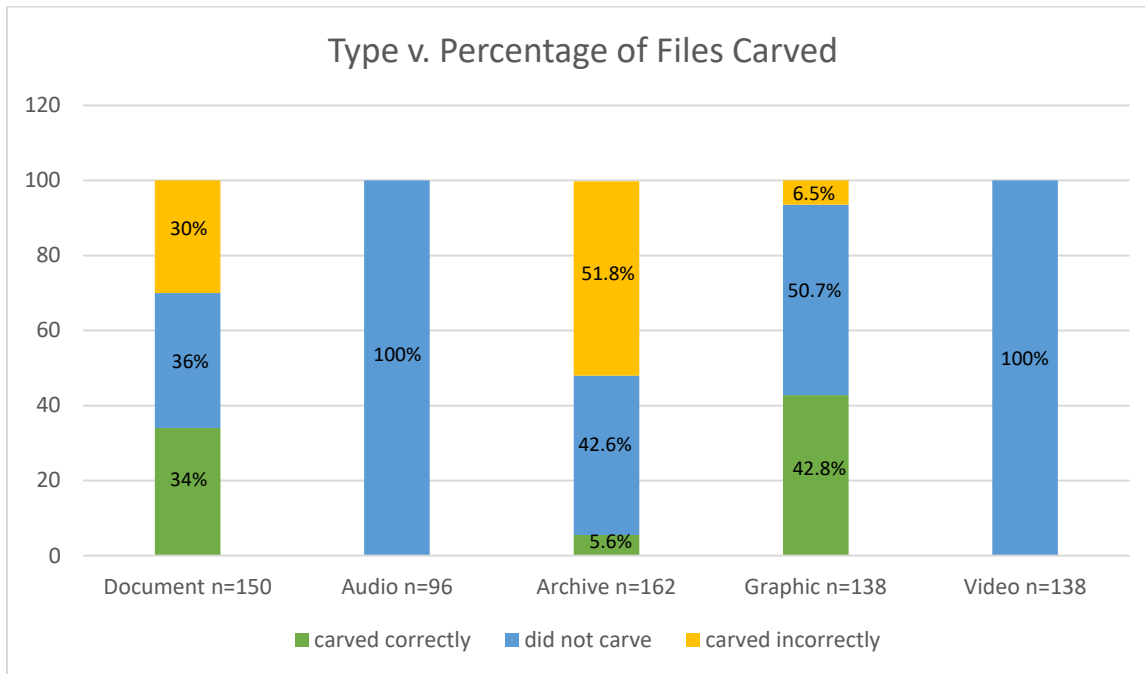


Figure 2
File type v. Percentage of Files Carved

All of the pdf files were carved correctly. The docx files were split 50/50 between carved correctly and carved incorrectly. Pptx and xlsx files were split 42.9/57.1. The xlsx files, when carved incorrectly, showed up as zip files in FTK. Docx and pptx file extensions were combined for the pairwise chi-square tests. There was no significant difference between doc_pptx and xlsx file extensions ($p=0.6559$). There was a significant difference between the remaining file extensions.

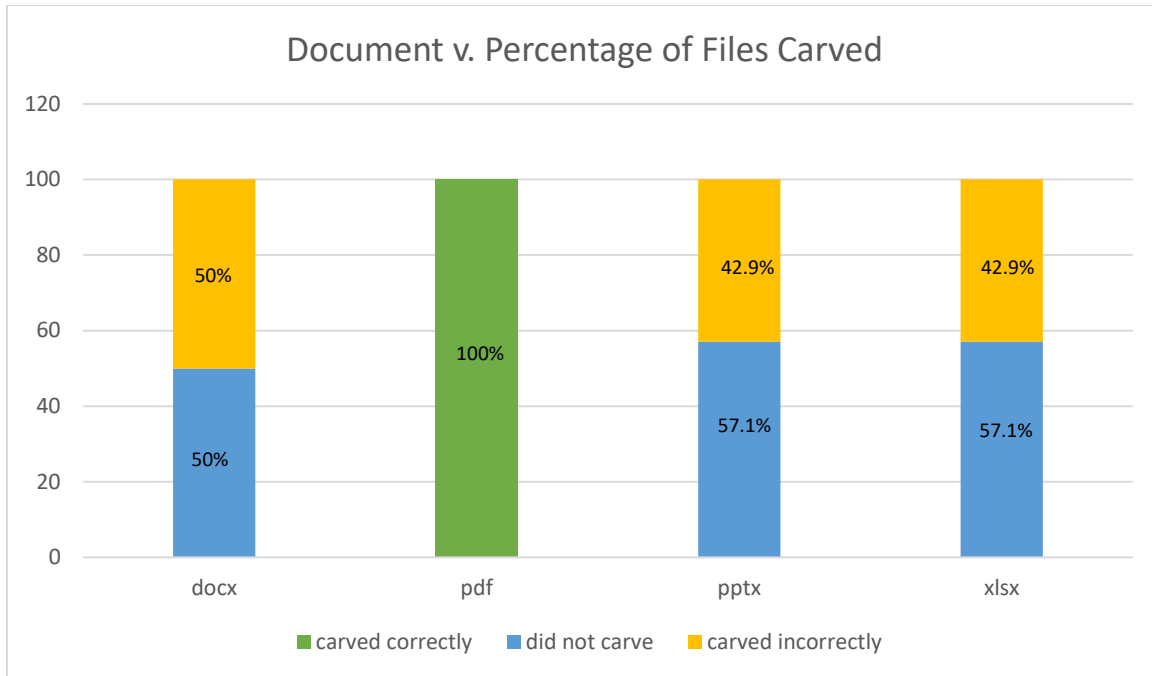


Figure 3

Document File Type v. Percentages of Files Carved

Graphic Files had the highest percentage of files carved correctly. Pcx files had the highest rate of files that were not carved. Bmp, gif and tif file extensions were combined for the pairwise chi-squared or Fisher's exact tests. There was no significant difference between bmp_gif_tif and png file extensions. The remaining file extensions' pairwise or Fisher's exact tests showed significant differences.

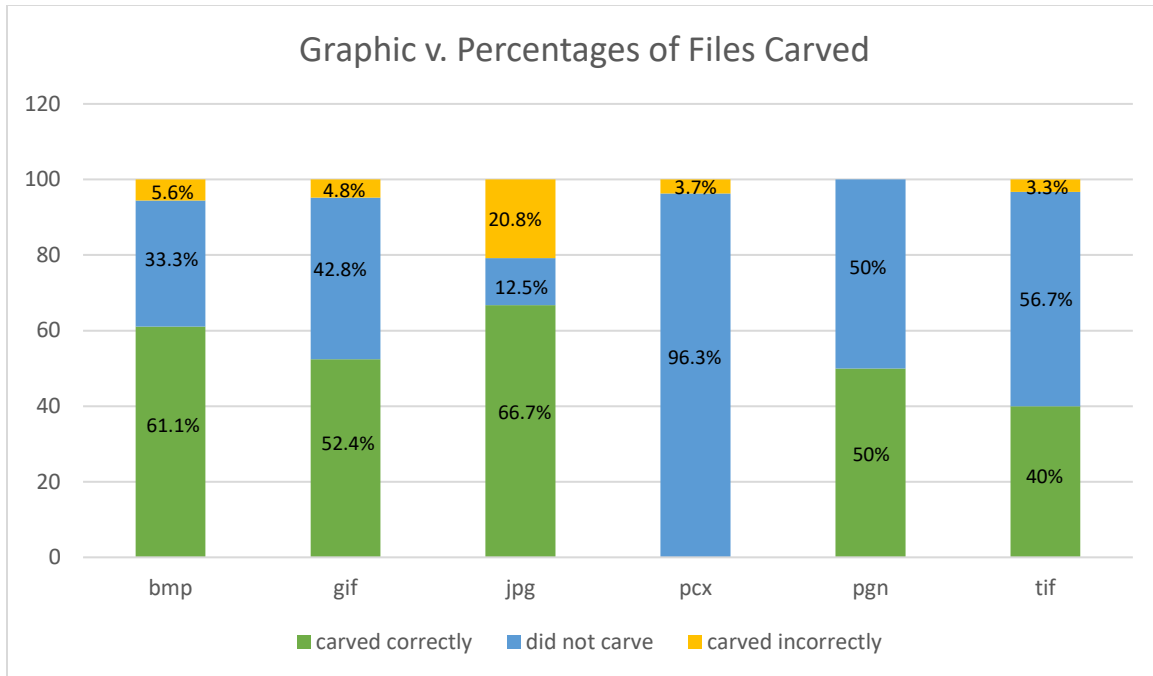


Figure 4

Graphic Files v. Percentages of Files Carved

The video files were not carved with FTK. These file extensions were not in FTK’s list of available extensions to carve.

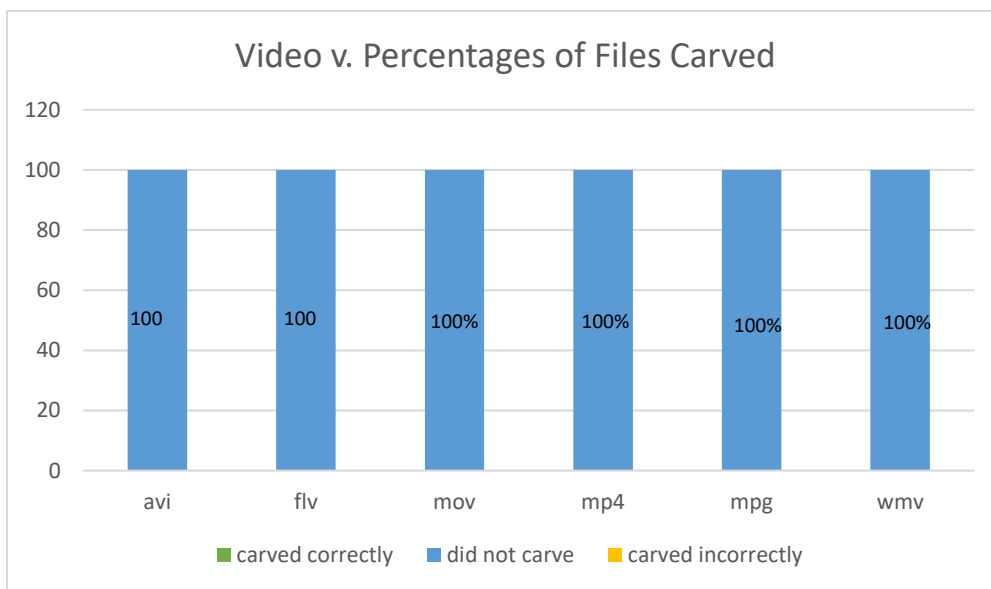


Figure 5

Video Files v. Percentages of Files Carved

The only Archive file that was able to be carved correctly were zip files. The other file extensions either did not carve or were carved incorrectly. The Archive files that were carved incorrectly showed up as jpg files. The pairwise chi-square tests showed significant differences between all file extensions except between 7z and rar files ($p=0.1306$)

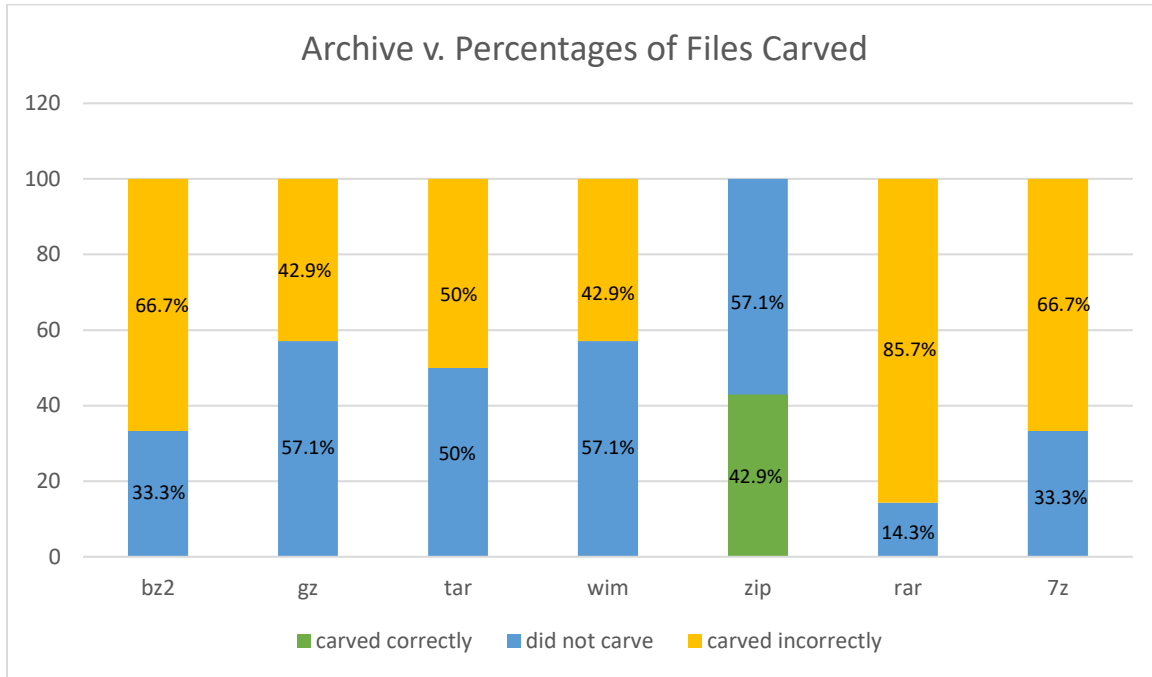


Figure 6

Archive Files v. Percentages of Files Carved

The Audio files were not able to be carved with FTK. These file extensions were not present in FTK's list of extensions to carve.

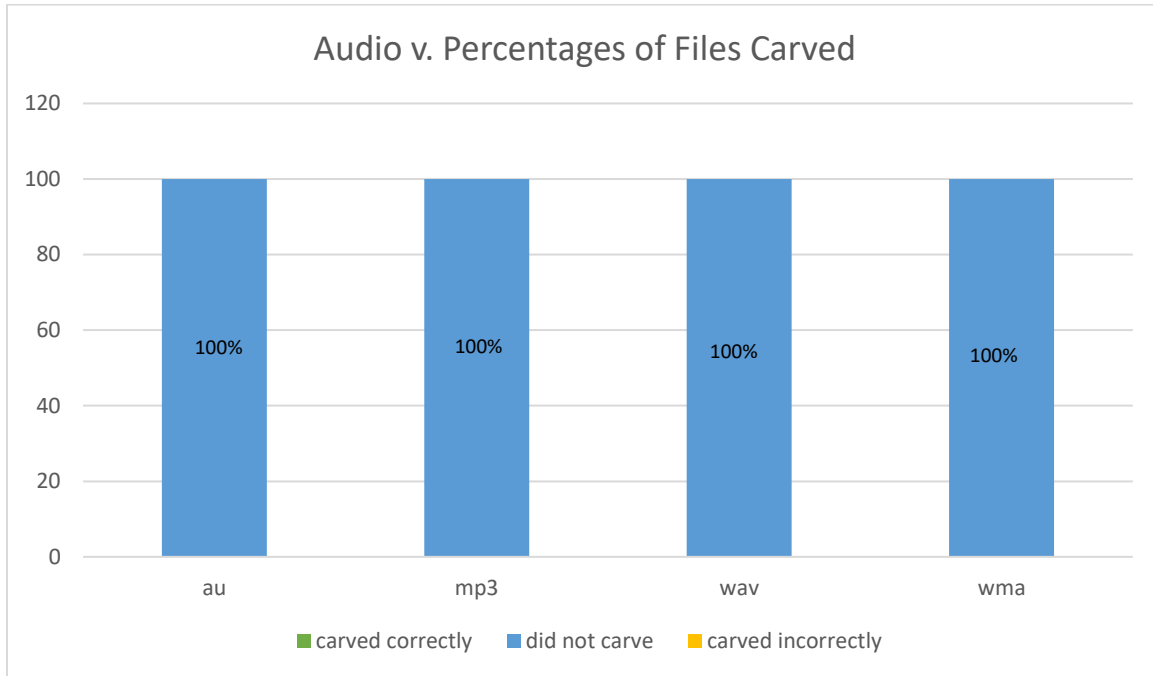


Figure 7

Audio Files v. Percentages of Files Carved

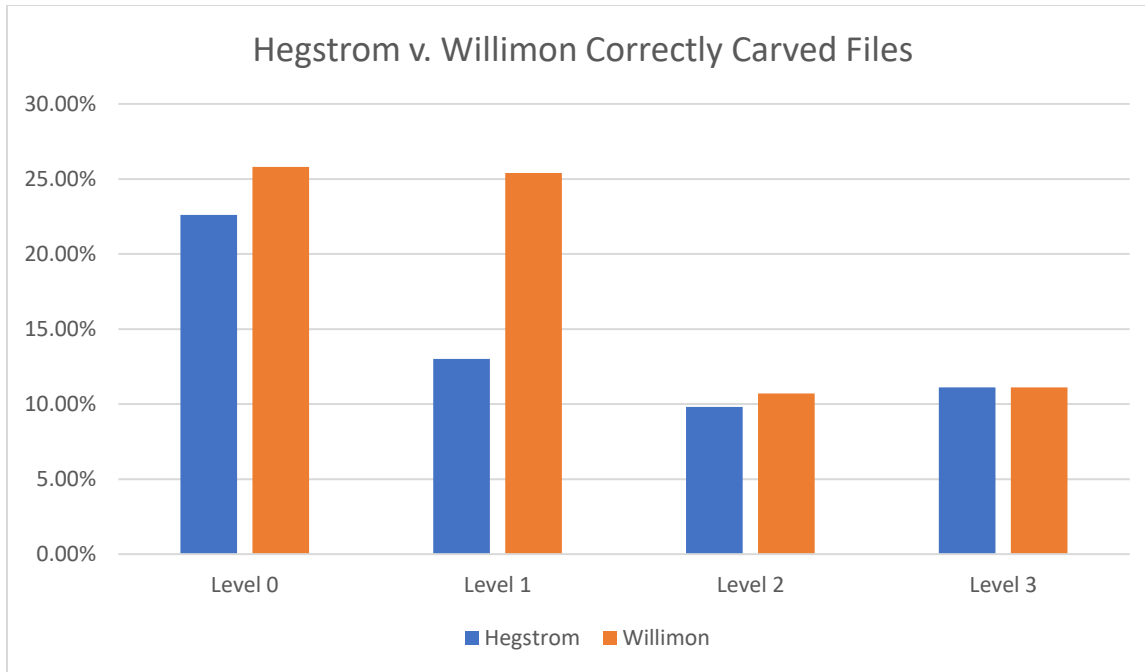


Figure 8

Hegstrom v. Willimon Correctly Carved Files

When comparing Hegstrom’s results with the results of this study, particularly the correctly carved files, the results are greatly similar with the exception of Level 1. Level 3 results were identical. This study did not report any false positive results as Hegstrom’s study did. This occurred when FTK reported to find files that were not actually present in the location stated.

CHAPTER 5

Discussion

Discussion

Within this study, FTK's file carving function performed fairly well when carving files from the SATA drives. There were many cases of files that were carved incorrectly. This showed up when the program read the file extension incorrectly or did not carve the entirety of the file. The former occurred mainly with the Archive file extensions and xlsx document files.

FTK's file carving function performed best when analyzing test images from Level 0. Level 1 was shortly behind by 0.4%. Levels 2 and 3 were more than half as successful than 0 and 1.

Graphic files had the highest percentages of test images carved correctly. Audio and Video file types were not able to be carved correctly. Archive had the highest percentage of files that were carved incorrectly, followed by Document files. Pdf extensions showed the highest percentage of correctly carved files at 100%.

The use of NIST's CFTT project's corpora was simple. The test images are easily available on the CFTT website. They are sized in a way that makes the process of restoring, copying, making forensic images a relatively quick task. The website provides the contents of the test images as well as their layouts available which aids in the process used in this study. The simplicity and availability of these test images saved a lot of time in this research and can be used in many capacities. This shows how useful digital corpora can be within this community and the need to develop a wider variety of corpora in the digital forensics area.

Limitations

The main limitation of this study was the sole use of FTK's file carving tool. The data used was limited to the data sets provided by NIST's CFTT project.

Time was another limiting factor. Due to each drive being the same size, the processing times were the same. Each drive took approximately 10 minutes to wipe, another 10 to hash, approximately 40 minutes to create an image file, and another hour to run through FTK. Once the file carving was complete, it took another hour at least to add it into the case file in FTK.

Following Kristina Hegstrom's methodology, we skipped trying to complete the restoration and imaging process solely through FTK due to FTK imager not allowing disks to be physically edited. We used WinHex to restore each test image to the drive to maintain the original formatting.

A complication encountered early on in this study was the amount of space that these test images turned into once they were imaged in FTK. Due to no compression being performed, the images were the size of the drive they encompassed. The computer used quickly ran out of space so, the forensic images were saved on an external hard drive.

Another limitation was the use of only the file extensions provided in FTK. While the ability to add other file extensions to carve exists, this study solely used the provided extensions.

Recommendations for Future Research

In future studies, we recommend that researchers utilize the ability to add the file extensions that are included in the test images if they are not in the standard extensions provided in FTK. This would ensure a more accurate representation of the abilities of FTK's file carving

function. FTK gives the user the ability to add file extensions based on file headers so, any file extension that is utilized within a test image can be added and carved. Using this function could have increased the percentages of correctly carved files.

Using a F.R.E.D forensic computer, or something similar, would greatly benefit future researchers. This research was conducted using a classroom desktop computer attached to a F.R.E.D. but, had a forensic computer been used, it could have greatly cut down on the time needed to complete the processes. These machines are made with the purpose of running forensic programs in mind and therefore, typically have faster processors and have a larger capacity for data.

Future researchers could change the file system of the test images in WinHex in order to run them as NTFS files. This would show how FTK's file carving function works on different file systems and give more insight to the tool's ability.

Conclusion

Two major conclusions were reached in this study. The first conclusion is that the file carving function, when applied to SATA drives, varies based on both file type and level. Second, the base file extensions in FTK's file carving function do not fully encompass the varieties of file types encountered in the field. Lastly, digital corpora is a highly useful and effective tool in file carving.

The first conclusion is that the level of fragmentation as well as the file type had a major impact on the results. The higher the level of fragmentation, the less likely FTK was able to correctly carve the file with the exception of levels 2-3. Graphic files had the highest percentage

of correctly carved files followed by Document file types. Audio and Video file types yielded no files correctly carved due to the file extensions not being in FTK's standard file extension list.

The second conclusion is that the file extensions within FTK's file carving tool do not adequately represent the file extensions that would be encountered in all forensic analyses. The ability to add file extensions is extremely helpful, but the user must be aware of what file extension they are carving for.

This study shows the need for more digital corpora. The test images provided and used during the research aided in the analysis of FTK's file carving function and showed ways in which testing can be improved. The expansion of available digital corpora can only benefit future researchers in this field. The task of constructing such corpora takes time and effort but, it is a task that is not without benefits. Building and perfecting such data sets aids the digital forensic community in not only validation of tools but, progresses the integrity of the field.

In conclusion, while the results of this research were not on par with the projected results per AccessData, it brought the variables to improve upon to light for further researchers. This study showed that solely trusting the tool to perform it is assigned tasks is not enough. Validation and verification studies are necessary and one cannot simply assume that the tool is performing as advertised by the supplier. In order to bring the digital forensic discipline up to the standards or it is fellow forensic disciplines, examiners must do their part. This will be no small feat but, advancing the field in ways similar to fields such as Toxicology and DNA can only gain the field it is well deserved respect.

References

- Beckett, J., Slay, J. (2007). Digital Forensics: Validation and Verification in a Dynamic Work Environment. *HICSS'07*, 1-10.
- Garfinkel, S. (2010). Digital forensics research: The next 10 years. *Digital Investigation*, S64-S73
- Garfinkel, S., Farrell, P., Roussev, V., Dinolt, G. (2009). Bringing science to digital forensics with standardized forensic corpora. *Digital Investigation*, 6, S2-S11.
- Guo, Y., Slay, J. (2010). Data Recovery Function Testing for Digital Forensic Tools. *Advances in Digital Forensics VI, IFIP AICT 337*, 297-311
- Guo, Y., Slay, J., Beckett, J. (2009). Validation and verification of computer forensic software tools- Searching Function. *Digital Investigation*, 6, S12-S22.
- Hegstrom, K. (2016). Use of Forensic Corpora in Validation of Data Carving on Solid-State Drives.
- National Institute of Science and Technology. (2014). *FTK v4.1: Test Results for Graphic File Carving Tool*.
- SWGDE recommended guidelines for validation testing*. Version 2.0 (2014).
- Wilsdon, T., Slay, J. (2005). Digital Forensics: Exploring Validation, Verification & Certification. *SADFE'05*, 1-8.

Appendix A: File Carving Test Images

Image	Description	File Types	File System
FAT_L0_Graphic.dd.bz2	Non-fragmented graphics files	JPG,PNG,BMP,GIF,TIF,PCX	FAT32
FAT_L1_Graphic.dd.bz2	Sequentially fragmented graphics files	JPG,PNG,BMP,GIF,TIF,PCX	FAT32
FAT_L2_Graphic.dd.bz2	Non-sequentially fragmented graphics files	JPG,PNG,BMP,GIF,TIF,PCX	FAT32
FAT_L3_Graphic.dd.bz2	Graphics files with missing fragments	JPG,PNG,BMP,GIF,TIF,PCX	FAT32
FAT_L0_Document.dd.bz2	Non-fragmented document files	DOC, XLS, PPT, PDF	FAT32
FAT_L1_Document.dd.bz2	Sequentially fragmented document files	DOC, XLS, PPT, PDF	FAT32
FAT_L2_Document.dd.bz2	Non-sequentially fragmented document files	DOC, XLS, PPT, PDF	FAT32
FAT_L3_Document.dd.bz2	Document files with missing fragments	DOC, XLS, PPT, PDF	FAT32
FAT_L0_Archive.dd.bz2	Non-fragmented archive files	7Z,BZ2,GZ,TAR,WIM,RAR,ZIP	FAT32
FAT_L1_Archive.dd.bz2	Sequentially fragmented archive files	7Z,BZ2,GZ,TAR,WIM,RAR,ZIP	FAT32
FAT_L2_Archive.dd.bz2	Non-Sequentially fragmented archive files	7Z,BZ2,GZ,TAR,WIM,RAR,ZIP	FAT32
FAT_L3_Archive.dd.bz2	Archive files with missing fragments	7Z,BZ2,GZ,TAR,WIM,RAR,ZIP	FAT32
FAT_L0_Audio.dd.bz2	Non-Fragmented audio files	MP3,WAV,AU,WMA	FAT32

Image	Description	File Types	File System
FAT_L1_Audio.dd.bz2	Sequentially fragmented audio files	MP3,WAV,AU,WMA	FAT32
FAT_L2_Audio.dd.bz2	Non-Sequentially fragmented audio files	MP3,WAV,AU,WMA	FAT32
FAT_L3_Audio.dd.bz2	Audio files with missing fragments	MP3,WAV,AU,WMA	FAT32
FAT_L0_Video.dd.bz2	Non-Fragmented video files	MP4,AVI,MOV,FLV,MPG,WMV	FAT32
FAT_L1_Video.dd.bz2	Sequentially fragmented video files	MP4,AVI,MOV,FLV,MPG,WMV	FAT32
FAT_L2_Video.dd.bz2	Non-sequentially fragmented video files	MP4,AVI,MOV,FLV,MPG,WMV	FAT32
FAT_L3_Video.dd.bz2	Video files with missing fragments	MP4,AVI,MOV,FLV,MPG,WMV	FAT32

Appendix B: Recorded File Carving Results

File Name	ext	Level	Type	Carved
000_021	png	0	Graphic	1
000_021	png	0	Graphic	1
000_021	png	0	Graphic	1
02010025	pcx	0	Graphic	0
02010025	pcx	0	Graphic	0
02010025	pcx	0	Graphic	0
02010026	jpg	0	Graphic	(1)
02010026	jpg	0	Graphic	(1)
02010026	jpg	0	Graphic	(1)
09260002	jpg	0	Graphic	1
09260002	jpg	0	Graphic	1
09260002	jpg	0	Graphic	1
100_0304crop	bmp	0	Graphic	1
100_0304crop	bmp	0	Graphic	1
100_0304crop	bmp	0	Graphic	1
100_0018	tif	0	Graphic	1
100_0018	tif	0	Graphic	1
100_0018	tif	0	Graphic	1
100_0183	gif	0	Graphic	1
100_0183	gif	0	Graphic	1
100_0183	gif	0	Graphic	1
09260002 (1)	jpg	1	Graphic	1
09260002 (1)	jpg	1	Graphic	1
09260002 (1)	jpg	1	Graphic	1
09260002 (2)	jpg	1	Graphic	1
09260002 (2)	jpg	1	Graphic	1
09260002 (2)	jpg	1	Graphic	1
100_0018 (1)	tif	1	Graphic	1
100_0018 (1)	tif	1	Graphic	1
100_0018 (1)	tif	1	Graphic	1
100_0018 (2)	tif	1	Graphic	1
100_0018 (2)	tif	1	Graphic	1
100_0018 (2)	tif	1	Graphic	1
100_0018 (3)	tif	1	Graphic	1
100_0018 (3)	tif	1	Graphic	1
100_0018 (3)	tif	1	Graphic	1
100_0304crop (1)	bmp	1	Graphic	1
100_0304crop (1)	bmp	1	Graphic	1
100_0304crop (1)	bmp	1	Graphic	1
100_0304crop (2)	bmp	1	Graphic	1

100_0304crop (2)	bmp	1	Graphic	1
100_0304crop (2)	bmp	1	Graphic	1
02010025 (1)	pcx	1	Graphic	0
02010025 (1)	pcx	1	Graphic	0
02010025 (1)	pcx	1	Graphic	0
02010025 (2)	pcx	1	Graphic	0
02010025 (2)	pcx	1	Graphic	0
02010025 (2)	pcx	1	Graphic	0
02010025 (3)	pcx	1	Graphic	0
02010025 (3)	pcx	1	Graphic	0
02010025 (3)	pcx	1	Graphic	0
100_0183 (1)	gif	1	Graphic	1
100_0183 (1)	gif	1	Graphic	1
100_0183 (1)	gif	1	Graphic	1
100_0183 (2)	gif	1	Graphic	1
100_0183 (2)	gif	1	Graphic	1
100_0183 (2)	gif	1	Graphic	1
000_021 (1)	png	1	Graphic	1
000_021 (1)	png	1	Graphic	1
000_021 (1)	png	1	Graphic	1
000_021 (2)	png	1	Graphic	1
000_021 (2)	png	1	Graphic	1
000_021 (2)	png	1	Graphic	1
100_0018 (1)	tif	2	Graphic	0
100_0018 (1)	tif	2	Graphic	0
100_0018 (1)	tif	2	Graphic	0
100_0018 (3)	tif	2	Graphic	0
100_0018 (3)	tif	2	Graphic	0
100_0018 (3)	tif	2	Graphic	0
100_0018 (2)	tif	2	Graphic	0
100_0018 (2)	tif	2	Graphic	0
100_0018 (2)	tif	2	Graphic	0
09260002 (1)	jpg	2	Graphic	1
09260002 (1)	jpg	2	Graphic	1
09260002 (1)	jpg	2	Graphic	(1)
09260002 (3)	jpg	2	Graphic	1
09260002 (3)	jpg	2	Graphic	1
09260002 (3)	jpg	2	Graphic	(1)
09260002 (2)	jpg	2	Graphic	0
09260002 (2)	jpg	2	Graphic	0
09260002 (2)	jpg	2	Graphic	0
100_0304crop (2)	bmp	2	Graphic	0
100_0304crop (2)	bmp	2	Graphic	0

Running head: THE UTILIZATION OF FORENSIC CORPORA

100_0304crop (2)	bmp	2	Graphic	0
100_0304crop (1)	bmp	2	Graphic	1
100_0304crop (1)	bmp	2	Graphic	1
100_0304crop (1)	bmp	2	Graphic	(1)
02010025 (1)	pcx	2	Graphic	0
02010025 (1)	pcx	2	Graphic	0
02010025 (1)	pcx	2	Graphic	0
02010025 (3)	pcx	2	Graphic	0
02010025 (3)	pcx	2	Graphic	0
02010025 (3)	pcx	2	Graphic	0
02010025 (2)	pcx	2	Graphic	0
02010025 (2)	pcx	2	Graphic	0
02010025 (2)	pcx	2	Graphic	0
100_0183 (3)	gif	2	Graphic	0
100_0183 (3)	gif	2	Graphic	0
100_0183 (3)	gif	2	Graphic	0
100_0183 (1)	gif	2	Graphic	1
100_0183 (1)	gif	2	Graphic	1
100_0183 (1)	gif	2	Graphic	(1)
100_0183 (2)	gif	2	Graphic	0
100_0183 (2)	gif	2	Graphic	0
100_0183 (2)	gif	2	Graphic	0
000_021 (2)	png	2	Graphic	0
000_021 (2)	png	2	Graphic	0
000_021 (2)	png	2	Graphic	0
000_021 (1)	png	2	Graphic	0
000_021 (1)	png	2	Graphic	0
000_021 (1)	png	2	Graphic	0
09260002 (1)	jpg	3	Graphic	1
09260002 (1)	jpg	3	Graphic	1
09260002 (1)	jpg	3	Graphic	1
100_0018 (1)	tif	3	Graphic	(1)
100_0018 (1)	tif	3	Graphic	0
100_0018 (1)	tif	3	Graphic	0
100_0018 (2)	tif	3	Graphic	0
100_0018 (2)	tif	3	Graphic	0
100_0018 (2)	tif	3	Graphic	0
100_0304crop (2)	bmp	3	Graphic	0
100_0304crop (2)	bmp	3	Graphic	0
100_0304crop (2)	bmp	3	Graphic	0
02010025 (1)	pcx	3	Graphic	(1)
02010025 (1)	pcx	3	Graphic	0
02010025 (1)	pcx	3	Graphic	0

Running head: THE UTILIZATION OF FORENSIC CORPORA

02010025 (3)	pcx	3	Graphic	0
02010025 (3)	pcx	3	Graphic	0
02010025 (3)	pcx	3	Graphic	0
100_0018 (2)	tif	3	Graphic	0
100_0018 (2)	tif	3	Graphic	0
100_0018 (2)	tif	3	Graphic	0
100_0183 (3)	gif	3	Graphic	0
100_0183 (3)	gif	3	Graphic	0
100_0183 (3)	gif	3	Graphic	0
000_021 (3)	png	3	Graphic	0
000_021 (3)	png	3	Graphic	0
000_021 (3)	png	3	Graphic	0
D1	pdf	0	Document	1
D1	pdf	0	Document	1
D1	pdf	0	Document	1
D2	pdf	0	Document	1
D2	pdf	0	Document	1
D2	pdf	0	Document	1
D3	xlsx	0	Document	(1)
D3	xlsx	0	Document	(1)
D3	xlsx	0	Document	(1)
D4	xlsx	0	Document	(1)
D4	xlsx	0	Document	(1)
D4	xlsx	0	Document	(1)
D5	docx	0	Document	(1)
D5	docx	0	Document	(1)
D5	docx	0	Document	(1)
D6	docx	0	Document	(1)
D6	docx	0	Document	(1)
D6	docx	0	Document	(1)
D7	pptx	0	Document	(1)
D7	pptx	0	Document	(1)
D7	pptx	0	Document	(1)
D1 (1)	pdf	1	Document	1
D1 (1)	pdf	1	Document	1
D1 (1)	pdf	1	Document	1
D1 (2)	pdf	1	Document	1
D1 (2)	pdf	1	Document	1
D1 (2)	pdf	1	Document	1
D2 (1)	pdf	1	Document	1
D2 (1)	pdf	1	Document	1
D2 (1)	pdf	1	Document	1
D2 (2)	pdf	1	Document	1

Running head: THE UTILIZATION OF FORENSIC CORPORA

D2 (2)	pdf	1	Document	1
D2 (2)	pdf	1	Document	1
D2 (3)	pdf	1	Document	1
D2 (3)	pdf	1	Document	1
D2 (3)	pdf	1	Document	1
D3 (1)	xlsx	1	Document	(1)
D3 (1)	xlsx	1	Document	(1)
D3 (1)	xlsx	1	Document	(1)
D3 (2,3)	xlsx	1	Document	(1)
D3 (2,3)	xlsx	1	Document	(1)
D3 (2,3)	xlsx	1	Document	(1)
D4 (1,2)	xlsx	1	Document	(1)
D4 (1,2)	xlsx	1	Document	(1)
D4 (1,2)	xlsx	1	Document	(1)
D4 (3)	xlsx	1	Document	(1)
D4 (3)	xlsx	1	Document	(1)
D4 (3)	xlsx	1	Document	(1)
D5 (1)	docx	1	Document	(1)
D5 (1)	docx	1	Document	(1)
D5 (1)	docx	1	Document	(1)
D5 (3)	docx	1	Document	(1)
D5 (3)	docx	1	Document	(1)
D5 (3)	docx	1	Document	(1)
D6 (1)	docx	1	Document	(1)
D6 (1)	docx	1	Document	(1)
D6 (1)	docx	1	Document	(1)
D6 (2)	docx	1	Document	(1)
D6 (2)	docx	1	Document	(1)
D6 (2)	docx	1	Document	(1)
D7 (1)	pptx	1	Document	(1)
D7 (1)	pptx	1	Document	(1)
D7 (1)	pptx	1	Document	(1)
D7 (2,3)	pptx	1	Document	(1)
D7 (2,3)	pptx	1	Document	(1)
D7 (2,3)	pptx	1	Document	(1)
D1 (3)	pdf	2	Document	1
D1 (3)	pdf	2	Document	1
D1 (3)	pdf	2	Document	1
D1 (1)	pdf	2	Document	1
D1 (1)	pdf	2	Document	1
D1 (1)	pdf	2	Document	1
D1 (2)	pdf	2	Document	1
D1 (2)	pdf	2	Document	1

Running head: THE UTILIZATION OF FORENSIC CORPORA

D1 (2)	pdf	2	Document	1
D2 (2)	pdf	2	Document	1
D2 (2)	pdf	2	Document	1
D2 (2)	pdf	2	Document	1
D2 (3)	pdf	2	Document	1
D2 (3)	pdf	2	Document	1
D2 (3)	pdf	2	Document	1
D2 (1)	pdf	2	Document	1
D2 (1)	pdf	2	Document	1
D2 (1)	pdf	2	Document	1
D3 (2)	xlsx	2	Document	0
D3 (2)	xlsx	2	Document	0
D3 (2)	xlsx	2	Document	0
D3 (1)	xlsx	2	Document	0
D3 (1)	xlsx	2	Document	0
D3 (1)	xlsx	2	Document	0
D4 (2)	xlsx	2	Document	0
D4 (2)	xlsx	2	Document	0
D4 (2)	xlsx	2	Document	0
D4 (1)	xlsx	2	Document	0
D4 (1)	xlsx	2	Document	0
D4 (1)	xlsx	2	Document	0
D4 (3)	xlsx	2	Document	0
D4 (3)	xlsx	2	Document	0
D4 (3)	xlsx	2	Document	0
D5 (1)	docx	2	Document	0
D5 (1)	docx	2	Document	0
D5 (1)	docx	2	Document	0
D5 (3)	docx	2	Document	0
D5 (3)	docx	2	Document	0
D5 (3)	docx	2	Document	0
D6 (2,3)	docx	2	Document	0
D6 (2,3)	docx	2	Document	0
D6 (2,3)	docx	2	Document	0
D6 (1)	docx	2	Document	0
D6 (1)	docx	2	Document	0
D6 (1)	docx	2	Document	0
D7 (3)	pptx	2	Document	0
D7 (3)	pptx	2	Document	0
D7 (3)	pptx	2	Document	0
D7 (2)	pptx	2	Document	0
D7 (2)	pptx	2	Document	0
D7 (2)	pptx	2	Document	0

Running head: THE UTILIZATION OF FORENSIC CORPORA

D7 (1)	pptx	2	Document	0
D7 (1)	pptx	2	Document	0
D7 (1)	pptx	2	Document	0
D1 (1)	pdf	3	Document	1
D1 (1)	pdf	3	Document	1
D1 (1)	pdf	3	Document	1
D1 (2,3)	pdf	3	Document	1
D1 (2,3)	pdf	3	Document	1
D1 (2,3)	pdf	3	Document	1
D2 (1)	pdf	3	Document	1
D2 (1)	pdf	3	Document	1
D2 (1)	pdf	3	Document	1
D2 (3)	pdf	3	Document	1
D2 (3)	pdf	3	Document	1
D2 (3)	pdf	3	Document	1
D3 (2)	xlsx	3	Document	0
D3 (2)	xlsx	3	Document	0
D3 (2)	xlsx	3	Document	0
D3 (3)	xlsx	3	Document	0
D3 (3)	xlsx	3	Document	0
D3 (3)	xlsx	3	Document	0
D3 (3)	xlsx	3	Document	0
D4 (2)	xlsx	3	Document	0
D4 (2)	xlsx	3	Document	0
D4 (2)	xlsx	3	Document	0
D5 (1)	docx	3	Document	0
D5 (1)	docx	3	Document	0
D5 (1)	docx	3	Document	0
D6 (2,3)	docx	3	Document	0
D6 (2,3)	docx	3	Document	0
D6 (2,3)	docx	3	Document	0
D7 (3)	pptx	3	Document	0
D7 (3)	pptx	3	Document	0
D7 (3)	pptx	3	Document	0
arc1	7z	0	Archive	(1)
arc1	7z	0	Archive	(1)
arc1	7z	0	Archive	(1)
arc2	bz2	0	Archive	(1)
arc2	bz2	0	Archive	(1)
arc2	bz2	0	Archive	(1)
arc3	gz	0	Archive	(1)
arc3	gz	0	Archive	(1)
arc3	gz	0	Archive	(1)
arc4	tar	0	Archive	(1)

Running head: THE UTILIZATION OF FORENSIC CORPORA

arc4	tar	0	Archive	(1)
arc4	tar	0	Archive	(1)
arc5	wim	0	Archive	(1)
arc5	wim	0	Archive	(1)
arc5	wim	0	Archive	(1)
arc6	rar	0	Archive	(1)
arc6	rar	0	Archive	(1)
arc6	rar	0	Archive	(1)
arc7	zip	0	Archive	1
arc7	zip	0	Archive	1
arc7	zip	0	Archive	1
arc1 (1)	7z	1	Archive	(1)
arc1 (1)	7z	1	Archive	(1)
arc1 (1)	7z	1	Archive	(1)
arc1 (2)	7z	1	Archive	(1)
arc1 (2)	7z	1	Archive	(1)
arc1 (2)	7z	1	Archive	(1)
arc2 (1)	bz2	1	Archive	(1)
arc2 (1)	bz2	1	Archive	(1)
arc2 (1)	bz2	1	Archive	(1)
arc2 (2)	bz2	1	Archive	(1)
arc2 (2)	bz2	1	Archive	(1)
arc2 (2)	bz2	1	Archive	(1)
arc2 (3)	bz2	1	Archive	(1)
arc2 (3)	bz2	1	Archive	(1)
arc2 (3)	bz2	1	Archive	(1)
arc3 (1)	gz	1	Archive	(1)
arc3 (1)	gz	1	Archive	(1)
arc3 (1)	gz	1	Archive	(1)
arc3 (2)	gz	1	Archive	(1)
arc3 (2)	gz	1	Archive	(1)
arc3 (2)	gz	1	Archive	(1)
arc4 (1)	tar	1	Archive	(1)
arc4 (1)	tar	1	Archive	(1)
arc4 (1)	tar	1	Archive	(1)
arc4 (2)	tar	1	Archive	(1)
arc4 (2)	tar	1	Archive	(1)
arc4 (2)	tar	1	Archive	(1)
arc4 (3)	tar	1	Archive	(1)
arc4 (3)	tar	1	Archive	(1)
arc4 (3)	tar	1	Archive	(1)
arc5 (1)	wim	1	Archive	(1)
arc5 (1)	wim	1	Archive	(1)

Running head: THE UTILIZATION OF FORENSIC CORPORA

arc5 (1)	wim	1	Archive	(1)
arc5 (2)	wim	1	Archive	(1)
arc5 (2)	wim	1	Archive	(1)
arc5 (2)	wim	1	Archive	(1)
arc6 (1)	rar	1	Archive	(1)
arc6 (1)	rar	1	Archive	(1)
arc6 (1)	rar	1	Archive	(1)
arc6 (2)	rar	1	Archive	(1)
arc6 (2)	rar	1	Archive	(1)
arc6 (2)	rar	1	Archive	(1)
arc6 (3)	rar	1	Archive	(1)
arc6 (3)	rar	1	Archive	(1)
arc6 (3)	rar	1	Archive	(1)
arc7 (1)	zip	1	Archive	1
arc7 (1)	zip	1	Archive	1
arc7 (1)	zip	1	Archive	1
arc7 (2)	zip	1	Archive	1
arc7 (2)	zip	1	Archive	1
arc7 (2)	zip	1	Archive	1
arc6 (2)	rar	2	Archive	(1)
arc6 (2)	rar	2	Archive	(1)
arc6 (2)	rar	2	Archive	(1)
arc6 (1)	rar	2	Archive	(1)
arc6 (1)	rar	2	Archive	(1)
arc6 (1)	rar	2	Archive	(1)
arc7 (1)	zip	2	Archive	0
arc7 (1)	zip	2	Archive	0
arc7 (1)	zip	2	Archive	0
arc7 (3)	zip	2	Archive	0
arc7 (3)	zip	2	Archive	0
arc7 (3)	zip	2	Archive	0
arc7 (2)	zip	2	Archive	0
arc7 (2)	zip	2	Archive	0
arc7 (2)	zip	2	Archive	0
arc1 (2)	7z	2	Archive	(1)
arc1 (2)	7z	2	Archive	(1)
arc1 (2)	7z	2	Archive	(1)
arc1 (2)	7z	2	Archive	(1)
arc1 (2)	7z	2	Archive	(1)
arc1 (1)	7z	2	Archive	(1)
arc1 (1)	7z	2	Archive	(1)
arc1 (1)	7z	2	Archive	(1)

Running head: THE UTILIZATION OF FORENSIC CORPORA

arc1 (3)	7z	2	Archive	0
arc1 (3)	7z	2	Archive	0
arc1 (3)	7z	2	Archive	0
arc2 (2)	bz2	2	Archive	(1)
arc2 (2)	bz2	2	Archive	(1)
arc2 (2)	bz2	2	Archive	(1)
arc2 (3)	bz2	2	Archive	(1)
arc2 (3)	bz2	2	Archive	(1)
arc2 (3)	bz2	2	Archive	(1)
arc2 (1)	bz2	2	Archive	0
arc2 (1)	bz2	2	Archive	0
arc2 (1)	bz2	2	Archive	0
arc4 (3)	tar	2	Archive	0
arc4 (3)	tar	2	Archive	0
arc4 (3)	tar	2	Archive	0
arc4 (1)	tar	2	Archive	0
arc4 (1)	tar	2	Archive	0
arc4 (1)	tar	2	Archive	0
arc4 (2)	tar	2	Archive	0
arc4 (2)	tar	2	Archive	0
arc4 (2)	tar	2	Archive	0
arc4 (2)	tar	2	Archive	0
arc5 (3)	wim	2	Archive	0
arc5 (3)	wim	2	Archive	0
arc5 (3)	wim	2	Archive	0
arc5 (2)	wim	2	Archive	0
arc5 (2)	wim	2	Archive	0
arc5 (2)	wim	2	Archive	0
arc5 (1)	wim	2	Archive	0
arc5 (1)	wim	2	Archive	0
arc5 (1)	wim	2	Archive	0
arc3 (2,3)	gz	2	Archive	0
arc3 (2,3)	gz	2	Archive	0
arc3 (2,3)	gz	2	Archive	0
arc3 (1)	gz	2	Archive	0
arc3 (1)	gz	2	Archive	0
arc3 (1)	gz	2	Archive	0
arc6 (1)	rar	3	Archive	0
arc6 (1)	rar	3	Archive	0
arc6 (1)	rar	3	Archive	0
arc7 (2)	zip	3	Archive	0
arc7 (2)	zip	3	Archive	0
arc7 (2)	zip	3	Archive	0
arc3 (1)	gz	3	Archive	0

Running head: THE UTILIZATION OF FORENSIC CORPORA

arc3 (1)	gz	3	Archive	0
arc3 (1)	gz	3	Archive	0
arc3 (2)	gz	3	Archive	0
arc3 (2)	gz	3	Archive	0
arc3 (2)	gz	3	Archive	0
arc2 (1)	bz2	3	Archive	0
arc2 (1)	bz2	3	Archive	0
arc2 (1)	bz2	3	Archive	0
arc2 (3)	bz2	3	Archive	0
arc2 (3)	bz2	3	Archive	0
arc2 (3)	bz2	3	Archive	0
arc1 (2)	7z	3	Archive	0
arc1 (2)	7z	3	Archive	0
arc1 (2)	7z	3	Archive	0
arc1 (3)	7z	3	Archive	0
arc1 (3)	7z	3	Archive	0
arc1 (3)	7z	3	Archive	0
arc4 (3)	tar	3	Archive	0
arc4 (3)	tar	3	Archive	0
arc4 (3)	tar	3	Archive	0
arc5 (2,3)	wim	3	Archive	0
arc5 (2,3)	wim	3	Archive	0
arc5 (2,3)	wim	3	Archive	0
Audio1	mp3	0	Audio	0
Audio1	mp3	0	Audio	0
Audio1	mp3	0	Audio	0
Audio2	wav	0	Audio	0
Audio2	wav	0	Audio	0
Audio2	wav	0	Audio	0
Audio3	au	0	Audio	0
Audio3	au	0	Audio	0
Audio3	au	0	Audio	0
Audio4	wma	0	Audio	0
Audio4	wma	0	Audio	0
Audio4	wma	0	Audio	0
Audio1 (1)	mp3	1	Audio	0
Audio1 (1)	mp3	1	Audio	0
Audio1 (1)	mp3	1	Audio	0
Audio1 (2)	mp3	1	Audio	0
Audio1 (2)	mp3	1	Audio	0
Audio1 (2)	mp3	1	Audio	0
Audio2 (1)	wav	1	Audio	0
Audio2 (1)	wav	1	Audio	0

Running head: THE UTILIZATION OF FORENSIC CORPORA

Audio2 (1)	wav	1	Audio	0
Audio2 (2)	wav	1	Audio	0
Audio2 (2)	wav	1	Audio	0
Audio2 (2)	wav	1	Audio	0
Audio2 (3)	wav	1	Audio	0
Audio2 (3)	wav	1	Audio	0
Audio2 (3)	wav	1	Audio	0
Audio3 (1)	au	1	Audio	0
Audio3 (1)	au	1	Audio	0
Audio3 (1)	au	1	Audio	0
Audio3 (2)	au	1	Audio	0
Audio3 (2)	au	1	Audio	0
Audio3 (2)	au	1	Audio	0
Audio4 (1)	wma	1	Audio	0
Audio4 (1)	wma	1	Audio	0
Audio4 (1)	wma	1	Audio	0
Audio4 (2)	wma	1	Audio	0
Audio4 (2)	wma	1	Audio	0
Audio4 (2)	wma	1	Audio	0
Audio4 (3)	wma	1	Audio	0
Audio4 (3)	wma	1	Audio	0
Audio4 (3)	wma	1	Audio	0
Audio1 (2)	mp3	2	Audio	0
Audio1 (2)	mp3	2	Audio	0
Audio1 (2)	mp3	2	Audio	0
Audio1 (1)	mp3	2	Audio	0
Audio1 (1)	mp3	2	Audio	0
Audio1 (1)	mp3	2	Audio	0
Audio2 (1)	wav	2	Audio	0
Audio2 (1)	wav	2	Audio	0
Audio2 (1)	wav	2	Audio	0
Audio2 (3)	wav	2	Audio	0
Audio2 (3)	wav	2	Audio	0
Audio2 (3)	wav	2	Audio	0
Audio2 (2)	wav	2	Audio	0
Audio2 (2)	wav	2	Audio	0
Audio2 (2)	wav	2	Audio	0
Audio3 (2)	au	2	Audio	0
Audio3 (2)	au	2	Audio	0
Audio3 (2)	au	2	Audio	0
Audio3 (1)	au	2	Audio	0
Audio3 (1)	au	2	Audio	0
Audio3 (1)	au	2	Audio	0

Running head: THE UTILIZATION OF FORENSIC CORPORA

Audio3 (3)	au	2	Audio	0
Audio3 (3)	au	2	Audio	0
Audio3 (3)	au	2	Audio	0
Audio4 (3)	wma	2	Audio	0
Audio4 (3)	wma	2	Audio	0
Audio4 (3)	wma	2	Audio	0
Audio4 (2)	wma	2	Audio	0
Audio4 (2)	wma	2	Audio	0
Audio4 (2)	wma	2	Audio	0
Audio4 (1)	wma	2	Audio	0
Audio4 (1)	wma	2	Audio	0
Audio4 (1)	wma	2	Audio	0
Audio1 (1)	mp3	3	Audio	0
Audio1 (1)	mp3	3	Audio	0
Audio1 (1)	mp3	3	Audio	0
Audio1 (3)	mp3	3	Audio	0
Audio1 (3)	mp3	3	Audio	0
Audio1 (3)	mp3	3	Audio	0
Audio2 (1)	wav	3	Audio	0
Audio2 (1)	wav	3	Audio	0
Audio2 (1)	wav	3	Audio	0
Audio3 (2)	au	3	Audio	0
Audio3 (2)	au	3	Audio	0
Audio3 (2)	au	3	Audio	0
Audio3 (3)	au	3	Audio	0
Audio3 (3)	au	3	Audio	0
Audio3 (3)	au	3	Audio	0
Audio4 (1)	wma	3	Audio	0
Audio4 (1)	wma	3	Audio	0
Audio4 (1)	wma	3	Audio	0
Audio4 (2)	wma	3	Audio	0
Audio4 (2)	wma	3	Audio	0
Audio4 (2)	wma	3	Audio	0
vid1	mp4	0	Video	0
vid1	mp4	0	Video	0
vid1	mp4	0	Video	0
vid2	avi	0	Video	0
vid2	avi	0	Video	0
vid2	avi	0	Video	0
vid3	mov	0	Video	0
vid3	mov	0	Video	0
vid3	mov	0	Video	0
vid4	flv	0	Video	0

Running head: THE UTILIZATION OF FORENSIC CORPORA

vid4	flv	0	Video	0
vid4	flv	0	Video	0
vid5	mpg	0	Video	0
vid5	mpg	0	Video	0
vid5	mpg	0	Video	0
vid6	wmv	0	Video	0
vid6	wmv	0	Video	0
vid6	wmv	0	Video	0
vid1 (1)	mp4	1	Video	0
vid1 (1)	mp4	1	Video	0
vid1 (1)	mp4	1	Video	0
vid1 (2)	mp4	1	Video	0
vid1 (2)	mp4	1	Video	0
vid1 (2)	mp4	1	Video	0
vid1 (3)	mp4	1	Video	0
vid1 (3)	mp4	1	Video	0
vid1 (3)	mp4	1	Video	0
vid2 (1)	avi	1	Video	0
vid2 (1)	avi	1	Video	0
vid2 (1)	avi	1	Video	0
vid2 (2)	avi	1	Video	0
vid2 (2)	avi	1	Video	0
vid2 (2)	avi	1	Video	0
vid3 (1)	mov	1	Video	0
vid3 (1)	mov	1	Video	0
vid3 (1)	mov	1	Video	0
vid3 (2)	mov	1	Video	0
vid3 (2)	mov	1	Video	0
vid3 (2)	mov	1	Video	0
vid3 (3)	mov	1	Video	0
vid3 (3)	mov	1	Video	0
vid3 (3)	mov	1	Video	0
vid4 (1)	flv	1	Video	0
vid4 (1)	flv	1	Video	0
vid4 (1)	flv	1	Video	0
vid4 (2)	flv	1	Video	0
vid4 (2)	flv	1	Video	0
vid4 (2)	flv	1	Video	0
vid5 (1)	mpg	1	Video	0
vid5 (1)	mpg	1	Video	0
vid5 (1)	mpg	1	Video	0
vid5 (2)	mpg	1	Video	0
vid5 (2)	mpg	1	Video	0

Running head: THE UTILIZATION OF FORENSIC CORPORA

vid5 (2)	mpg	1	Video	0
vid5 (3)	mpg	1	Video	0
vid5 (3)	mpg	1	Video	0
vid5 (3)	mpg	1	Video	0
vid6 (1)	wmv	1	Video	0
vid6 (1)	wmv	1	Video	0
vid6 (1)	wmv	1	Video	0
vid6 (2)	wmv	1	Video	0
vid6 (2)	wmv	1	Video	0
vid6 (2)	wmv	1	Video	0
vid2 (2)	avi	2	Video	0
vid2 (2)	avi	2	Video	0
vid2 (2)	avi	2	Video	0
vid2 (1)	avi	2	Video	0
vid2 (1)	avi	2	Video	0
vid2 (1)	avi	2	Video	0
vid6 (1)	wmv	2	Video	0
vid6 (1)	wmv	2	Video	0
vid6 (1)	wmv	2	Video	0
vid6 (3)	wmv	2	Video	0
vid6 (3)	wmv	2	Video	0
vid6 (3)	wmv	2	Video	0
vid6 (2)	wmv	2	Video	0
vid6 (2)	wmv	2	Video	0
vid6 (2)	wmv	2	Video	0
vid1 (2)	mp4	2	Video	0
vid1 (2)	mp4	2	Video	0
vid1 (2)	mp4	2	Video	0
vid1 (1)	mp4	2	Video	0
vid1 (1)	mp4	2	Video	0
vid1 (1)	mp4	2	Video	0
vid1 (3)	mp4	2	Video	0
vid1 (3)	mp4	2	Video	0
vid1 (3)	mp4	2	Video	0
vid3 (2)	mov	2	Video	0
vid3 (2)	mov	2	Video	0
vid3 (2)	mov	2	Video	0
vid3 (3)	mov	2	Video	0
vid3 (3)	mov	2	Video	0
vid3 (3)	mov	2	Video	0
vid3 (1)	mov	2	Video	0
vid3 (1)	mov	2	Video	0
vid3 (1)	mov	2	Video	0

Running head: THE UTILIZATION OF FORENSIC CORPORA

vid5 (3)	mpg	2	Video	0
vid5 (3)	mpg	2	Video	0
vid5 (3)	mpg	2	Video	0
vid5 (2)	mpg	2	Video	0
vid5 (2)	mpg	2	Video	0
vid5 (2)	mpg	2	Video	0
vid4 (3)	flv	2	Video	0
vid4 (3)	flv	2	Video	0
vid4 (3)	flv	2	Video	0
vid4 (2)	flv	2	Video	0
vid4 (2)	flv	2	Video	0
vid4 (2)	flv	2	Video	0
vid4 (1)	flv	2	Video	0
vid4 (1)	flv	2	Video	0
vid4 (1)	flv	2	Video	0
vid4 (1)	flv	3	Video	0
vid4 (1)	flv	3	Video	0
vid4 (1)	flv	3	Video	0
vid2 (2)	avi	3	Video	0
vid2 (2)	avi	3	Video	0
vid2 (2)	avi	3	Video	0
vid1 (1)	mp4	3	Video	0
vid1 (1)	mp4	3	Video	0
vid1 (1)	mp4	3	Video	0
vid1 (2)	mp4	3	Video	0
vid1 (2)	mp4	3	Video	0
vid1 (2)	mp4	3	Video	0
vid5 (1)	mpg	3	Video	0
vid5 (1)	mpg	3	Video	0
vid5 (1)	mpg	3	Video	0
vid5 (3)	mpg	3	Video	0
vid5 (3)	mpg	3	Video	0
vid5 (3)	mpg	3	Video	0
vid6 (2)	wmv	3	Video	0
vid6 (2)	wmv	3	Video	0
vid6 (2)	wmv	3	Video	0
vid6 (3)	wmv	3	Video	0
vid6 (3)	wmv	3	Video	0
vid6 (3)	wmv	3	Video	0
vid3 (3)	mov	3	Video	0
vid3 (3)	mov	3	Video	0
vid3 (3)	mov	3	Video	0