

Open Source Software in Digital Forensics

Jorielle Scott

Auburn University

3101D Shelby Center

Auburn, AL 36849

+1 (334) 202-3806

scottjm@auburn.edu

Patrick Carpenter

Auburn University

3101D Shelby Center

Auburn, AL 36849

+1 (334) 328-9722

carpept@auburn.edu

ABSTRACT

In this paper, we discuss various digital forensics toolkits with an emphasis on their usefulness and admissibility in a legal setting. When using digital evidence in a legal case, one of the main considerations by all parties is to ensure that all software is properly licensed. There are many misconceptions about open-source software; among these are that "open-source is less secure", "open-source is less functional", and "open-source software does not require a license." The output

generated by a digital forensics tool is tested before admitting it into evidence by applying Daubert guidelines to measure its reliability. This paper will thoroughly examine the Daubert guidelines in an attempt to answer the question of whether or not open-source digital forensics tools better satisfy the Daubert guidelines than do closed-source tools. We begin by discussing relevant background information in open- source software and digital forensics. Then we survey open source toolkits used in digital forensics and compare them to commercially available products in terms of availability, usability, and reliability. We then list and discuss various legal arguments for and against the use of open source forensics toolkits in court proceedings, especially with respect to the Daubert guidelines. All the while, we attempt to bring into the discussion any relevant court decisions and endeavor to draw conclusions from these. We use these conclusions to argue that open source forensics toolkits better satisfy the Daubert criteria in some respects, and in other respects may rest on a somewhat shakier legal foundation than do closed-source tools. We conclude by acknowledging that open-source and closed-source tools both have a valuable role to play in digital forensics for the foreseeable future.

Categories and Subject Descriptors

K.4.2 [Computers and Society]: Social issues - abuse and crime involving computers.

K.5.0 [Legal Aspects of Computing]: General

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Conference'04, Month 1–2, 2004, City, State, Country.

Copyright 2004 ACM 1-58113-000-0/00/0004...\$5.00.

General Terms

Reliability, Legal Aspects, Security, Verification.

Keywords

Open source, digital forensics, GNU general public license, Daubert guidelines, evidence, software, commercial

1. INTRODUCTION

In a society where the notion of technology is becoming very innovative and widely adopted, the use of computers and other digital devices is easily accessible and affordable. As a result of the worldwide growth in technologies, new forms of crime and exploitations are also increasing. Therefore, it is becoming increasingly vital to government and law-enforcement agencies that they have the tools and techniques to combat high-tech crime.

To enact as a proactive defense against digital crime, digital forensic investigations are thoroughly exercised. When conducting a digital forensic investigation, investigators undergo three high-level phases. These phases include acquisition, analysis, and presentation. The acquisition phase is actually the most important phase because it saves the state of the system in its original, pristine form. This step is very important because if it is not exercised properly, evidence found may become inadmissible in court. The analysis phase actually takes the data that was acquired and thoroughly examines it to identify pieces of evidence. The evidence itself is categorized into at least three different types. Evidence types include inculpatory, exculpatory, and tampering evidence. Inculpatory evidence actually supports a given theory while exculpatory evidence contradicts a given theory. Evidence of tampering cannot be related to any theory, but it still shows that the system was tampered with to avoid identification. [4] The last phase in digital forensics, the presentation phase, focuses more on the

legal aspects of an investigation. The presentation phase is the phase in which conclusions are presented to the court based on the technical data found in the previous two phases. Although the concept of digital forensics is a fairly new one, it has become very important and relevant in society, due to the advancement in technology mentioned previously. It has been proven that the origin of the International Journal of Electronic Security and Digital Forensics (IJESDF) and its new regulations and proven methodologies in digital forensics have objectively benefited many real companies. These new methodologies and regulations have been proven to help companies with:

- Electronic security
- Anti-forensics
- Systems and network security
- Vulnerability research
- Ethical hacking
- Information security systems
- Authentication authorisations
- Security in mobile platforms
- Attack pattern recognition
- Strategic approaches to security
- Security policies and procedures
- Identity theft

- Identity management systems
- Identity and access management systems
- Phishing, pharming, spearphishing
- Cyber war
- Digital cities
- Criminal data mining
- Criminal network analysis
- Cybercrime detection and analysis
- Hidden data
- Steganography
- Computer, mobile device, network and software forensics
- Digital forensics tools and technique
- Testing and approvals process for forensic tools
- Crime scene/search and seizure processes
- Criminal investigation of mobile devices
- Investigative techniques and judicial processes
- Legal and ethical issues
- Cyber crime legislations

- Criminal intelligence
- Digital and physical surveillance
- Digital image manipulation
- Security requirements engineering

These benefits have led to tangible reductions in cyber crime, violation of company policies, fraud, hate crime, extremism, child pornography, and terrorism [31]. The facts have shown that digital forensics is a worthwhile pursuit.

The ultimate reasoning or goal behind digital forensics is to locate the criminally relevant information on a digital system. One of the most prominent tasks in the analysis of digital forensics is locating the relevant information within the computer system. The relevancy of the data being analyzed is often dependent upon the identification of the type of data being examined. However, due to the sensitivity and ability to easily modify digital devices, the procedures and techniques used to analyze in digital evidence is done with special digital forensic software tools. These tools are designed specifically for the analysis of files and data, the extraction of the data, and the preservation of the data that may be stored on investigated computers or other digital devices.

Generally, software can be classified in different ways, such as open source, closed source, commercial, or non-commercial. Due to various misconceptions, commercially-driven software programs are often thought of as having more credibility than open source software when dealing with issues of the court. However, through research, this paper provides relevant information to prove that open source software is reliable and creditable in digital forensics.

2. OPEN SOURCE – BACKGROUND

2.1 Introduction

Open source software (OSS) is commonly identified as software that has been licensed to allow its users to utilize the however they choose, to make changes to the program, and to redistribute such modifications all at no cost. "Free Software" is related to the freedom with which the software may be used, not the price. Although the term "free" software is used, a charge can still occur [16]. In other words, "free" software is "free" in the same way in which "free speech" is free, not the way in which "free beer" is (i.e., zero cost). In many cases, the development of open source software begins as part of a commercial project which allows peer production or allows its users to make changes and aid in the development of the project to help offset costs and maintenance.

Typical developers of OSS/FS are people from different organizations working together toward a common goal [41].

Open-source software is becoming broadly recognized for the opportunities it has created for information and communication technologies (ICTs). The popular interest in OSS has led to the evolution and ubiquity of the Internet as a virtual meeting place. This has compelled the Internet to stretch beyond the notion of being only a library source to become a forum for expression, allowing programmers to interact through their works. Due to the dominating factor of successful open-source development efforts, business and intellectual products are being threatened in terms of how they should be run, created, and protected. "OSS has affected the social organization of software production by expanding the realms of possibility, as well as offered individuals and institutions new options to choose from." [24]

Due to the fact that open-source software is generally known to be "free" in terms of cost, it is often misunderstood as being free software (FS). OSS and FS are not interchangeable. In FS, the term "free" is in reference to users' "freedom" when it comes to using, viewing and redistributing software. "FS is

understood within this context as relating to a user's freedom to run, copy, distribute, study, change, and improve software." [24] However, OSS is associated as being "free" in relation to having "no cost".

Figure 2.1 illustrates the relationship between different types of software.

Figure 2.1 - Illustrates the relationship between

software sources [24]

2.2 Open Source Software License

The General Public License (GPL) protects users of open source software by providing them with various rights to open OSS, including redistribution and modification. Most software licenses are established to take away one's ability to modify and share works. However, the GNU Public License was designed to have the opposite effect. "The GNU public License is intended to guarantee your freedom to share and change all versions of a program to make sure it remains free software for all its users" [16]. [...] Developers that use the GNU GPL may protect their rights with two steps: (1) assert copyright on the software, and (2) offer you this License giving you legal permission to copy, distribute and/or modify it." [16] The GNU GPL protects its developers and its users by explaining that there is no warranty for the software, and requiring modifications of projects to be marked as changed. This rule was established to ensure that no author or developer will be mistakenly accused for another author's or developer's version of the program if it achieves lower levels of usability, reliability or functionality.

3. USE OF OPEN-SOURCE SOFTWARE IN DIGITAL FORENSICS

3.1 Introduction

Open source software is categorized with a number of benefits including its being cheaper and faster; however, it is often misjudged in terms of performance and/or reliability. These potentially unjustified misgivings are especially detrimental in the field of digital forensics, where proven reliability is often

the standard to which software is held in courts of law. In this section, we explore some of the issues surrounding using open-source software in digital forensics investigations, and seek to address some of the concerns which most often lead individuals to question the reliability of open-source tools.

3.2 Digital Forensics Background

3.2.1 Introduction

Digital forensics is often identified as the underlying analysis and investigation of computers or digital devices that may have been used in criminal or related acts. The investigation may include retrieving data from existing or deleted files, extracting the contents of files and interpreting the meaning thereof.

Although the enforcement of digital forensics has been widely adopted in the criminal justice realm of society, the idea is fairly new. In fact, the practice of analyzing electronic evidence did not emerge until the early 1980s, and was due to the rapid increase in the ubiquity and degree of utilization of computers in homes and businesses [42].

The basic process of digital forensics includes preserving the state of the computer or digital device in the state in which it was found while at the crime scene, surveying the data for obvious evidence, searching for more detailed evidence based on the survey results, and using any relevant evidence found during this process to reconstruct events. Findings can be admissible per se as evidence or can be used by the digital forensics investigator to aid in forming expert opinion to be entered as testimony.

Skills for those who are considered digital forensics professionals include:

- Being able to identify electronic evidence associated with violations of specific laws.
- Being able to identify and articulate probable cause necessary to obtain a search warrant and recognize the limits of warrants.

- Being able to locate and recover relevant electronic evidence from computer systems using a variety of tools.
- Being able to recognize and maintain a chain of custody.
- Being able to follow a documented forensics process.

Personnel that deal with forensic evidence are categorized in three different categories. These categories include: Technicians, policy makers, and professionals. Technicians are those that carry out the technical aspects of gathering evidence. They must have sufficient technical skills to gather information from digital device, as well have a general understanding of software, hardware, and networks. Policy makers are those individuals involved in digital forensics that establish forensic policies that reflect broad considerations. Their main focus is generally on crime itself; however they must be familiar with computing and forensics. Professionals serve as the actual link or mediator between the policy and its execution. They must have extensive technical skills as well as a broad understanding of legal procedures.

3.2.2 Obtaining Digital Evidence

In the United States, the Fourth Amendment in the Bill of Rights guards against unreasonable searches and seizures, and additionally requires any warrant to be judicially sanctioned and supported by probable cause. It provides:

"The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized." [22]

Searches of computers or other digital devices can raise significant Fourth Amendment issues [2]. Computers have the capacity to store vast amounts of information which may include not only evidence of criminal activity but also private information that law enforcement should not be allowed to view. "Problems of capacity and privacy are compounded with networked computers in which a server may store personal information belonging to thousands of individuals." [2] If the owner of a computer is absent and has not given proper consent, law enforcement will generally need to obtain a warrant to conduct a search of the computer. The Fourth Amendment states that if there is "probable cause" a warrant may be issued which describes exactly what information is to be sought. The United States Supreme Court has determined that the probable cause standard is satisfied if the judge who issues a warrant is able "to make a practical, common-sense decision whether, given all the circumstances set forth in the affidavit before him, including the 'veracity' and 'basis of knowledge' of persons supplying hearsay information, there is a fair probability that contraband or evidence of a crime will be found in a particular place." [18]

The scope of the search authorized by a warrant is very important to the digital forensics examiner because the Fourth Amendment requires a particular description of what is to be searched. A warrant is legally invalid and not useful if it is so general that it would attempt to authorize a "general exploratory rummaging in a person's belongings [7]. Indeed, one of the Founding Fathers' primary concerns in the definition of the fourth amendment was to make illegal a certain British practice generally referred to as a bill of attainder, which is similar to the American concept of a warrant but which is not specific regarding the kind of evidence which is being sought or where the investigator can search. In any event, "[i]t is imperative for a forensic examiner to stay within the bounds specified in the search warrant." [2] The requirement that a forensic examination of a computer must not exceed the scope of the search warrant was emphasized in *United States v. Carey* [37]. "The search warrant in *Carey* authorized a search of the defendant's computers for 'names, telephone numbers, ledger

receipts, addresses, and other documentary evidence pertaining to the sale and distribution of controlled substances." The officers used key words to conduct a search of text-based files and found no files that were related to drugs. However, they did find numerous JPG files with sexually suggestive titles, and when they opened them, they found child pornography. The defendant was then prosecuted for possession of child pornography, but the appellate court reversed his conviction on the grounds that the officers had exceeded the scope of the search warrant by opening the JPG files." [2]

According to Adams:

"From the perspective of a digital forensic tool developer, the Carey decision highlights the desirability of designing forensic tools so that they record information about the steps taken during a forensic examination. Generating a clear audit trail will permit verification that the forensic examiner did not exceed the scope of a warrant while analyzing the data on a hard drive. An audit trail will also reinforce the reliability of the results from a forensic examination, because it will allow the forensic examiner's steps to be retraced so that the results of the forensic examination may be duplicated. While the Fourth Amendment requires a particular description of what type of information is to be searched for during a digital forensic examination, it does not require a description of the search protocol to be used during the examination [36]. It will generally be impractical for a warrant to specify a particular search protocol to be followed because information may be stored many different ways on a computer. Therefore, a forensic examiner will ordinarily not be limited to using a particular methodology. Even though a forensic examiner will generally not be required to follow a particular methodology, a recording of the examiner's methodology would still be desirable, because it would enable the examiner's analysis to be recreated and thus bolster its results." [2]

This is good to keep in mind for digital forensics, since so much of the digital forensics investigative process – and so much of the question of the use of software tools – depends on a clear understanding

of the legality, or rather the reliability from the Daubert perspective, of the methodologies used to collect evidence. In other words, while a warrant doesn't spell out the methodology which must be employed to carry out an investigation, it does not obviate the need for a proven process and set of techniques. Improper procedure and techniques can be just as damaging to the admissibility and weight as digital evidence as a violation of the scope of the warrant.

3.2.3 Examining Digital Evidence

The examination of digital evidence should always be done on data that has been acquired using accepted forensic procedures, such as safe cloning of the original evidence without altering it.

Extraction and analysis are concepts that are intended to assist the examiner in developing procedures and structuring the examination of the digital evidence. The term extraction refers to the interpretation of the recovered data from the media. Analysis refers to the interpretation of the recovered data and placement of it in a format that is logical and which may be useful. The selected examination approach taken in each criminal case is left to the discretion of the examiner. However, the following two steps should be considered when conducting and evidence examination: preparation and examination. Preparation entails preparing working directories on separate media to which evidentiary files and data can be recovered or extracted. Extraction can be either physical or logical. The physical extraction phase identifies and recovers data across the entire physical drive without regard to the file system. The logical extraction phase identifies and recovers files and data based on the installed operating system, file system, and applications.

"During the stage of physical extraction, the extraction of the data from the drive occurs at the physical level regardless of file systems present on the drive." [34] The following methods may be used during this phase: keyword searching, file carving, and extraction of the partition table and unused space on the physical drive. The keyword search allows the examiner to extract data that may not be accounted

for by the operating system and the file system. File carving utilities processed across the physical drive may assist in recovering and extracting useable files and data that may not be accounted for by the operating system and file system. Examining the partition structure may identify the file systems present and determine if the entire physical size of the hard drive is accounted for.

Figure 3.2.3 – Digital Forensics/Evidence Response Methodology. This is discussed in more detail later.

"During the logical extraction stage, the extraction of the data from the drive may include data from active files, deleted files, file slack, and unallocated file space. The steps may include:

- Extraction of the file system information to reveal characteristics such as directory
- Structure, file attributes, file names, date and time stamps, file size, and file location.
- Data reduction to identify and eliminate known files through the comparison of calculated hash values to authenticated hash values.
- Extraction of files pertinent to the examination. Methods to accomplish this may be based on file name and extension, file header, file content, and location on the drive.
- Recovery of deleted files.
- Extraction of password protected, encrypted, and compressed data.
- Extraction of file slack.
- Extraction of the unallocated space." [2]

Following all of these steps, or rather, taking all of these points into account during a digital forensics investigation can go far in ensuring the digital forensics examiner does not miss out on any potentially useful digital evidence.

3.3 Challenges in Digital Forensics

3.3.1 Introduction

Although digital evidence has proven to be extremely helpful in solving various crimes, there are challenges that arise when trying to analyze digital evidence. One of the challenges that is faced in digital forensics is the identification of file types [27]. Other challenges include but are not limited to the fact that digital evidence is latent in nature, fragile and time sensitive, can be easily destroyed and modified, can be easily mishandled, can be easily misinterpreted, and is often identified as misleading.

3.3.2 Easily Manipulated

Digital evidence can easily be manipulated in a sense that it can be changed by the suspect during the actual criminal act, or accidentally altered by investigators while collecting evidence. As shown in Figure 3.3.2a – 3.3.2b, opening a file to view or copying a file can easily alter the original file access time and dates. Both situations could occur without any obvious signs of distortion. According to Moody et al. [27], "[...] digital evidence has several features that mitigate this problem, which include:

- Digital evidence can be duplicated exactly and a copy can be examined as if it were the original. It is common practice when dealing with digital evidence to examine a copy, thus avoiding the risk of damaging the original.
- With the right tools it is very easy to determine if digital evidence has been modified or tampered with by comparing it with an original copy.
- Digital evidence is difficult to destroy. Even when a file is "deleted" or a hard drive is formatted, digital evidence can be recovered.

- When criminals attempt to destroy digital evidence, copies and associated remnants can remain in places that they were not aware of."

Figure 3.3.2a – A digital forensics tool

showing files in an initial condition.

Figure 3.3.2.b – The same digital forensics tool showing files after opening and viewing. This demonstrates that evidence can be destroyed or changed, if proper care is not taken.

Figure 3.3.3c – The same digital forensics tool showing files after saving. Timestamps of files containing evidence can be irrecoverably lost in many circumstances.

Given these facts, it is not unreasonable to hold out hope of finding digital evidence even when it is expected that anti-forensic techniques have been employed.

3.3.2.1 Case Study: Blanton 1995

In the celebrated North case, Rosenbaum reports in the New York Times that:

"When Colonel Oliver North was under investigation during the Iran Contra affair in 1986, he was careful to shred documents and delete incriminating emails from his computer. However, unbeknown to him, electronic messages sent using IBM Professional Office System (PROFS) were being regularly backed up and were later retrieved from backup tapes." [31]

This is a prime example of how attempts to conceal digital evidence of crimes can fail due to a lack of understanding or knowledge of all the places where digital evidence can remain. Even though North attempted to hide his tracks, digital forensics investigators were able to find the incriminating messages in back-up logs.

3.3.3 Circumstantial

Digital evidence is usually circumstantial, making it difficult to attribute computer activity to one individual in particular. Because of this, digital evidence can only be one component of an investigation. Basing an entire investigation on one form of digital evidence, such as a time stamp, is automatically identified as a weak case and could easily be argued against. [13] In fact, the authenticity argument against the introduction of digital evidence is a fairly common one and will be discussed in a later section.

3.3.3.1 Case Study: Grant 2000

In another celebrate case, we recover the following:

"In an investigation into the notorious online Wonderland Club, Grant argued that all evidence found in his home should be suppressed because investigators had failed to prove that he was the person associated with the illegal online activities in question. However, the prosecution presented enough corroborating evidence to prove their case." [31]

The necessity to provide corroborating evidence is one which stems from the fact, mentioned above, that digital evidence is inherently circumstantial. It can be used to strengthen a case, but attempts to use digital evidence to make a case can fail.

3.4 Difficult to Handle

Digital evidence can be a slippery form of evidence in a sense that it can be in layers, making it difficult to handle. As demonstrated in Figure 3.4, only a small portion of the blended evidence may be relevant to the use of a case, therefore making it necessary to extract useful pieces of evidence, fit them together, and then translate them into a form that may be interpreted [8].

Figure 3.4 – Evidence relevant to the case must be extracted and unpacked from the large amount of raw data. [8]

A common issue with digital forensics investigations is dealing with the explosion of information when trying to consider all potentially useful digital media. With the density of media on the rise, this trend shows no sign of slowing or changing directions. Better tools and methodologies are needed to cope with the ever-increasing amount of data which must be analyzed and explored.

3.5 Data Identification by Statistical Analysis

The type of data being analyzed in a criminal investigation is often a factor to be considered when determining relevancy of the information in terms of its admissibility and weight in a legal setting [27]. The identification of general file types is based on typical techniques such as file extensions or magic numbers/keys. Generally, file extensions are applied to indicate the format of the contents within the file. The development of Statistical Analysis Data Identification (SADI) provides the capability to identify what digitally stored data actually represents and can also allow for the selective extraction of portions of the data for additional investigation. SADI is able to be successful with such capabilities by applying statistical analysis of the byte values of the data in such a way that the accuracy of the technique relies on the values of the data itself, rather than on any potentially misleading metadata.

3.6. Open-Source – Legal Arguments

3.6.1 Introduction

The common goal of digital forensic analysis is to identify digital evidence for an investigation. When using digital evidence or relevant expert opinion retrieved using digital forensics tool, the reliability of the output that is produced from the tool is determined by the judge instead of a jury. The judge's decision is made in a pre-trial hearing referred to as a "Daubert Hearing", which will be explained later

in the paper. It is the main responsibility of the judge to identify whether the methodology and techniques that were used to gather the evidence are relevant and reliable [4]. The role of scientific gatekeeper attributed to judges during a Daubert hearing has been hotly contested by some legal and scientific experts, and some of these issues are explored later on.

Example Open source software tools include:

- Advanced Forensic Format (AFF) – AFF is an open and extensible file format designed to store disk images and associated metadata. This site also lists tools that work with AFF.
- Autopsy – The Autopsy Forensic Browser is a graphical interface to the command line digital investigation analysis tools in The Sleuth Kit. Together, they can analyze Windows and UNIX disks and file systems (NTFS, FAT, UFS1/2, Ext2/3).
- The Coroners Toolkit – TCT is a collection of programs by Dan Farmer and Wietse Venema for a post-mortem analysis of a UNIX system after break-in.
- Mac-robber – mac-robber is a digital investigation tool that collects data from allocated files in a mounted file system.
- Live-view – Live-view is a Java-based graphical forensics tool that creates a VMware virtual machine out of a raw (dd-style) disk image or physical disk. This allows the forensic examiner to "boot up" the image or disk and gain an interactive, user-level perspective of the environment.
- Open Source Digital Forensics - This site is a reference for the use of open source software in digital investigations (a.k.a. digital forensics, computer forensics, incident response). This site is a tool repository for Open Source tools on both Windows and Unix platforms.

- Open Computer Forensics Architecture – The main goal is to automate the digital forensic process to speed up the investigation and give tactical investigators direct access to the seized data through an easy to use search and browse interface.
- Sleuth Kit – The Sleuth Kit (previously known as TASK) is a collection of UNIX-based command line file and volume system forensic analysis tools.
- TULP2G – TULP2G is a .NET 2.0 based forensic software framework for extracting and decoding data stored in electronic devices. Along with the framework this version includes several plug-ins in the area of retrieving data from mobile phones.

3.6.2 Reliability

The United States Supreme Court has ruled in many cases that the trial judge must serve as a gatekeeper in digital forensic cases to ensure that all scientific evidence is not only relevant to the case but that it is also reliable [32] [2]. "The identified factors that could become the determining decision maker for the judge include:

- Whether or not the theory or technique on which the evidence is based has been tested;
- Whether or not it has been subject to peer review and publication;
- The known or potential rate of error and the existence and maintenance of standards that control the technique's operation,
- Whether or not the theory or technique used enjoys general acceptance within the relevant scientific community [2]."

Later, in the case of *Kumho Tire Co. v. Carmichael*, [21] the Court emphasized that these factors are not a definitive checklist. The Court observed that the absence of peer review and publication would not

necessarily disqualify scientific evidence, while the general acceptance of an unreliable technique would not make the results obtained from its use admissible. Instead, trial courts should focus on reliability as the touchstone for ruling on the admissibility of evidence 10". "A legal consideration for forensic tools is that they provide adequate safeguards for privacy interests.

Other requirements that are related to reliability may also affect the admissibility of evidence. Federal Rules of Criminal Procedure and the Federal Rules of Civil Procedure require a party to provide to other parties a pretrial disclosure of a summary of any evidence that is "based on scientific, technical, or other specialized knowledge." [2]

"Another issue related to reliability is the possibility that a court may require the defense counsel to be given an opportunity to review the source code for forensic tools." [2] The reason a court would order that defense counsel should be given access to the source code would be to enable them to assess the software's reliability. Similarly, the Supreme Court of Minnesota and the Court of Appeals of Kentucky have ruled that defense counsel in drunk driving cases should be permitted to obtain the source code for the Intoxilyzer 5000, which is an instrument used to determine whether a driver was intoxicated based on a breath test. [19] [38] Unfortunately, disclosure of the source code for a forensic tool might enable criminal suspects to avoid electronic surveillance by exploiting weaknesses in the software. In addition, disclosure could result in loss of trade secret protection. These concerns might be reduced to some extent by appropriate protective orders that require the defense not to disclose the source code to any other persons. [38]

3.6.3 Daubert's Reasoning of Reliability

The basis of the Daubert process is identified with four general categories that are utilized as basic guidelines when going through a procedure. The categories include testing, error rate, publication, and acceptance. General questions about these categories should be asked when assessing a digital

forensic procedure, such as: Has the procedure that was used been tested before? If so, does the procedure contain a known error rate? Has the procedure been published and subject to peer review? Lastly, is the procedure generally accepted in the relevant scientific community? Each guideline should be addressed in great detail in reference to digital forensics as a whole, then examined for both analysis and acquisition tools [4].

The two main categories of tests that should be performed on the tool output include both false negatives and false positives. Due to the fact that computers themselves are so complex, testing guidelines is a complex problem with digital forensics but it must be done to ensure if a procedure can provide accurate results or not. The false negative tests will ensure that the tool provides all available data from the input, while the false positive tests will ensure that the tool does not introduce new data to the output [4]. The false positive test is the more difficult category of test, while the acquisition tool is the easiest.

The two categories of errors that can exist in digital forensic tools include Tool implementation Error and Abstraction Error. Tool Implementation error is done in the software development stage which comes from bugs in the code or from using the wrong specification while developing the software. An Abstraction Error comes from the tool making decisions that are not one-hundred percent certain.

"Generally this occurs from data reduction techniques or by processing data in a way that it was not originally designed for." [4] It can be very difficult to maintain an accurate error rate in closed source software applications, due to the fact that if the bug was never identified to the public, it could be easily fixed and never added to the error rate. To add to the difficulty, commercial tools are driven by revenue and volume of sales, so publishing error rates are guarded topics because of the fear of sales being lost. However, when using open source software the error rate is easily maintained. Even if the

actual bug in the software is not documented, the latest source released can be compared with the previous one to find out which code contains modifications.

4. CLOSED-SOURCE FORENSICS TOOLKITS – AN OVERVIEW

4.1 Commercial Products

4.1.1 Introduction

This section discusses the development and features of some of the most widely-used proprietary digital forensics software toolkits in service today. Two particularly important tools, EnCase and Forensic Toolkit, receive detailed treatment. Other generally-available tools are also given brief mention. Some attention is also paid to the existence of forensics toolkits available only to law-enforcement agencies. Some familiarity with the available forensics tools will be useful later in the paper for grasping certain comparisons between open-source and closed-source digital forensics software tools.

The CFTT constructed a project that was executed to evaluate open source alternatives to ensure the accuracy and of open source software during a forensic investigation. [25]

The following categories were compared during the project:

- Calculates MD5 and SHA1 image
- hashes
- Provides hash value for individual files
- Verifies image integrity
- Finds deleted and encrypted files

- Identifies deleted and encrypted files
- clearly
- Recovers deleted files
- Identifies file extension mismatches
- Searches for strings (ASCII and
- Unicode)
- Includes HEX level viewer
- Organizes files into predetermined
- categories
- Provides an image gallery
- Shows file modified, accessed, and
- creation dates and times
- Logs investigator activity
- Identifies and analyzes slack/free space
- Finds and identifies overwritten files
- Finds cookies and URLs in registry
- Image import speed
- Initial import data

These are some of the many ways in which digital forensics software tools can be compared, although neither Manson et al. nor these authors mean to suggest that they are comprehensive.

4.1.2 Forensic Toolkit 3.0

Produced by Access Data, Forensic Toolkit 3.0 is marketed as being "the industry-standard computer forensics software used by government agencies and law enforcement around the world," and "available to all investigators and analysts, whether they are in law enforcement, education, a government agency, a Fortune 500 corporation, or performing digital investigations as a computer forensics service provider." [1] Desirable features of FTK include its being an integrated forensics solution, powerful password-recovery and decryption features, a responsive and full-featured GUI, robust operation, support for 32-bit and 64-bit Windows machines, RAM dump analysis, and many more. With abundant training options and a commitment to serving a widely-varying audience of forensics communities, FTK 3.0 is a powerful tool for digital forensics investigations.

Figure 4.1.2 – A screen shot of the FTK's GUI [25]

According to Manson et al. [25], some of FTK's product features include support for multiple file systems and disk image formats, tools for performing e-mail analysis, wide support for data file formats, including compressed files and password recovery, and advanced searching and indexing techniques, including cryptographic hashing of files. Deleted file recovery and MAC-time analysis are also standard.

4.1.3 EnCase

Produced by Guidance Software, EnCase Forensic Suite is marketed as being "used by governments, corporations, and law enforcement to conduct thorough, network-enabled, and court-validated computer investigations, e-discovery requests, internal investigations, regulatory inquiries, as well as

data and compliance auditing." [17] EnCase Forensic facilitates acquisition of digital evidence from many different kinds of sources, viewing of a wide range of file formats, analysis of data and reporting via automated tools. Benefits include powerful search and automation features through the EnScript scripting library and an unsurpassed record of court acceptance. With a variety of licensing and training options provided by Guidance Software, EnCase Forensic Suite can fit the needs of a variety of academic, industrial and governmental organizations.

Manson et al. [25] indicate that EnCase and FTK offer many of the same features, with some small differences (e.g., the EnScript automation language) which have already been mentioned. Compared to FTK, those authors find that FTK is generally more intuitive and easy to use, and provides a few more features (e.g., SHA1 hashing) than does EnCase.

Figure 4.1.3 – A screen shot of EnCase's GUI [25]

4.1.4 Other Publicly-Available Tools

Attempting to list all commercially-available tools which could be considered as digital forensics tools would be a long and tedious endeavor, and one almost certainly doomed to failure. Indeed, attempting to list just those tools which have been successfully used in court proceedings and which market themselves as digital forensics tools would require more effort from the authors to find, and more effort from the reader to comprehend, than the benefit from doing so would justify. As such, only a few of the other available tools are described at all, so as not to give the casual reader the impression that FTK and EnCase are the only commercially-available toolkits in existence (although, at the time of this writing, they certainly seem to dominate the market).

Paraben Corporation's P2 Commander offers many of the same features as FTK and EnCase, and with training, enterprise and support, and interoperability with EnCase's imager, it is in many ways a viable

digital forensics solution. [30] A screen show of this system's registry and system analyzer is given in Figure 4.1.4 for comparison purposes.

Figure 4.1.4 – Screen shot of P2 Commander's Registry and System Analyzer GUI [30]

The Datalifter Forensicware Solution is used by "federal, state and local governments from around the world," and is built on years of experience in performing digital forensics investigations. [9] The solution offers many of the same features as FTK and EnCase (indeed, the bundled tool ships with the FTK Imager), although generally with less enterprise support and fewer licensing and training options. The tool sells for a fraction of the cost of FTK and EnCase, however, and as such is targeted at smaller organizations or groups with forensic needs.

Digital Intelligence Software also markets a line of bare-bones software tools which allow users to image drives, view summary information about partitions, block writing and wipe digital media, all from an extended DOS interface. [12] The list of companies selling digital forensics toolkits goes on.

4.1.5 Tools for Law Enforcement: COFEE

Microsoft's Computer Online Forensic Evidence Extractor (COFEE) is marketed as being "designed exclusively for use by law enforcement agencies," and "brings together a number of common digital forensics capabilities into a fast, easy-to-use, automated tool for first responders." [26] Perhaps more surprisingly, "COFEE is being provided—at no charge—to law enforcement around the world." Note that a software's being provided free of cost does not uniquely determine its status as free and open-source versus closed-source and proprietary; indeed, COFEE is less "free" than other offerings, such as those listed earlier in this subsection, in that it is not available to the general public. In any event, this tool is interesting to note in that it is an example of a tool which does not fit neatly into either broad category of software, and some of the questions we will be addressing later on must be interpreted

somewhat differently for systems such as this. For instance, questions regarding the correctness of a tool faced with the reality that closed-source operating systems (such as Windows) are opaque to the forensics toolkit designer simply do not apply to Microsoft's COFEE, for obvious reasons.

4.2 Certification and Licensing

4.2.1 Introduction

This subsection describes certification and licensing issues of digital forensics investigators on the one hand, and of digital forensics software toolkits on the other. Certification and licensing are important concepts from a legal standpoint as they add to the credibility of and help establish precedent for the tools, techniques and organizations involved in the digital forensics process. As will be discussed later, certification and licensing are areas where the distinctions between free and open-source and closed-source and proprietary software are most pronounced.

4.2.2 Certification & Licensing of Investigators

There are myriad certifications available for digital forensics investigators, including those applicable to specific tools (e.g., FTK and EnCase) and those which are more generally applicable (e.g., Global Information Assurance Certification program's forensic certification and the Infosec Institute's certification).

Figure 4.2.2a, 4.2.2b, 4.2.2c – Digital forensics investigators can become certified either in specific tools (as in 4.2.2a) or more generally (as in 4.2.2.b and 4.2.2c). Courtesy of Guidance Software, Global Information Assurance Certification Program and the Infosec Institute.

As with all certifications, the primary reason for obtaining certifications in digital forensics is to demonstrate that a respected third-party organization recognizes and endorses your knowledge, skills and abilities and your status as a professional [6]. Some states require digital forensics expert

witnesses to possess certain certifications, and in any event possessing certain certifications can lead to better job opportunities and more job marketability. Additionally, the role of higher educational institutions in educating budding digital forensics investigative professionals should not be overlooked.

[5] With the growing job market and exponential increase in digital incident worldwide, academia can go a long way in increasing the status of and refining the scientific basis of digital forensics.

The Computer Forensics Tool Testing Program was established to provide researchers of computer investigation tools with measures of affirmation to ensure that the tools used provide accurate results.

"Accomplishing this requires the development of specifications and test methods for computer forensic tools and subsequent testing of specific tools against those specifications [35]. The CFTT is a project that is joined by the National Institute of Justice (NIJ), the Department of Justice (DOJ), and the information technology laboratory. It is supported by other organizations such as the Federal Bureau of Investigation (FBI), the U.S Department of Defense Cyber Crime Center, the U.S. Internal Revenue Service Criminal Investigation Division Electronic Crimes Program, and the U.S Department of Homeland Security's Bureau of Immigration and Customs Enforcement, U.S Customs and Border Protection, and U.S. Secret Service [35].

The results of the tests conducted by the CFTT [28] aid in improving the tools by providing developers with valuable information. They also provide information for users to make informed choices, and the community of legal advisors and others to understand the tools' capabilities. "The approach used in testing computer forensic tools is based on well organized methodologies for conformance and quality testing". [35] Figure 4.2.2 is an example of a test report. It shows the report completed in June 2008 from testing EnCase, version 5.05f, against the Digital Data Acquisition Tool assertions and Test Plan version 1.0.

Figure 4.0: Table 1 lists the features available in

EnCase and the linked cases selected for execution.

Table 2 list the features not available in Encase and

the test cases not executed. [35]

4.2.3 Certification & Licensing of Software

In order not to raise grave questions over the chain-of-custody of digital evidence, software tools used by digital forensics investigators often require some degree of certification. [39] This is an important fact and will be mentioned later in the paper, in that certification is a costly process and some capital is generally required to undergo the certification process. Closed-source and proprietary software can usually justify the expense of certification by using projected sales income; free and open-source software does not usually have such a sure source of income. Perhaps more importantly, closed-source and proprietary software has an inherently more centralized development effort and in that sense can control changes and versioning of the system to a far greater extent than is possible with free and open-source software. The implications of this will be elaborated upon in later sections.

In the United States, certification of software tools for digital forensics is handled primarily by the National Institute of Standards and Technology's Computer Forensics Tool Testing Program [28]. According to NIST, the program responds to a "critical need in the law enforcement community to ensure the reliability of computer forensic tools," and does so by "establish[ing] a methodology for testing computer forensic software tools by development of general tool specifications, test procedures, test criteria, test sets, and test hardware." The results of their tests and evaluations are freely available, as well as the detailed specifications against which the tools were tested. Among the reports of disk imaging tools are both open-source and closed-source offerings. Please consult NIST CFTTP for more information about the results of these and other evaluations.

5. RULES OF DIGITAL EVIDENCE

5.1 Introduction

This section describes some of the rules and judicial findings affecting the admissibility of and weight given to evidence in legal proceedings, with an eye towards applying this to digital evidence obtained in the course of a digital forensics investigation by an expert in digital forensics. We first trace major milestones in the development of legal opinion and court precedent concerning the admissibility of and weight given to scientific and/or technical evidence. Then, we enumerate some of the most common objections offered against digital evidence and addresses how various courts have responded to these.

5.2 History

5.1.1 Introduction

This subsection traces major milestones in the development of legal opinion and court precedent concerning the admissibility of and weight given to scientific and/or technical evidence. Specifically, the court decisions in two landmark court cases – *Frye v. United States* and *Daubert v. Merrel Dow Pharmaceuticals* – are cited and discussed in terms of how they might influence modern thinking regarding digital evidence. The Federal Rules of Evidence are presented and contrasted with the findings in the previous cases. The subsection concludes with a brief discussion of what "digital forensics" means, and what it ought to mean, in a legal context and, specifically, with respect to the rules governing testimony by expert witnesses.

5.1.2 *Frye v. United States*

The Frye v. United States case is an incredibly valuable source of information on the development of modern ideas concerning expert witnesses and their testimony. The interested court of appeals cites the following rule which describes the privilege afforded expert witnesses in offering opinions:

"The rule is that the opinions of experts or skilled witnesses are admissible in evidence in those cases in which the matter of inquiry is such that inexperienced persons are unlikely to prove capable of forming a correct judgment upon it, for the reason that the subject matter so far partakes of a science, art, or trade as to require a previous habit or experience or study in it, in order to acquire a knowledge of it. When the question involved does not lie within the range of common experience or common knowledge, but requires special experience or special knowledge, then the opinions of witnesses skilled in that particular science, art, or trade to which the question relates are admissible in evidence." [15]

In other words, expert witnesses are allowed to present testimony when it is deemed that reliable interpretation of evidence requires knowledge or skill available only through education, training and practice in such interpretation. For example, digital forensics examiners are allowed to interpret the results of digital forensics investigation because they possess specialized knowledge and skill which allows them to make reliable judgments regarding the meaning of digital evidence. Their opinions regarding the integrity, weight, and inculpatory or exculpatory nature of digital evidence is itself admissible. The main contribution of the Frye case lies in the court's decision regarding the appeal by Frye that a certain scientific test – an early lie-detector test – be admissible as evidence. The court's landmark decision has profoundly influenced admissibility standards for the introduction of expert witnesses and their opinions. From Frye v. United States:

"Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of

the principle must be recognized, and while courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, the thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs." [15]

This landmark decision implies at least the following about the testimony of expert witnesses not required by the rule cited earlier: the admissibility and evidentiary value of testimony offered by an expert witness is related to, if not determined by, its time-varying degree of acceptance or rejection by the scientific community (in other words, not all "science" is equal in the eyes of the law). The rule cited earlier requires only that the expert witness possess knowledge and training in an art or science relevant to the interpretation of evidence; it says nothing about which techniques or theories from said arts and sciences are allowed. From the point of view of digital forensics, this criterion might disallow testimony based on e.g. the procurement or analysis of digital evidence using techniques not generally accepted by the computer forensics community, such as a logical backup of the medium or an analysis which alters the medium.

5.1.3 Federal Rules of Evidence

The Federal Rules of Evidence (FRE) were adopted in 1975 and supersede the test outlined in the Frye decision (this is not to say that they are incompatible; indeed, the sections of the FRE dealing with expert testimony are in much the same spirit as the Frye test) [10]. Of particular interest to the question of admissibility and weight of digital evidence and the testimony offered by digital forensics specialists are rules 702 through 705. Rule 702 of the FRE deals generally with testimony of experts and is codified as follows:

"Rule 702 Testimony by Experts. If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an

expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise, if (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case." [14]

The requirements here are the following: that the expert witness must be qualified by knowledge, skill, experience, training, or education; that the basis of the expert's credentials is knowledge (i.e., of facts or assertions reasonably believed to be factual) of the scientific, technical or specialized kind claimed by the expert; that the expert's opinions must be relevant in the sense that they contribute to the set of facts the jury can reasonably be expected to use to determine the guilt or innocence of the defendant; that the expert must have access to enough information to draw the conclusion; that the expert must use the aforementioned knowledge correctly in forming the aforementioned opinions.

This represents a significant departure from the Frye rule – whereas the Frye rule requires that the technique used by the expert witness in forming the opinion presented as testimony to have gained "general acceptance in the particular field in which it belongs," the FRE requires only that the methods used be reliable. So in many ways, this is a step back towards the pre-Frye notion of admissibility in the sense that it shifts the focus from admissibility (determined by the Judge before trial) to weight (determined by the trier of fact, i.e. the jury based on the sum of all evidence admitted in the trial).

However, these rules do not change the fact that the Judge must ultimately decide what expert witness is admissible and what is not. This rule has clear implications on digital forensics: first, that digital forensics must be recognized as containing scientific, technical or specialized knowledge, which today is fairly well-understood given the large body of work on the subject; second, only testimony which assists the trier of fact in determining facts is admissible (e.g., timestamps might not be required to aid the jury in understanding the identification of illegal pornographic images the existence of which is demonstrated through other means, etc.); third, digital forensics examiners must be qualified by

knowledge, skill, experience, training or education (certification of computer forensic examiners will be discussed later); fourth, the testimony offered must be backed by enough data to make it seem reliable (e.g., if only parts of a deleted file are recoverable, or if they have experienced irreversible corruption due to other means, opinions based on these data may be thrown out); fifth, the data must be analyzed using approved and understood techniques and procedures (clearly this applies to the stages of the digital forensics investigation; to what extent this applies to software tools employed by the investigator in conducting the digital forensics analysis is at the heart of the question this paper addresses); finally, it must be demonstrable that approved techniques were applied to sufficient data in this case (e.g., logs of investigators' activities and a documented understanding of the processing of the data by the tools is required). To see why these last two are interesting in discussing digital forensics toolkits, consider Rule 705:

"Rule 705 Disclosure of Facts or Data Underlying Expert Opinion. The expert may testify in terms of opinion or inference and give reasons therefore without first testifying to the underlying facts or data, unless the court requires otherwise. The expert may in any event be required to disclose the underlying facts or data on cross-examination." [14]

Upon cross-examination, digital forensics investigators might be required to testify to the data and methods used to arrive at the opinions presented as evidence. This implies that the investigator has knowledge of the methods used to arrive at the evidence, which at its most fundamental level requires a knowledge of the internal structure of the software toolkit used in the digital forensics investigation. This is not explicitly addressed in the Frye rule, and as such represents an additional hurdle to the introduction of testimony by expert witnesses.

5.1.4 Daubert v. Merrell Dow Pharmaceuticals

This case is significant in that it recognizes the differences between the FRE and the Frye Test, favoring the former over the latter for determining admissibility of expert testimony in federal courts. The sentiment is captured neatly by the following:

"Cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof, rather than wholesale exclusion under an uncompromising "general acceptance" standard, is the appropriate means by which evidence based on valid principles may be challenged. That even limited screening by the trial judge, on occasion, will prevent the jury from hearing of authentic scientific breakthroughs is simply a consequence of the fact that the Rules are not designed to seek cosmic understanding but, rather, to resolve legal disputes." [10]

The Frye test is therefore recognized to be too restrictive in the sense that the question of admissibility is often better left as a question of weight to the trier of fact. The court does stress, however, several important factors which the Judge should take into account when acting as gatekeeper of scientific evidence [10]: first, whether the theories or techniques used to arrive at opinions falsifiable (e.g., the claim that a bitstream copy of a physical drive contains the same data as the original drive can be tested, or falsified, by doing a bit-by-bit comparison or by using hashes); second, whether the theories of technique have been subject to peer review or scrutiny of the community via publication (this is an interesting question at the heart of the question this paper seeks to answer: what implications are there for the publishing test within the Daubert guidelines for software toolkits embodying the techniques or theories used by digital forensics examiners?); third, whether the technique's error rate is reasonable (this can go beyond physical limitations and involve human error e.g. of the software development team in correctly encoding approved methods or of the digital forensics investigator using a system with which s/he is unfamiliar). The Daubert guidelines, therefore, relax the requirements on the admissibility of scientific evidence by expert witnesses, define a clear precedent

for the interpretation of the federal rules of evidence, and elaborate greatly on the Judge's role as a gatekeeper in determining what kinds of evidence are admissible.

5.1.5 Defining Digital Forensics

In their 2006 paper entitled "Building a National Forensics Case Repository for Forensic Intelligence," Biros and Weiser define digital forensics as "scientific knowledge and methods applied to the identification, collection, preservation, examination, and analysis of information stored or transmitted in binary form in a manner acceptable for application in legal matters". [40] There are several components of this definition which require careful analysis.

Figure 5.1 – The study of digital forensics is inherently multidisciplinary and involves scientific, technical, security and legal expertise. Courtesy of the ADFSL, <http://www.adfsl.org/>.

First, digital forensics is to be thought of as a collection of "scientific knowledge and methods". Connotations of the words "scientific" and "knowledge" are important from a legal standpoint, in that they define what forensic principles and techniques are allowed in legal proceedings (indeed, the connotations are discussed in detail in the opinion on the *Daubert v. Merrell Pharmaceuticals* case [10]).

Second, a digital forensics investigation consists of distinct activities - at least identification, collection, preservation, examination and analysis - and, under an appropriate ordering, these can be understood as constituting a digital forensics investigative process (discussed at length in Section III).

Third, digital forensics investigations deal with only certain kinds of information, in particular, that "stored or transmitted in binary form". This author finds this part of the definition somewhat unsatisfactory in that information stored or transmitted in binary form is not necessarily related to computing technology at all; for instance, if a student's answers to a true-false quiz on a sheet of paper,

this can be thought of as storing and potentially transmitting - e.g., to the grader - binary information which, in this author's opinion, falls - or ought to fall - outside the realm of digital forensics. Forgiving this lack of specificity, we understand the definition to limit digital evidence to evidence recoverable from a computer system or related electronic system.

Fourth, digital forensics concerns itself only with recovering digital data - according to a process based on sound scientific principles - in such a way as to preserve this data's evidentiary value in a legal setting. Therefore, techniques or approaches which render otherwise valuable digital evidence inadmissible in a court of law, including any scientific principles which require or are based in such techniques, fall outside the realm of digital forensics as defined by Biros and Weiser. This is an especially important distinction in that it is a (potential) limitation on the kinds of data which can be recovered, or on the ways in which this data can be recovered. For instance, while opening a web browser and viewing browsing history on an individual's system might be a useful way to recover data about the pages the individual has been viewing, this technique may result in a change of this data, thereby rendering it inadmissible.

Conducting a digital forensics investigation is challenging [20], and the challenge is understandable given the definition discussed above. Conducting a digital forensics investigation requires multidisciplinary understanding.

5.3 Objections to Digital Evidence

5.2.1 Introduction

This subsection discusses some of the most commonly-used arguments against the introduction of digital evidence obtained from digital forensics investigations in court proceedings. The two main categories of objections, those based on authenticity and hearsay, are discussed first. Then, other less-

frequently used objections are mentioned. This subsection is vital to the discussion of the relative merits of open-source versus closed-source software, as it provides some indication of the typical hurdles which must be overcome even after the credibility and reliability of the expert witness has been established.

5.2.2 The Authenticity Objection

One category of objection raised against digital evidence in court proceedings deals with the authenticity of the evidence. Here, authenticity refers both to the integrity of the evidence (in terms of tampering and/or misrepresentation) and to the authorship of evidence (in terms of the individual who created or caused to be created the evidence). [11] These issues are complicated due to the fact that digital evidence is, by comparison to physical evidence, very easy to alter, hide or otherwise destroy. In general, courts have not allowed the mere possibility of tampering or misrepresentation stand in the way of introduction of digital evidence - even though the possibility is clearly greater than for physical evidence.

5.2.3 The Hearsay Objection

Another objection which is raised against the introduction of digital evidence, or testimony based on the results of a digital forensics investigation, is that digital evidence inherently constitutes hearsay. According to the Department of Justice's "Searching and Seizing Computers and Obtaining Electronic Evidence Manual", hearsay is defined as "a statement, other than one made by the declarant while testifying at the trial or hearing, offered in evidence to prove the truth of the matter asserted." [11] Courts have often found that computer evidence is admissible by the hearsay exception of the Federal Rules of Evidence applying to "business records", i.e. documents produced in the normal course of conducting business. [14] In other cases, courts have overruled hearsay objections by observing that digital evidence generated automatically, i.e. by a computer program or associated machinery, does

not constitute hearsay since only statements made by human beings constitute hearsay. This is a significant observation for software developers in that it implies software is inherently trustworthy. This author is unconvinced, however, by this argument; to make the argument convincing, there needs to be a set of criteria for determining the degree to which the output of a computer program is to be considered hearsay. In a sense, evidence generated from hearsay evidence continues to be hearsay evidence so long as the truth or falsity of the derived evidence depends on the validity of the underlying hearsay.

5.2.4 Other Objections

Authenticity objections and hearsay objections constitute the vast majority of all objections to digital evidence presented by digital forensics investigators. [11] That being said, two other objections are sometimes made, albeit with even less success, in general, than the aforementioned ones: objections based on the "Best Evidence Rule" and those which contest that computer printouts are subject to the rules governing summaries.

In simple terms, the "Best Evidence Rule" usually requires that to prove the existence of a piece of evidence, the original must be presented as evidence. This is clearly problematic for digital evidence in that digital evidence is inherently unknowable except via representations (electronic or printed) thereof. That is to say, digital evidence can be thought of as consisting of streams of binary digits stored on a medium (magnetic, solid-state, etc.). These bit-streams cannot be viewed in their original form, at least not without requiring inordinate amounts of time and at considerable expense. Digital evidence is usually presented by allowing computer programs to interpret the evidence and produce human-readable output on an electronic display or via printing hard copies. The basis of this argument against digital evidence is that a picture of a murder weapon is not the same, and does not carry the same weight or meet the same admissibility requirements, as the physical weapon itself; by the same

token, then, ought human-readable representations of digital evidence be treated in the same way? Courts have usually overruled such objections on grounds that the digital evidence itself is more akin to a photographic negative, and that printouts are more like photographic prints. Given the common usage and practicality, then, computer representations of digital evidence are as admissible as the original. This rule also applies to copies of the original, so long as the copying can be demonstrated to be reliable. This gives an indication of why digital forensics examiners often go to considerable lengths to ensure and document the faithfulness to the original of all copies of the media they investigate.

It is sometimes contested that computer representations of digital evidence constitute summaries of the digital evidence, in that they take digital evidence in raw form and produce human-readable content that is usually a meaningful digest of or simply a useful subset of the sum of all evidence. In general, representations of digital evidence have not been, per se, considered summaries in most cases [11]. However, summaries of other admissible evidence – such as summaries or automated analyses of evidence by forensics toolkits – are subject to restrictions in the Federal Rules of Evidence [14].

6. SOFTWARE AND THE DIGITAL FORENSICS PROCESS

6.1 Introduction

This section outlines and discusses the digital forensics process, which consists in identifying, acquiring, analyzing and presenting evidence during a digital forensics investigation, and what role software tools – be they open-source or proprietary – play in this process. We first follow the forensic investigative process from identification and preparation through reporting. We then explore what roles software products play, if any, during each stage of the forensic investigative process. In so doing, we will illustrate with examples and, possibly, relevant examples from the legal or academic literature the ways in which these investigative stages manifest themselves in the field. A thorough understanding of

these stages will be useful in Section 7 when discussing the technical and legal issues associated with closed-source versus open-source software.

6.2 The Digital Forensics Process

6.2.1 Introduction

This subsection provides a general-purpose overview of the digital forensics investigative process. The understanding recoverable from this overview will be vitally important in understanding how software assists the digital forensics investigator in conducting his investigation. Note that this implies more coverage than a simple discussion of the ways in which software assists the investigator at various stages, since the role of software in some stages may be unimportant or coincidental rather than fundamental. Included are some of the most important techniques and practices governing the execution of each of the phases. Given that Section 7 compares open-source and closed-source digital forensics software toolkits in terms of these stages of investigation, these techniques will certainly come to bear.

Figure 6.2.1 – A simplified model of the digital forensics investigative process. Note that, as with any sufficiently complex process, feedback and several iterations are generally required. Courtesy of National Digital Forensics.

6.2.2 Identification and Preparation

As reported by Ami Narh et al. [Ami Narh], the identification and preparation phase occurs when an "incident is recognised as needing investigation." This occurs after irregularities or other potential evidence of crimes has occurred. Clearly, this step must occur before any digital forensics investigation can begin. The input at this stage of the process is in the form of external stimuli (initial indications that some incident worth investigating has taken place) and the output is a set of plans, documentation and

shared understanding regarding the nature, scope, and importance of the investigation. The primary concern of this phase is in developing enough understanding – in terms of where to look, who to look at and what to look for – to be granted a lawful warrant to search for and seize digital evidence. In a sense, this can be thought of as occurring before the digital forensics investigation takes place, although this author prefers an inclusive approach.

This is the part of the digital forensics investigative process where questions of jurisdiction and positive identification come into play. [Ami Narh] In cases where a digital forensics investigation is conducted pursuant to a non-digital crime, this is not generally at issue. When the case involves application of laws concerning cyber crime, however, it must be remembered that regulations differ from region to region. Given the largely anonymous nature of the Internet, these kinds of investigations can be difficult to get off the ground.

6.2.3 Search and Seizure

The search and seizure phase consists of obtaining necessary search warrants, as well as formulating and carrying out plans to obtain as much relevant digital evidence as possible with as little negative impact on the owners of hardware systems as is possible. [Ami Narh]. This occurs after an initial plan of action has been formulated in the identification and preparation phase. As with the identification and preparation phases, this is sometimes not considered as part of a digital forensics investigation.

However, since it is clearly a required step in the broader legal process, this author feels that inclusion is mandatory when discussing the legal issues surrounding digital forensics investigations, which is this paper's primary topic. The input to this phase consists of the plans, strategies and understandings developed in the previous phase. The output from this stage is a set of digital media and/or related property which is believed to contain valuable digital evidence, obtained legally and in as non-invasive a manner as possible from owners thereof.

6.2.4 Preservation

The preservation phase consists of developing plans and using techniques to ensure that the set of media obtained as the result of the previous phase retains its evidentiary value in potential court proceedings and that the loss of valuable digital evidence is reduced to a minimum. This stage includes making admissible copies of the digital evidence on seized devices and verifying the authenticity of the copied information. This is a crucial phase in the digital investigative process in that a technical or human error at this stage can partially or completely invalidate findings by the investigator in later stages. This stage occurs immediately after the search and seizure phase. As investigators seize media, preservation begins immediately so as to minimize the impact on owners and ensure the highest levels of data integrity. The input of this stage is a stream of media under investigation, and the output is an assurance that the integrity of the seized media is maximal (or, at least, maximal subject to realistic constraints).

What makes preservation interesting from a digital forensics standpoint is the dynamic and transient nature of electronic computer systems. For instance, memory in RAM can be lost and digital evidence can be altered or destroyed without trace. [3] Special techniques employed at this phase can help prevent forensic examiners from inadvertently rendering potential evidence inadmissible, but the problem of deliberate tampering – on both the prosecution and defense sides – is more a much more serious one which, in some instances, may not be solvable. Anti-forensic techniques are all the more common among those being investigated of committing computer crime, artificially increasing the seriousness of this problem for digital forensics investigators.

6.2.5 Examination

The examination phase consists of duplicating the preserved media and mining it for the information of interest, i.e. the information spelled out in the warrant. [3] This is one of the most widely-recognized

phases of digital forensic investigation and is one for which many of the tools most clearly motivated by digital forensics exist. This phase occurs after preservation of the media and is the first step in collecting admissible digital evidence. The input to this phase is a preserved set of digital media from which all possible digital evidence will be collected. The output of this phase is a set of digital media containing a copy of the original data (made in such a way as to ensure digital evidence derived therefrom will defeat any authenticity or "Best Evidence" objections it encounters in a court of law) and a function with all the potentially useful digital evidence discovered in the copied data as the domain and the corresponding locations of these pieces of potential evidence as the range.

6.2.6 Analysis

The analysis phase consists of defining a function which is a subset of the function output as a result of the examination phase and in developing a set of conclusions based thereon. Specifically, this function will consist only of those (evidence, location) pairs which correspond to useful evidence, and the conclusions will be made in accordance with approved digital forensics techniques (so as to satisfy the FRE) and support testimony given as expert witness opinion. [Ami Narh] The input to this phase is the set of media and function defining all potentially useful digital evidence, both of which were generated as output during the examination phase. The output of this phase consists of the set of digital evidence and conclusions to be presented in the form of evidence and expert testimony, respectively, in court or legal proceedings.

6.2.7 Reporting

The reporting phase consists of taking everything of interest that has been learned during the digital forensics investigative process up to this point and presenting these findings in such a way as to convince a judge and jury of peers that the evidence being presented meets standards for admissibility outlined in Section 5 and weighs heavily enough on one side of the issue as to influence the court's

decision. The presentation must be made in such a way as to be comprehensible to the jury (as discussed earlier), that is to say, the testimony must be understandable to an individual without technical or scientific expertise. Additionally, this phase entails extensive preparation on the part of the legal team and expert witness in preparing to overcome common challenges to the introduction and weight of expert testimony, some of which were discussed in Section 5. The input to this stage consists of the body of digital evidence and derived expert witness testimony, along with an understanding on the part of the expert witness of the scientific and professional framework of digital forensics. The output is the presentation of these facts and opinions which meets with the expectations of the legal system and which assists the legal team in proving its case.

Figure 6.2.7 – A more detailed diagram describing the digital forensics process, with explanations and emphasis on documentation. Courtesy of [3], with adaptation.

6.3 The Role of Software Tools

6.3.1 Introduction

This subsection uses the framework established in the previous subsection to explain the ways in which software tools are used at various stages of a digital forensics investigation. Software tools, after all, exist to extend and facilitate the capabilities of the investigator, and as such reflect the same techniques and principles used by the investigator.

Figure 6.3.1 – The idea behind forensics tools is not a difficult one: sophisticated tools plus a massive amount of potential evidence can lead to human understanding. As with most things, the devil is in the details. Courtesy Microsoft Clip Art Gallery, with adaptation.

6.3.2 Identification and Preparation

This stage of the digital forensics process is largely non-technical, and such the role of software tools in this stage of the digital forensics process should not be overstated. Although lots of software is involved in this stage – office applications, networked applications (such as web browsers and e-mail clients), intrusion-detection systems, etc. – at least this author does not consider such software to fall under the umbrella of software directly applicable to the digital forensics investigative process. Of the software that is used at all, the category that comes closest to being useful to the digital forensics investigation directly (as opposed to indirectly, e.g., an e-mail client to collaborate with colleagues) includes detection systems (IDS, AntiVirus, Firewall, etc.) which may produce enough useful evidence to prove useful in its own right, or in obtaining a warrant to search and seize pieces of media which may contain other more inherently valuable information. Although such kinds of software have been and should be considered in a discussion such as this, we have elected to focus on more widely-recognized phases of the digital investigative process.

6.3.3 Search and Seizure

This stage of the digital forensics investigative process is, to the best of the authors' knowledge, a mostly law-enforcement and administrative activity not requiring the use of any specialized software tools. Therefore, we will not directly discuss this stage for the remainder of the paper, despite its overarching importance in any effective investigation.

6.3.4 Preservation

This is the first stage of the digital forensics investigative process in which a clear need for specialized, reliable, and effective software (and possibly hardware) tools are required, or could be used to great effect. For instance, non-invasive software/hardware tools can be used to capture the contents of physical memory before attempting to attempt offline the more lengthy process of creating a complete bit-stream copy. Software tools can be used to analyze seized devices for evidence-

destroying malware (viruses, rootkits, etc.) and to save as much information as possible about the history of interactions with a network. Hardware and/or software write blockers can be used to ensure that no changes are made to certain pieces of media. Software tools can be used to help document the activities performed at this stage so as to provide a record of the chain of custody, which will be useful in establishing the use of proper procedures.

6.3.5 Examination

This stage is a heavy hitter in terms of the importance of how software tools are used by the digital forensics investigator. Specialized tools are needed to ensure that devices containing large amounts of nonvolatile memory are copied in such a way as to preserve evidence. Usually, bit stream imaging is performed to ensure that all potential evidence, including that which may have been hidden either intentionally (e.g., by writing directly to slack space or unused areas of the disk) or unintentionally (e.g., by deleting a file) are present in the forensic copy. Software is then used to find evidence outlined in the warrant (e.g., searching is used to find files) and specialized techniques are used as necessary to make positive identifications of data (e.g., decompression of zipped directories, password recovery and/or decryption, identification of applications of steganography used to hide data in images, etc.). Tools not only recover file data, but also MAC times and cryptographic hashes for identification and documentation purposes. At this stage, software allows for a truly massive amount of information to be recovered from media.

6.3.6 Analysis

This stage of the forensics process can also benefit greatly from the appropriate use of specialized software tools. Software excels at taking a large amount of data and processing it to produce more immediately meaningful results. For instance, files can be sorted according to size, percentage match on a search, file type, MAC time, etc. This information can be used to help the investigator in forming

opinions about the data and implications this may have on the case. Issues surrounding summaries may complicate the use of some of the tools in this phase, but as long as the evidence itself isn't presented as evidence, few problems should be encountered.

6.3.7 Reporting

As is the case with the search and seizure phase, the authors do not have much to say regarding the use of specialized software tools during this stage. By this stage in the investigation, if all has gone well, the evidence has been extracted from the media and the technical activity is largely complete. In many ways, reporting is the single most important part of the entire process; however, since we are limiting ourselves to a technical discussion on the interplay between technical and legal factors in the digital forensics investigative process, we will not pursue this stage of the process any more than has already been done.

7. OPEN-SOURCE VERSUS CLOSED-SOURCE DIGITAL FORENSICS TOOLS

7.1 Introduction

This section seeks to address the fundamental thesis of the paper by utilizing the framework defined in previous sections to compare and contrast open-source and closed-source (proprietary) digital forensics toolkits in legal and evidentiary terms. We first summarize some of the arguments for and against free and open-source software (FOSS) in a general context. We then look at specific issues related to the use of FOSS during the relevant stages of the digital forensics process, and weigh the advantages and disadvantages of free and open-source software against those of closed-source and proprietary software from a legal viewpoint. In so doing, we will conclude that the advantages of FOSS far outweigh the disadvantages, and indeed that in many respects the balance is tipped in favor of

FOSS. However, we will admit limitations of FOSS from a digital forensics standpoint when such admission is due, and suggest potential avenues for alleviating this

7.2 Free and Open-Source Software (FOSS)

7.2.1 Introduction

This subsection exposes some of the debate surrounding and tradeoffs involved in the use and development of free and open-source, versus closed-source and proprietary, software. Generally speaking, arguments for and against free and open-source software and closed-source and proprietary software can be roughly categorized as philosophical, economic, and legal, depending on whether they have to do with ethical foundations, reliability and availability or licensing and usage, respectively. Since digital forensics software is a proper subset of the set of all software tools, these issues apply in broad terms to the thrust of this paper and will be revisited in the next subsection.

7.2.2 Philosophical Issues

Up until the early 80s, free and open-source software was the rule, not the exception. [23] Software was commonly given to customers by vendors of computer hardware in an attempt to make their systems more marketable. Since then and to this day, the Free Software Foundation (FSF) has been standing by what they consider to be the individual's right to "use, study, copy, modify, and redistribute computer programs," despite the trend towards proprietary and closed-source software. [33] The FSF, then, makes the issue of free and open-source software an issue concerning rights, which is a classic and well-studied problem in the field of ethics. The Open Source Initiative (OSI) takes a slightly different tact, attempting to argue for the use of FOSS from the quality standpoint. [29] However, this is simply an ethical argument in disguise, in that its true basis lies in the utilitarian ethical philosophy that holds whatever creates the most utility is best (i.e., since FOSS leads to higher quality

software, and higher quality software is good for society, FOSS is good in an ethical sense). This is subtly distinct from the FSF's position, which holds that the openness of software is inherently good, per se (and is, therefore, a deontological ethical argument).

Figure 7.2.2 – The debate between Linux and Windows is one of the most visible fronts of the open-source debate.

7.2.3 Economic Issues

According to Vlastos et. al [39], "[t]he open source model works best when the users of software are also developers, and when there are enough of them to sustain and share the development and maintenance of effort." They go on to say that "[s]mall, highly specialised non-technical (or, more precisely, non-programmer) user communities are served by closed source software manufacturers who make, sell and support high margin, low volume, niche software." The authors of that paper point out that the digital forensics community is more nearly described by the latter of these characterizations, and we agree. As digital forensics matures as a field, we anticipate that the divide between software developers on the one hand, and digital forensics investigators on the other, will continue to grow. We therefore expect that closed-source software will actually increase in use in digital forensics, rather than decrease.

7.2.4 Legal Issues

Many legal issues involving FOSS can be stated in terms of licenses and licensing. This is due in part to the fact that there are many different such licenses, each with its own legal issues.

Figure 7.2.4 – Open source means different things to different people. This table compares several common open-source licenses. Courtesy of Shafqat Ahmed in 2008 (<http://www.shafqatahmed.com/>)

Simply stated, open source software licenses are generally such that there is a lack of centralized, coordinated control over the development of a software product, particularly in terms of branching or diverging behavior. [23] It is at least partially for reasons like these that software controlling elevators, airplanes and pacemakers is almost always closed-source; certain standards for reliability and responsibility (or, perhaps more precisely, liability in the event of failure) apply in these cases. FOSS licenses are generally more concerned with keeping licensed software and derivative works free than they are with establishing guarantees of merchantability for an express purpose.

7.3 Open-Source Software for Digital Forensics

7.3.1 Introduction

This subsection is the culmination of all the foundation laid in this paper in that it finally addresses the paper's fundamental question: what are the benefits of free and open-source software versus closed-source and proprietary software in the context of digital forensics? This question is addressed by considering the differences between free and open-source software and closed-source and proprietary software at each stage of the digital forensics investigative process. This author recommends careful study of the previous material before attempting this section, as the arguments are based directly in the facts and discussion conducted in earlier sections. The authors also recommend that the interested reader consult the referenced papers for more information on the subject.

7.3.2. Arguments for Open-Source

The arguments for the acceptance and use of open-source software in the digital forensics investigative process, in general, hinge on the idea that open-source leads to an inherently better understanding of how digital forensics toolkits actually work, in several dimensions: for peer review, for admission as evidence or in response to a request by an opposing legal team, and to accommodate

groups with dynamic needs, such as for educational or training purposes in an academic setting. Each of these dimensions bears some additional discussion.

Open-source digital forensics software is inherently more open to peer review, and as a result, is ultimately more suitable for use according to the Daubert test [10] according to the FRE [14]. This is, in our opinion, the single best argument for the use of open-source digital forensics software as opposed to closed-source digital forensics software. To gain general acceptance in the community, software's source code must be exposed to examination by experts. Such examination is by definition restricted in the context of closed-source software, and usually greatly so. Fundamental questions of the soundness of underlying scientific/technical principles, reliability, error rate, etc. of closed-source software can only be answered via black-box software testing strategies. Since black-box software testing cannot test the logic of the software itself (in terms of, e.g., branch coverage, condition coverage, etc.), common test suites may give misleading results for certain unusual, exceptional, or boundary cases. In a digital forensics investigation, these cases are precisely the one which should require the strictest testing, especially if identifying and overcoming anti-forensic techniques is an issue.

Open-source digital forensics software is inherently safer to rely on in a legal setting, since requests by the court to examine the tools under the rules of discovery is guaranteed not to encounter significant resistance. [4] It has been held in some courts that software used in preparing, presenting, extracting or otherwise processing evidence can be subject to the rules of discovery in that the source code of such software can be requested by the court in order to determine its adherence to scientific, technical, etc. standards for admissibility. Although courts understand and in general will take reasonable steps to protect the proprietary interests of software companies, some companies have been – and, at any rate, could be – unwilling to cooperate with such requests. Open-source software does not suffer from

this limitation since, by definition, the source code can be made readily available at any individual's request.

Open-source digital forensics software is easier to deploy in a variety of organizations due to the inherent nature of open-source software licenses. This is especially true for dynamic organizations, such as for education or training of digital forensics investigators in an academic setting. [5][25] With the growing need to educate and train digital forensics investigators, and with the growing need for such experts in general, many organizations – academic, private and governmental – will require quickly expanding capacity to use digital forensics software. With closed-source software, licensing can become an issue if multiple copies of the software will be used, particularly if done so in a distributed setting. Since open-source software packages are less restrictive in this respect, many have found open-source digital forensics attractive for small and/or dynamic organizations.

Open-source encourages involvement in developing and expanding software, and this can lead to a better understanding of the tool and its limitations. In the context of digital forensics, this can reduce the error rate due to investigator error and help the expert witness demonstrate expertise under cross-examination [that guy who had a website... find his name]. After digital evidence has passed the Daubert test and is admitted in court, the question of weight given to that evidence must be decided. A tactic typically used by legal teams when faced with the damning testimony of an expert witness, and especially in the case of a digital forensics expert, is to try to cast as much doubt on the meaning and significance of the findings as possible. One particularly effective way of accomplishing this can be to expose the expert's lack of understanding of the tools (analytic or, in our case, software) which were used in forming expert opinions. Open-source software permits and, in a sense, encourages the thorough understanding of the underlying logical and procedural mechanisms employed inside tools,

and as such its use can have beneficial effects during the presentation phase of the digital forensics investigative process.

7.3.3. Arguments for Closed-Source

The arguments for the continued dominance of closed-source software in the digital forensics investigative process, in general, hinge on the idea that closed-source software is inherently better in terms of usability and support and that this leads to a reduction in the error rate and to an increase in the perceived reliability of such systems. Several of these aspects require closer inspection.

Closed-source digital forensics software has traditionally enjoyed richer, more intuitive graphical user-interfaces (GUIs), and this can lead to a reduction in the error rate from using such tools, in addition to a lower learning curve for non-programmers. [25] Many researchers, developers and users of open-source and closed-source software have noted that closed-source software is generally more user-friendly than open-source software, and therefore appeals to a broader, non-technical user-base. The situation is not so different in the realm of digital forensics software. Software which is easier to use out-of-the-box, and which requires less configuration and manipulation to use, is easier for non-programmers to learn to use. This means that costs – time and money – associated with training of personnel are reduced. Additionally, the use of more intuitive, higher-level software can help reduce human error, thereby reducing the overall error rate of the tool. Since tools with lower error rates are "better" than tools with higher error rates, closed-source software GUIs can be seen as beneficial in this light.

Closed-source digital forensics software is maintained in a relatively centralized fashion, thereby reducing the susceptibility to change and increasing the perceived reliability of such systems. When fewer versions, varieties, branches, etc. of a system exist, it is easier to define the strengths, weakness, capabilities and limitations of the system in a clear, self-contained and succinct fashion. Closed-source

software licenses typically preclude the alteration of software by third parties, thereby establishing a clear and centralized versioning history. Open-source software, on the other hand, can usually be viewed, executed, modified and redistributed by any individual or group of individuals. Questions of changes to the reliability, weaknesses, strengths, capabilities and limitations of such software can be raised, and rightly so. Certain software systems outside the realm of digital forensics – such as the software used to control elevators, airplanes and pacemakers – generally require a thorough understanding of the organization involved their construction, and digital forensics software may one day be thought of in the same light. Another dimension to this observation is that, in a common-law legal system such as that widely used in the United States, case law – or precedents – can be used to refine or, in some cases, define the law. With fewer versions or varieties of closed-source tools being used, a greater percentage of positive court precedents are established for specific software tools and procedures, leading to better chances of their passing the Daubert test on grounds that they are widely used in the field.

Closed-source software is developed and sold by a group of individuals with expertise who practice the development of such tools as a profession. As the divide between digital forensics experts and computer programmers continues to grow, the reliance on closed-source tools should be expected to grow. [23] The development of tools used in digital forensics investigations and the use of such tools in ways which preserve the evidentiary value of potentially useful data are related, but separate, activities. In many ways, the field of digital forensics is still in its infancy and has not clearly separated itself from the field of computer programming or software development. This is less true of other forensics fields which, while clearly relying on the underlying scientific theories and related tools, have separated to the extent that (for the most part) their practitioners are not considered scientists or engineers, per se. As digital forensics and computer programming or software development continue to diverge, there will be less of an economic or social argument to use open-source software, since the

user-base will be largely unable to contribute to the development process. This process has already begun and it does not show any signs of abating.

8. CONCLUSIONS

In this paper, it has been our goal to provide an adequate context for understanding the distinction between closed-source and open-source software and its use in the digital forensics investigative process. To this end, we have provided a wealth of background material on both open-source software and digital forensics, given an account of some of the tools in use today and their feature sets, traced the history of legal thinking on evidence and related testimony by expert witnesses, explained in some detail the digital forensics investigative process and how software informs its various stages, and presented arguments for and against the use of open-source software as compared to closed-source software. The journey has certainly been a lengthy one, but hopefully the reader will have found it useful.

Figure C1 – Unfortunately, the question of whether or not open-source software is superior to closed-source software for digital forensics investigations remains an open one.

There is no clear answer to the question of whether open-source digital forensics tools are better or worse than closed-source alternatives, nor is there any clear indication of what the future holds in store (we cannot, for instance, make any statements as definitive as the one given in Figure C1).

Convincing arguments exist on both sides, and it is possible – these authors think it likely – that the two sides will continue to exist side-by-side as long as digital forensics is practiced professionally.

Understanding the tradeoffs involved in selecting one kind of tool over another, however, is as important now as it ever has been, and will continue to be so as long as choice continues to exist. The legal constraints on the use of tools, after all, are in place only to help ensure that, to the greatest

extent possible, the truth, the whole truth, and nothing but the truth is used in determining guilt or innocence – and selecting the right tool for the job is of paramount importance.

9. REFERENCES

- [1] Access Data. Forensic Toolkit (FTK) Computer Forensics Software. Retrieved online 4/16/2011 from <http://accessdata.com/products/forensic-investigation/ftkLang>. Syst. 15, 5 (Nov. 1993), 795-825.
- [2] Adams, C. W. 2008. Legal Issues Pertaining to the Development of Digital Forensic Tools. Third International Workshop on Systematic Approaches to Digital Forensic Engineering (May 2008), pp. 123-132.
- [3] Ami-Narh, J. T., Williams, P. A. H. 2008. Digital forensics and the legal system: A dilemma of our times. Australian Digital Forensics Conference (2008).
- [4] Carrier, B. 2002. Open source digital forensics tools: The legal argument. Retrieved online 4/13/2011 from http://www.digital-evidence.org/papers/opensrc_legal.pdf
- [5] Chi, H., Dix-Richardson, F., Evans, D. 2010. Designing a Computer Forensics Concentration for Cross-disciplinary Undergraduate Students . InfoSecCD '10 (October 2010), pp. 52-57.
- [6] Computer Forensics Certifications. List of computer forensics certifications. Retrieved online 4/16/2011 from <http://www.computerforensicscertification.net/computer-forensics-certifications.php>
- [7] Coolidge v. New Hampshire (1971), 403 U.S. 443, 467 (1971)
- [8] Daniel, L. Digital Forensics (powerpoint presentation). Retrieved online 4/15/2011 from <http://www.aoc.state.nc.us/www/ids/>
- [9] DataLifter. DataLifter. Retrieved online 4/16/2011 from <http://www.datalifter.com/>

- [10] Daubert v. Merrell Dow Pharmaceuticals, Inc. (1993) 509 U.S. 579, 589.
- [11] Department of Justice, Computer Crime and Intellectual Property Division. 2009. Searching & Seizing Computers and Obtaining Electronic Evidence in Criminal Investigations, 3rd Edition. Retrieved online 4/16/2011 from <http://www.cybercrime.gov/ssmanual/index.html>
- [12] Digital Intelligence. Digital Intelligence Forensic Software. Retrieved online 4/16/2011 from <http://www.digitalintelligence.com/software/disoftware.php>
- [13] Eoghan, C. 2004. Digital Evidence and Computer Crime. Academic Press (2004).
- [14] Federal Rules of Evidence. 1999. Fed. R. Evid. 7xx.
- [15] Frye v. United States (1923) 293 F. 1013 (DC Cir. 1923)
- [16] GNU. The GNU Public License (GPL). Retrieved online 4/16/2011 from www.gnu.org/licenses/gpl.html
- [17] Guidance Data. EnCase. Retrieved online 4/16/2011 from <http://accessdata.com/products/forensic-investigation/ftk>
- [18] Illinois v. Gates (1983), 462 U.S. 213, 238 (1983)
- [19] In re Commissioner of Public Safety (2007), 735 N.W.2d 706 (Minn. 2007)
- [20] Kenneally, E. (2002). Computer forensics - beyond the buzzword. Retrieved 4/16/2011, from <http://www.usenix.org/publications/login/2002-08/pdfs/kenneally.pdf>.
- [21] Kumho Tire Co. v. Carmichael (1999), 526 U.S. 137
- [22] LectLaw. Fourth Amendment Defined and Explained. Retrieved online 4/16/2011 from <http://www.lectlaw.com>

- [23] Letellier, F. 2008. Open Source Software: the Role of Nonprofits in Federating Business and Innovation Ecosystems. Association for Financial Markets in Europe (2008). Retrieved online 4/16/2011 from <http://flet.netcipia.net/xwiki/bin/download/Main/publications%2Dfr/GEM2008%2DFLetellier%2DSubmittedPaper.pdf>
- [24] Machado, C. Thompson, K. 2005. The Adoption of Open Sources within Higher Education in Europe and a Dissemination Case Study. Turkish Online Journal of Distance Education (2005), pp. 34-51.
- [25] Manson, D., Carlin, A., Ramos, S., Gyger, A., Kaufman, M., Treichelt, J. 2007. Is the Open Way a Better Way? Digital Forensics using Open Source Tools. Proceedings of the 40th Hawaii International Conference on System Sciences (2007).
- [26] Microsoft. Computer Online Forensic Evidence Extractor (COFEE). Retrieved online 4/16/2011 from <http://www.microsoft.com/industry/government/solutions/cofee/default.aspx>
- [27] Moody, S. J., Erbacher, R. F. 2008. Statistical Analysis for Data Type Identification. SADFE '08.
- [28] National Institute of Standards and Technology. Computer Forensics Tool Testing Program. Retrieved online 4/16/2011 from <http://www.cftt.nist.gov/>
- [29] Open Source Initiative. Mission. Retrieved online 4/16/2011 from <http://www.opensource.org/>
- [30] Paraben Forensic Software. P2 Commander. Retrieved online 4/16/2011 from <http://www.paraben.com/p2-commander.html>
- [31] Rosenbaum, D. E. 1987. Iran-Contra Hearings; North Says His Shredding Continued Despite Presence of Justice Department Aides: Shadow on Casey. The New York Times (July 8, 1987).

- [32] Ryan, D. J., Shpantzer, Gal. Legal Aspects of Digital Forensics. Retrieved online 4/16/2011 from <http://euro.ecom.cmu.edu/program/law/08-7332/Evidence/RyanShpantzer.pdf>
- [33] Sullivan, J. 2011. Free software is a matter of liberty, not price. Free Software Foundation (March 2011). Retrieved online 4/16/2011 from <http://www.fsf.org/about/>
- [34] U.S. Department of Justice, Office of Justice Programs, National Institute of Justice. 2004. Forensic Examination of Digital Evidence: A Guide for Law Enforcement (April 2004). Retrieved online 4/15/2011 from <http://www.ncjrs.gov/pdffiles1/nij/199408.pdf>
- [35] US Department of Justice, Office of Justice Programs, National Institute of Justice. 2008. Forensic Examination of Digital Evidence: A Guide for Law Enforcement (August 2008). Retrieved online 4/16/2011 from <http://www.ncjrs.gov/pdffiles1/nij/223433.pdf>
- [36] United States v. Brooks (2005), 427 F.3d 1246, 1251 (10th Cir. 2005)
- [37] United States v. Carey (1999), 172 F.3d 1268 (10th Cir. 1999)
- [38] United States v. Ganier (2006), 486 F.3d 920 (6th Cir. 2006)
- [39] Vlastos, E., Patel, A. 2008. An open source forensic tool to visualize digital evidence. Computer Standards and Interfaces Volume 30 Issue 1-2 (January 2008), pp 8-19.
- [40] Weiser, M., Biros, D., Mosier, G. 2006. Building a National Forensics Case Repository for Forensic Intelligence. Journal of Digital Forensics, Security, and Law (2006), 1 (2)
- [41] Wheeler, D. A. 2007. Why Open-Source Software / Free Software (OSS/FS, FLOSS, or FOSS). Retrieved online 4/16/2011 from http://www.dwheeler.com/oss_fs_why.html
- [42] Willassen, Y., Mjolswes, S. F. 2005. Digital Forensics Research. Teletronikk 2005:1, pp. 92-97

10. ABOUT THE AUTHORS

Jorielle Scott is a graduate of Selma High School, which is located in Selma, AL. Graduating with honors, she was ranked in the top 3% of her senior class. At Selma High School, she served as vice president of the National Honor Society, and president of Mu Alpha Theta(honorary math society), and a member of Who's Who Among High School Students. Currently, she attends Auburn University and will complete her course of study in May 2011. She is pursuing a Bachelor of Science degree in Computer Science with a concentration in Mathematics. As a sophomore at Auburn, she completed an internship at the Walt Disney World Resort. There, she served as a college program employer (CP) where she shadowed superiors in the development of software programs which were later released and executed throughout registers within the resort. As an intern, her duty was to test the newly develop software to ensure that the functionality of the software met its specifications one-hundred percent. Currently, Jorielle is employed at Cary Woods Elementary School which is located in Auburn, AL. At Cary Woods, she works under the America Reads Tutor program. She assists grade level students in both reading in math, helping to aid in the increase of students individual standardized test scores. In addition to working as a tutor, Jorielle also serves as the after school sports coordinator. As a sports coordinator, she teaches students basic fundamental skills in sports, as well as aid in their development athletically.

Patrick Carpenter received his Bachelor of Science in Computer Science degree from Auburn University in the Spring of 2010, graduating with honors. During the Summer of 2010, he worked in a research and development capacity at Los Alamos National Laboratory (LANL) in the area of High Performance Computing (HPC). Beginning in the Fall of 2010, he began a graduate program of study at Auburn University in the area of Computer Science and Software Engineering as a Master's degree student. He is currently a member of the Parallel Architecture and System Laboratory (PASL) and works under the guidance of Dr. Weikuan Yu, where he studies high performance computing, scientific computing and

parallel computing. Current research projects include GPU acceleration of a parallel ecosystem model using nVidia's CUDA toolkit and measuring the performance of the ecosystem model. Other research interests include computer architecture and operating systems and the mathematical and theoretical foundations of computer science. Patrick expects to graduate from his current program of study in the Spring of 2012.