Kandidatuppsats

IT-Forensik och informationssäkerhet 180 hp

Standards and methodologies for evaluating digital forensics tools

Developing and testing a new methodology

Digital forensik 15 hp

2018-06-29

Victor Andersson

HÖGSKOLAN
I HALMSTAD

Standards and methodologies for evaluating digital forensics tools

Developing and testing a new methodology

Kandidatuppsats

2018 Juni

Författare: Victor Andersson

Handledare: Stefan Axelsson

Examinator: Urban Bilstrup

Sektionen för informationsvetenskap, data- och elektroteknik

Högskolan i Halmstad

Box 823, 301 18 HALMSTAD

Kandidatuppsats
Sektionen för informationsvetenskap, data- och elektroteknik
Högskolan i Halmstad

Abstract

Standards play a big role in a lot of professions and when it comes to most aspects of law enforcement and forensic investigation, it’s no different. Despite that, for some reason, there aren’t any for when it comes to evaluating and choosing forensic tools.

The lack of an international standard for evaluating forensic tools has a clear negative impact on the digital forensics community as it lowers the value of tool tests and evaluations and hinders both the reproducibility and verification of their results.

Most tool evaluations are performed with custom forensic images and measures metrics that are not scientifically motivated, but rather made up based on the evaluator's personal preferences.

By examining current standards and related work done in the field, a new methodology is proposed. It builds on scientific principles and the strengths of existing literature. The methodology is then tested in a practical experiment.

The result of the paper is a solid foundation for a new standard to be built upon.

Contents

Tables

1 Introduction

Many of the tools and the software that forensics use have a long history and have evolved just as computers and digital systems have, and are just as useful for non-professionals and amateurs as they are for trained experts. Many of them weren't developed specifically for forensic purposes but were adopted over time by professionals due to their usefulness and utility. Examples of such utilities include tools present in various operating systems which were created with different tasks and problems in mind but proved suitable and useful for the work forensics do [21].

There is no internationally recognized standard that all forensics follow that dictate which programs and functions should be used and how instead it comes down to what organizations and often the individuals in those organizations prefer to use [1]. It has often been suggested that such a standard should be created and enforced around the world to make it easier for the digital forensics community to collaborate and reduce the amount of variance found in forensic investigations, but many logistical issues present themselves [2]. This thesis will discuss standards and techniques proposed by various organizations and researchers aimed at countering this problem.

The market for forensic tools is continually growing in size [3], and there is now a plethora of tools available for a computer forensic to use in their work. Because of that, choosing the right tool has become a more difficult and important task than ever before.

The breakneck pace at which technological advancements are made has led to a significant increase in the size and time requirement of each case. As cybercrime continues to become a progressively larger threat that affects more and more people [4], the importance of computer forensics to be able to process and analyze cases effectively is paramount.

As technological advances are made, and criminals find new exploits and strategies they can employ, the forensic software developers have their own arms race amongst each other to update their software and offer new solutions and features to combat the new rising threats and attacks [21].

Despite their similarities and the fact that they're built for the same purpose, no two tools are the same. Some tools are open source; some require a license from the creators. Some tools support a wider array of functions, while some others are designed to be better at fewer, more specific things. Some tools work on every operating system, whereas some others are limited to less.

Tests have been performed in the past to determine which of the many computer forensic tools on the market is the best, most worth it at its price point, most user friendly, has the most features and so on and so forth, and coming to any meaningful conclusion is difficult because the scope is so big and the number of things you have to account for quite overbearing [5][12][27].

Additionally, because the criteria in studies are oftentimes subjective and software updates, both big and small, so frequently change the programs in major ways, studies risk becoming outdated and misinforming, which makes an evolving standard for tool testing important for forensics to make sure they can keep up with the changes and work as efficiently as possible.

1.1 Purpose

The aim of this paper is to compare and assess existing standards and methodologies for testing and evaluating forensic tools and attempt to create a new one based on their strengths with the goal of helping forensics and organizations make the most informed decision they possibly can when choosing their tools.

After creating and testing the new methodology the goal is to have a solid foundation to further build upon in collaboration with the entire digital forensics community. Hopefully, this small contribution to the field can turn into something a lot bigger with far-reaching effects.

1.2 Limitations

Nowadays there are a lot of different forensic tools on the market to choose from. Their capabilities and functionalities vary from tool to tool, and performing tests on each one would be tremendously time-consuming, which is why limitations will have to be made.

This study will use two tools to test the new proposed methodology. The aim is to test a methodology, not compare every tool on the market, and to that end, two tools will suffice. Several aspects were taken into consideration when picking the tools, and ultimately EnCase and Autopsy were chosen for the experiment. Both tools have extensive features and can perform many different tasks. Because of that, both tools can be used in most situations, which makes a comparison between the two apt. Finally, EnCase is a proprietary, licensed product while Autopsy is free for anyone to download and use without restrictions, which adds an additional interesting aspect to the test.

The study will also be limited to the file recovery, also known as carving, capabilities of the tools. Because modern tools have so many features and capabilities, creating a methodology that covers each aspect of a tool would be a gargantuan task, far out of the possible scope of this paper. As such, focusing on one particular feature will create a more manageable task. Carving was chosen partly because it’s an important part of digital forensic investigations and partly because it’s interesting to work with.

2 Background

As the use of computers became more and more common during the 1970s and 1980s, so did crimes that involved the use of them. New types of crimes such as hacking and cracking developed, but older, traditional crimes such as fraud and child pornography started to shift into more of a digital crime as well. Because criminals knew the chances of being caught were small, and the chance of being prosecuted even if you were smaller still, crime via computers and the internet increased rapidly and still continues to affect more and more people today [4] [13].

As the spread and use of computers began the terms "digital forensics" and "computer forensics" were used interchangeably, however as technological advancements were made and more units such as the cellular phone became common everyday items they became distinct separate branches of forensic science. [14]

The American government agency US-CERT defines computer forensics as [15]:

> *We define computer forensics as the discipline that combines elements of law and computer science to collect and analyze data from computer systems, networks, wireless communications, and storage devices in a way that is admissible as evidence in a court of law*

The group Digital Forensic Research Workshop defines digital forensics as [16]:

> *The use of scientifically derived and proven methods toward the preservation, collection, validation, identification, analysis, interpretation, documentation and presentation of digital evidence derived from digital sources for the purpose of facilitating or furthering the reconstruction of events found to be criminal, or helping to anticipate unauthorized actions shown to be disruptive to planned operations*

As this new technology was developed and a new type of crime came with it, so did the need for new laws. Laws and their definitions were introduced separately in countries around the world. In the USA, the *Computer fraud and abuse act* of 1986 was the response to the growing threat of computer crimes and the potential difficulty in prosecuting those cases [17] [18]. In the United Kingdom it the *Computer Misuse Act* of 1990 was the response to the same issues [18].

Computer and digital forensics are rather unique when it comes to forensic sciences due to the fact that they, up until the 2000s, were developed both by law enforcement professionals and computer hobbyists and as such didn't follow the same scientific methods as the other forensic sciences [19]. That said, development wasn't slow. Despite the problems facing the field of digital forensics in today's landscape, digital forensics has seen a tremendous amount of evolution and development since its inception [21][23][25].

Once it became apparent that there was a need for standardization in the field, and guidelines for how investigations should be conducted, several organizations, including the International Organization for Standardization (ISO), proposed standards and practices that have made the digital forensics development more scientific in nature. As a result, the methods of

investigation have become more refined, and the credibility of digital forensics as evidence in court is unquestionable. The tools forensics use have also seen a marked increase in quality and quantity. However, one aspect of digital forensics that is lagging behind and isn't seeing much development is the creation of a standard and methodology for evaluating digital forensics tools [21][22][24]. As for the tools themselves, there are some guidelines such as the ISO 25010 standard which dictates which characteristics software should have. Several issues with forensic tool development have been discussed in research papers; some of those issues include the difficulty in keeping up with the changes in technology, high research-and-development costs, and short product lifetimes [21][23].

However, despite the tremendous amount of development and growth in the field of digital forensics, the lack of standards and its negative impact on the forensic community has been well documented by government agencies, professionals and scholars alike. The National Academies of Sciences, Engineering, and Medicine, which is the collective scientific national academy of the United States, found in a major 2009 study that the forensic science community in the U.S was severely underfunded, underdeveloped and undereducated as a whole, which is an opinion many academics share [1][2][14][21][22][23][24][26].

3 Question Statement

Due to the number of tools available on the market and the various tasks they can perform it may seem like an impossible job to figure out which ones to use and why. With the lack of international standards, and varying policies from company to company, forensics will most likely end up working with a multitude of tools during their careers, something that can be detrimental to their work as constant re-training costs both time and money, two valuable and finite resources.

The forensic community should attempt to adopt a standard protocol that outlines how forensics should evaluate forensic tools and how investigations should be conducted, in order to help facilitate easier, more efficient cooperation and collaborations between companies, law enforcement and government agencies around the globe.

In this thesis three specific questions will be answered, they are:

- What are the current methodologies and standards for evaluating forensic tools?
- Is it possible to improve on the state of the art?
- Can a useful result be produced in a practical experiment using the new proposed methodology?

3.1 Problematization

- What are the current methodologies and standards for evaluating forensic tools?

In order to create a new methodology, it is first necessary to learn about those that already exist and what they look like. Gaining a fundamental understanding of the field being researched is an essential part of any project, which is why this is the first, and likely most important, question posed in this thesis. In order to gain that understanding, thorough research of empirical material will be conducted.

Regardless of the answer, the question is sure to produce valuable results. If there are a lot of standards, there will be a lot of information to use and learn from. If there are only a few standards, it offers the opportunity to try and figure out why an area seemingly important is so lacking.

As with most research, the biggest issue is making sure all relevant information is gathered. Failure to find important sources and information will taint a paper and yield a lower quality result. If the information gathering process is not comprehensive enough this question cannot be answered to the degree that is sufficient and satisfactory.

- Is it possible to improve on the state of the art?

Finding a possible improvement, however large or small, would be an important step in the much larger process of creating a global standard. As such, the point of this question is to determine if the existing methodologies and standards can be improved upon, and if so, propose a new version. By first gathering all relevant information described in the opening question statement it can then be analyzed and hopefully turned into something new and improved. This new proposed methodology would be firmly based in scientific principles and

literature, something that is surprisingly rare in the field of digital forensics because of the organic and unstructured way in which the field came to be and has evolved [21].

The question is not without weak points, however. The most glaring one is the fact that it's possible that one or several methodologies or standards are discovered and determined to be unable to be improved upon, or the potential improvement too small to be of any greater value. While that would make creating a new standard a wasted effort, it still presents an opportunity to review existing work and discuss their strengths and weaknesses, which is a result in itself.

If there is a lack of scientifically rooted resources, there's a risk a proposed methodology ends up being too subjective or simply misses crucial information.

- Can a useful result be produced in a practical experiment using the new proposed methodology?

Once the research has been done, and the methodology has been produced, the final logical step is to perform an evaluation, which in this case is most appropriately done by conducting a practical experiment. The goal is to produce a result that is reliable, consistent, meaningful and, most importantly, reproducible. It's also important for the experiment and methodology to function despite the scale. A good methodology works both when testing every function of a tool or just a select few. One potential problem with this question is the use of the term "useful result" as it can lead to misinterpretation. A "useful" result could mean different things to different people. In this practical experiment only the carving capabilities of two tools will be tested, using specific metrics to compare them, and as such determining if the result is "useful" or not is simple, but the methodology must be well defined in order to avoid confusion, especially when working with several tools on a large scale.

4 Method

The methods for achieving answers to the questions posed in this work are straightforward. Thorough research of empirical material and a practical experiment cover all bases and meet the demands set. Once the gathering of information is complete it will be analysed and turned into something new. This new methodology will be based on lessons learned and knowledge gained during the course of the project. As such, no strict guidelines or rules will be followed.

Highly valuing peer-reviewed material based on scientific principles and methods, for this paper the relevant information is most likely to be found in various academic journals, conference papers or scholarly journals. The potential policies of different government agencies will also be included in the study. Journalistic articles and other work not published in any journal or presented at any conference, however, will not be examined as they do not meet these criteria.

Performing an interview study would be a viable alternative to the chosen method, however, due to the reliance on others to achieve a satisfactory result, it was not done.

4.1 Problematization

The emphasis on peer review can be debated, and there isn't necessarily a right or wrong answer. Different methods are better suited for different situations. The reasoning behind the decision to focus on peer-reviewed material in this work is based on several factors. As standards should be of the highest utmost quality, a collaborative effort from the community they affect is required. That means a lot of input from a lot of people. Personal blogs or secluded online communities is not the correct type of forum for that kind of discussion. Furthermore, when the material is peer-reviewed, flaws and weaknesses are more easily identified and rectified. The way digital forensics developed is clear proof that peer review isn't always necessary or important, but for something as serious and wide reaching as a standard, it is worth serious consideration.

4.2 Related Work

4.2.1 NIST

The biggest effort in this particular field is that of the National Institute of Standards and Technology (NIST) in the United States.

In 2004, NIST, in collaboration with various other American agencies, released their methodology for how to evaluate and test forensic programs, specifically focusing on data acquisition tools and write blockers [6]. With a focus on accuracy and completeness, the methodology offers detailed scientific instructions for how to perform tests and evaluate the results.

NIST has since increased its resources and now includes several other proposed methods for various types of tasks such as forensic file carving and string search [7][8]. NIST develops these methods in categories based on the functions of the program.

NIST's process for developing standards is very methodical and includes 14 steps divided into two separate phases. They are as follows:

1. Specification development process

After a tool category and at least one tool are selected by the steering committee the development process is as follows:

1. NIST and law enforcement staff develop requirements, assertions and test cases document (called the tool category specification).
2. The tool category specification is posted to the web for peer review by members of the computer forensics community and for public comment by other interested parties.
3. Relevant comments and feedback are incorporated into the specification.
4. A test environment is designed for the tool category.

2. Tool test process

After a category specification has been developed and a tool selected, the test process is as follows:

1. NIST acquires the tool to be tested.
2. NIST reviews the tool documentation.
3. NIST selects relevant test cases depending on features supported by the tool.
4. NIST develops test strategy.
5. NIST executes tests.
6. NIST produces test report.
7. Steering Committee reviews test report.
8. Vendor reviews test report.
9. NIST posts support software to the web.
10. DHS posts test report to the web.

The major advantage of NIST's method is how thorough it is and the detail in which it goes into. The method is scientific and meticulous with several reviews by both parties involved (the vendors and the organization performing the tests) along the way which leads to results likely to be trustworthy and just.

However, there are disadvantages to NIST's process, most importantly time and coverage.

With constant reviews of the methodology being performed by vendors and the organization performing the test, development can take several years before it’s completed. By the time a new methodology has been developed, it risks already being outdated in certain areas due to the rapid development of the forensic tools it’s meant to evaluate. Additionally, NIST develops methodologies on a per-function basis, meaning it is not necessarily suited for

testing the overall worth of a tool, just parts of it. The methodology for testing and evaluating programs focusing on data acquisition and write blockers was published in 2005 [6] while file carving came four years later in 2009 [7]. An example of how this could leave NIST unprepared is found in cloud computing. Cloud computing is a big and rapidly growing technology which isn't represented in any of NIST's literature. As cloud computing has exploded onto the scene, the slow development cycle of NISTs process has failed to catch up and produce any standard.

With only a small part of the tools having a methodology for testing, NIST's procedure is certainly incomplete but is still of great value.

An example of a NIST standard that's appropriate for this thesis is the one for file carving. It contains only four bullet points, but they are, as expected, well thought out and reasoned. The specifications detailed in the NIST standard for forensic file carving tools [3] are:

• The tools are used in a forensically sound environment.

• The individuals using these tools adhere to forensic principles and have control over the environment in which the tools are used.

• The carving tool input is a file or set of files that might be produced by a forensic acquisition tool acquiring digital media such as secondary storage or volatile memory.

• The files used to test input to carving tools were created in a process that places file data blocks in a manner similar to how end-user activity would locate file data blocks.

NIST does not recommend any specific forensic image to perform tests with, something that can be considered a flaw. If testers don't work with the same data sets, especially ones that have been specifically generated for forensic tool testing purposes, results will be impossible to compare in any meaningful way.

On top of government agencies attempting to create standards, academics have also performed studies and experiments to produce appropriate literature.

4.2.2 Batten and Pan

In a 2007 paper Batten and Pan attempt to create a methodology for correctness testing for forensic tools [9]. They use their own methodology in which they decide to test the file recovery capabilities of four forensic tools using the two metrics they deem the most valuable and reliable, accuracy rate and precision rate. They define accuracy rate as "the number of correctly recovered files out of the total number of original files to be recovered" and precision rate as "the total number of recovered files out of the total number of original files" [9]. To test the programs, they then performed file carving in all four programs on two independently generated forensic images and compared the results. The forensic images were gathered from Digital Forensics Tool Testing Images [10] and Digital Forensic Research Workshop [11]. The results of the study seemed to indicate it was a success and that the method could be used in similar tests in the future.

The use of publicly available pre-made forensic images is especially commendable in this work, as it makes reproducing and verifying results easier. The methodology has deep roots in scientific principles, and the authors are clear in regard to the strengths and weaknesses of the results they managed to produce.

One problem with the study is the choice of metrics. If the study were to be expanded to encompass other functions of the tools, accuracy and precision rate would be unsuitable for producing results for many of them.

Another potential problem with the study is the small sample size. Testing two functions on four tools twice on two separate forensic images may not be enough to present an accurate picture on a bigger scale. With the number of tools on the market and the complexity of them and their inner workings, it would be appropriate if the study considered more factors and tested them in more tools using a larger number of forensic images. Granted, this might put the scope out of their reach and be considered unrealistic, but if a methodology were to be standardized and globally adopted it should be thoroughly tested.

4.2.3 Creutzburg and Kröger

Creutzburg and Kröger performed an experiment in their 2013 paper "A practical overview and comparison of certain commercial forensic software tools for processing large-scale digital investigations," in which they compared four different forensic tools using several functions and four unique forensic images [12]. Their goal was to test the performance, functionality, usability, and capability of the tools. They created their own forensic images to work with. Those forensic images were of varying sizes and contained several different kinds of file types, including .jpg and .mp3. They also provided a good overview of how unforeseen issues could arise with their experiment.

Unlike Batten and Pan and NIST, they didn't set out to create a standard or methodology, but their paper is useful regardless. Similar to Batten and Pans paper, they created a test with variables they chose themselves and deemed most important to answer the questions they posed and were able to reach a legitimate conclusion. They thoroughly explain their testing process, which paints a clear picture of how the tests were performed and what the results were.

Although this study was useful in some sense, it does have several deep flaws.

The creation and use of their own forensic images that they did not make publicly available severely diminish the ability for others to recreate the study and produce results that can verify their findings.

Some of the criteria for judging the programs were subjective, such as how intuitive the graphical user interface (GUI) was, and they didn't calibrate the programs they used resulting in programs with more functionalities receiving a lower score due to their increased processing time over simpler programs.

This kind of study also becomes outdated quickly as the tools develop and change significantly over time, even from the time of writing to the time of publication.

4.2.4 Leahy Center Students

Students at the Leahy Center for Digital Investigation at the Champlain College in the U.S performed a study in 2016 in which they investigated the differences and capabilities of three different forensic tools using methods based on the research of professors there [20]. They set out to answer four research questions.

In the first question, they compared the speed at which the different tools performed specific keyword searches.

In the second question, they looked at how many "hits" each tool got for each search and what potential differences could mean.

In the third question, they investigated the timeline features of each tool.

In the fourth question, the exporting features of each tool were measured.

The questions are well thought out and easy to produce results for, and the reasoning for them well explained in the paper.

They opted to use their own self-made data sets which consisted of downloading programs, files, and pictures, installing various programs, performing searches on different websites using different web browsers, and copying and deleting several files in the computer.

After the data sets were created, they analyzed them using EnCase, Forensic ToolKit (FTK) and Magnet AQUIRE. Performing Imaging and keyword searches, comparing timeline features and exporting options, they were able to answer the questions they posed and found useful data that proved the differences between the tools and their meaning. The tools work in different ways, and each offers some functionalities that the others lack.

However, though thorough and thought out throughout, there are drawbacks and problems with their study. Specifically, it's structure and usefulness in the long run.

Much like Creutzburg and Kröger's study, the results will lose its relevance quickly. As a matter of fact, during the compiling of their results, one of the tools released a software update that affected several functions. In order for these kinds of studies to keep their usefulness, they would have to be updated and re-done every time a major software update is released, which is a rather gargantuan task that requires a lot of resources, both time-and-money-wise.

As for the structure of the experiment, creating their own data sets rather than use pre-existing, known forensic data sets as Batten and Pan did is a weakness. The difference between this study and the one performed by Creutzburg and Kröger is that the students made the creation process of the data sets available, technically enabling any reader to recreate them to use themselves. This, however, would be a wasted effort, as some instructions are vague or no longer relevant due to the evolution of the software they used in the creation of the data sets.

4.2.5 Dykstra and Sherman

In a 2012 paper, Dykstra and Sherman sought out to evaluate eight different forensic tools ability to remotely acquire forensic evidence from cloud computing and measure their effectiveness [26]. At the time, tests involving cloud computing were rare, and because of that, they created an experiment highly customized and informal. They performed three experiments in total and measured the results using two metrics. They measured if the tool was able to collect evidence remotely and how accurately the data compared to a standalone control machine.

They were able to produce clear results and could prove differences between tools. However, the experiment had several serious flaws.

For starters they weren't transparent about which tools they used during the tests, nor did they expressly state the differences between their capabilities. They also didn't mention the specifications of their web server or the websites they created.

The experiment would not be possible to recreate due to vagueness regarding many technical aspects of the building of their testing environment. It's also a study that quickly becomes outdated as the equipment and tools they work with develop over time.

The reason why this study is noteworthy is that working with the cloud introduces a third party to the mix. Testing and verification become more difficult because different service providers offer different tools to work with. Technological challenges that don't exist during traditional testing suddenly appear and need to be solved somehow. On top of that, there's now an ethical aspect to these tests. Dykstra and Sherman correctly point out that you will have to rely on trust with the provider to ensure data is correct and unaltered compared to what's in the cloud and that disturbance of the cloud infrastructure might be necessary. As cloud computing becomes a bigger part of the internet, and how we as a society use it, new research will have to be made and answers to new questions need to be figured out.

4.2.6 Brorsson and Wernebjer

In a 2010 paper, Brorsson and Wernebjer attempted to evaluate two forensic tools, EnCase and BackTrack, to determine if a licensed and an open-source product would yield equally valuable results [5]. They did so by choosing the three metrics they deemed most important, which were keyword searching, file signature analysis, and hash verification. They then performed a practical experiment in each program.

On top of the practical experiment, they also conducted two interviews with experienced professionals in the field of digital forensics to learn what tools they use and what their thought process is when determining which ones to use.

The study is useful because they clearly documented the testing process and were able to produce viable results using their own made up technique. On top of that, having more papers to analyze helps increase the quality of this work. However, the usefulness ends there.

Despite the study sounding good in theory, several major flaws reduce its value. The authors use their own forensic image that they share no information about except for its size which is only described as “small.” They follow no methodology or apparent structure in their testing, and little of what they do seems to be based on any scientific grounds. Additionally, because this paper only compares two tools in their current state, it runs into the same issues as several other previous works have, wherein it gets outdated quickly as tools receive updates and technological advances are made.

4.2.7 Kiper

Kiper, in a 2018 paper for the SANS Institute, sets out to create a forensic tool topology to learn what members of the forensic community look for in a tool when they decide which one to use [27]. The study is based on what tools are able to do in theory, not how they behave when performing those functions.

In order to produce results that clearly indicate what professionals in the field of digital forensics deem most essential in a tool, he performs both in-person interviews and posts an online questionnaire. The interviews were conducted with 13 people while the online questionnaire was answered by 46 individuals. The approach is scientific and based on previous research done in the field. The results from the interviews and questionnaire were analyzed and proved that the characteristics valued the highest were similar in both cases.

The aim of the paper, once a topology is created, is to help facilitate the choice of tools for forensics worldwide. If the person trying to decide what tool to use understands their needs and can find the tool that seems most appropriate based on those criteria, they can narrow down the list of tools they are interested in testing which makes the actual testing phase faster and easier.

Criticisms towards the paper include the fact that the sample size of people is small and answers are inherently subjective.

While no methodology for performing the actual testing of the tools is discussed, it is still a useful work for this paper.

4.2.8 Common themes

Only NIST's, as well as Batten and Pan's papers, actually proposed a methodology for forensic tool testing, the others simply performed tests without attempting to help create a standard.

The methodologies used in the papers vary in many ways but are similar in some others. With the exception of Batten and Pan’s paper, the metrics tested and determined to be the most important are based on personal opinion, not scientific literature.

Efficiency was a metric highly valued in every paper, and it comes as no surprise as forensic images are continuingly increasing in size and the time required to process them getting longer.

Very few authors seemed to care about the future of their work and how it would age. Almost all tests were only concerned with how the tools performed in their current state and iteration, with no regard for how to keep results relevant in the future. This is also evident in the fact that most studies use their own data sets and forensic images which makes recreating the results near impossible for others.

4.2.9 Question statement positioning

The question statements posed in this paper vary quite greatly from those in the related works. Most authors seek to compare forensic tools and do so using whatever metrics they deem the most important, rather than create a methodology or standard that could be used by the entire digital forensics community. Batten and Pans paper is the most similar and do attempt to find answers to questions that aims to solve the same problem , but even then, their focus is more on the technical aspect of a methodology and is viewed on a smaller scale. The differences between this paper and the related work thusly is quite clear.

4.3 Proposed Methodology

The methodology presented in this work draws inspiration from all previous works examined as well as from the standards proposed by NIST. Because of the limited scope of this thesis, the focus will be on carving specifically, but it applies to any aspect of forensic tool testing with some modifications.

The methodology proposed is as follows:

- The tools are used in a forensically sound environment.
- The individuals using these tools adhere to forensic principles and have control over the environment in which the tools are used.
- The individuals performing the tests do so multiple times on multiple systems.
- Pre-made, legitimate forensic images from trusted sources should be used to perform tests with.
- Limit the scope of the essential metrics. For carving those are accuracy rate, precision rate and performance.

The methodology is based on both scientific research as well as the documented opinion of those in the digital forensic community. Building upon the lessons learned from evaluating previous work, a strong foundation is created. Utilizing this methodology, evaluations both big and small should produce useful results.

As with everything else, however, this methodology does have weaknesses.

Different metrics will matter when you test different aspects of a tool, and this methodology isn't expanded enough to the point where those are specified. Accuracy rate, precision rate and performance are the metrics chosen for carving purposes specifically, but they are not terms that are all-encompassing and will not be ideal for all aspects of a tool.

The methodology assumes that forensic images either already exist or are created that closely resemble real cases and can be used regardless of what function of a tool is being evaluated. Since real case files are unlikely to be used for testing, new, advanced images would have to be used, which causes an interesting potential problem. If forensic images are designed with the purpose of being used for tool evaluation in mind, it's possible they become too perfect and convenient to use. If the images used are too unrealistic or make up for tools weaknesses, the evaluation and grading of tools could be swayed and become less valuable.

Furthermore, the methodology hasn't been peer-reviewed, and no input from the forensic community has been gathered.

5 Experiment

Based on the proposed methodology and previous works, the practical experiment in this paper will be performed as follows:

In accordance with the proposed methodology, testing will take place in a secure, forensically sound environment on two separate computers.

The computers are identical in their specifications. The computers used are Dell OptiPlex 5040 with an Intel i5-6500 CPU @ 3.20GHz and 8 GB RAM running 64-bit Windows 10 Enterprise.

Because of the extensive number of features available in each forensic tool, performing a complete test of each one would be unrealistic as the time commitment would be overwhelming. As such, this experiment is limited to the file carving capabilities as well as the general performance of each tool. For the purpose of this test, performance is measured in the amount of time it takes to process the images and produce the results.

EnCase v. 8.06 and Autopsy v. 4.6 will be used to analyze four separate pre-made forensic data sets stored as disk images on an external USB-drive.

The first two data sets, Basic Data Carving Test #1 and Basic Data Carving Test #2, were gathered from Digital Forensics Tool Testing (DFTT) [10]. The other two data sets, L0_Graphic.dd and L0_Documents.dd were gathered from the Computer Forensic Reference Data Sets (CFReDS), which is a part of NIST [25].

Basic Data Carving Test #1 is 62 MB in size and contains 15 files. The image contains several allocated and deleted files and the header one JPEG file was modified. The different file types are .doc, .wav, .jpg, .xls, .pdf, .jpg, .gif, .mov, .wmv, .ppt and .zip.

Basic Data Carving Test #2 is 124 MB in size and contains ten files. This file system image contains several allocated and deleted files, none of which have been modified. The different file types are .bmp, .doc, .jpg, .pdf, .gif, .xls and .ppt.

L0_Graphic is 64 MB in size and contains 6 files. The image contains non-fragmented graphical files. They are .jpg, .png, .bmp, .gif, .tif and .pcx.

L0_Documents is 42 MB in size and contains 7 files. The image contains non-fragmented document files. They are .doc, .xls, .ppt and .pdf.

Each image will be hashed and compared with the legitimate, known hashes publicly available to ensure their integrity before, during, and after the tests.

The results will then be measured using the metrics “accuracy rate” and “precision rate” as well as overall performance.

5.1 Hypothesis

The hypothesis is that every file will be recovered, and processing time will take under a minute.

The forensic images used in the experiment are small, old and have well-documented contents. Those three factors should make it easy for the tools to successfully carve the files without many, if any, errors.

5.2 Problematization

The phrase “forensically sound environment” sounds vague, and it is. The phrase lacks a real definition but in this experiment it is going to be defined as: “An environment free from contamination and under control by the tester”. Essentially, an environment that the tester controls access to and that they can prevent contamination in.

The decision to put the forensic images on an external USB hard drive does introduce an additional potential source of failure and contamination, but it was made in order to try to keep the test as similar to a real case as possible.

The choice of hardware is largely determined by circumstance. The computers were chosen primarily due to availability, as they were the only computers in the secure lab environment at the university this study was performed in. Computers found in forensic labs are assumed to vary greatly in specifications, which makes it tough to judge how these compare on average.

The images were chosen primarily for three reasons. First, they’re used by Batten and Pan in their paper. By using the same images it’s possible to compare the results produced in this experiment with the results that they got. Secondly, the images are easily available to download. They’re free and come ready to use immediately after download. Finally, they are described in great detail. The contents of each image is both known and hashed which makes verification very easy to perform. The small sizes (the images are all under 1GB each) of the images is a weakness. It is completely unrealistic to work with forensic images that only have a handful of files on them in the real world. However, for the sake of this experiment they are deemed sufficient

The biggest weakness of the experiment is the small scale of it. Testing one function of two tools on two computers using four smaller forensic images, while sufficient to produce results, is simply put not sufficient for a global standard. For a serious evaluation and comparison between tools a lot of functions would have to be tested, something the methodology in its current form does not support, and something that is much too resource-demanding than this thesis allows for. The results produced from the experiment will not be useless, but they are diminished in value because of these limitations.

6 Results

The hypothesis that every file would be carved turned out to be almost correct.

Overall the tools performed very similarly. In 33 out of the 34 total tests performed every file was recovered correctly. The one instance where a file wasn't correctly carved was in the first test of Autopsy on the first computer. The file failed to be carved was domopers.wmv, which was a deleted file.

The images and files were hashed and compared to the original files hashes and were all matches.

A more detailed description of the result of the practical experiment is detailed in four tables on the page below.

The tables have seven columns with different variables. They are:

Test. Numbers each test.
Image. The forensic image used in the test.
Recoverable Files. The number of known files contained in the forensic image.
Recovered Files. The number of files successfully recovered.
Precision Rate. The total number of recovered files out of the total number of original files.
Accuracy rate. The number of correctly recovered files out of the total number of original files to be recovered.
Performance. The time required to complete the carving process.

Additionally, the forensic images used are labeled A, B, C, and D.
A is Basic Data Carving Test #1, the image that is 62 MB in size and contains 15 files.
B is Basic Data Carving Test #2, the image that is 124 MB in size and contains ten files.
C is L0_Graphi, the image that is 64 MB in size and contains six files.
D is L0_Documents, the image that is 42 MB in size and contains seven files.

Test	Image	Recoverable Files	Recovered Files	Precision Rate	Accuracy Rate	Performance
1	A	15	14	93%	93%	< 30 seconds
2	A	15	15	100%	100%	< 30 seconds
3	B	10	10	100%	100%	< 30 seconds
4	B	10	10	100%	100%	< 30 seconds
5	C	6	6	100%	100%	< 30 seconds
6	C	6	6	100%	100%	< 30 seconds
7	D	7	7	100%	100%	< 30 seconds
8	D	7	7	100%	100%	< 30 seconds

Table 1. The result of Autopsy test on PC #1

Test	Image	Recoverable Files	Recovered Files	Precision Rate	Accuracy Rate	Performance
1	A	15	15	100%	100%	< 30 seconds
2	A	15	15	100%	100%	< 30 seconds
3	B	10	10	100%	100%	< 30 seconds
4	B	10	10	100%	100%	< 30 seconds
5	C	6	6	100%	100%	< 30 seconds
6	C	6	6	100%	100%	< 30 seconds
7	D	7	7	100%	100%	< 30 seconds
8	D	7	7	100%	100%	< 30 seconds

Table 2. The result of Autopsy test on PC #2

Test	Image	Recoverable Files	Recovered Files	Precision Rate	Accuracy Rate	Performance
1	A	15	15	100%	100%	< 30 seconds
2	A	15	15	100%	100%	< 30 seconds
3	B	10	10	100%	100%	< 30 seconds
4	B	10	10	100%	100%	< 30 seconds
5	C	6	6	100%	100%	< 30 seconds
6	C	6	6	100%	100%	< 30 seconds
7	D	7	7	100%	100%	< 30 seconds
8	D	7	7	100%	100%	< 30 seconds

Table 3. The result of EnCase test on PC #1

Test	Image	Recoverable Files	Recovered Files	Precision Rate	Accuracy Rate	Performance
1	A	15	15	100%	100%	< 30 seconds
2	A	15	15	100%	100%	< 30 seconds
3	B	10	10	100%	100%	< 30 seconds
4	B	10	10	100%	100%	< 30 seconds
5	C	6	6	100%	100%	< 30 seconds
6	C	6	6	100%	100%	< 30 seconds
7	D	7	7	100%	100%	< 30 seconds
8	D	7	7	100%	100%	< 30 seconds

Table 4. The result of EnCase test on PC #2

7 Discussion

The field of forensic tool testing standards and methodologies is largely unexplored and underdeveloped. Aside from work done by NIST, there isn't any fully fledged out body of work that deals with the topic. Whether that is because it is a topic deemed academically unimportant and insignificant or because it's a hidden gem waiting to be discovered and explored is unclear from the literature.

The impression gained from reading related work and standards is that the people in charge of performing the tests and evaluations prefer to arrange their experiments themselves based on their subjective opinions of what is best and most relevant for them in each case. One explanation for why that is might be that for a professional evaluating a tool, it's apparent what it should do and how it should behave. If the tester is interested in the disk imaging feature of a tool, it is obvious that one of, if not the most important aspect is how accurately the tool creates the forensic disk image. How much of the original content is on the image, did any data change and how long did the tool take to complete the process? The metrics come naturally. The same seems true for the topic explored specifically in this work, carving. Batten and Pan were the only ones to propose metrics to base the evaluation of the tools on and motivate them scientifically, but every other paper reached the same conclusion anyway. When you carve files, you want to see how many out of the potential files the tool was able to carve. As long as the thought process behind the choice of metrics is sound and well-reasoned an argument can be made that standards aren't needed, or at least not needed to govern which metrics should be tested. That said, if it really is that obvious and that people come to the same conclusions independently, then those metrics may as well be incorporated into a standard. A standard would ensure everyone is on the same page and helps with reproducibility and verification, which is why it is still included in the proposed methodology in this work.

7.1 Proposed methodology

The methodology proposed in this work contains five points, the first of which is:

The tools are used in a forensically sound environment.

This step is copied from NIST's file carving standard, and for obvious reasons. It is imperative to perform tests in an environment that minimizes the risk of contamination and outside interference. Following standard forensic procedures and practices will yield better results and produce results that are more trustworthy. What a forensically sound environment is not is not explained in greater detail as it is assumed that those reading this work and using this methodology is already familiar with the concept.

The second point is:

The individuals using these tools adhere to forensic principles and have control over the environment in which the tools are used.

The second step is also copied from NIST's file carving standard and expands upon that idea introduced in the first step and again assumes some sort of previous knowledge or

background in forensic work. For tests to be considered accurate and trustworthy, it's important that no tampering, accidental or otherwise, is possible and no outside interference could jeopardize the integrity of the test. Working in an environment that the tester has full control over is therefore of paramount importance.

The third point is:

The individuals performing the tests do so multiple times on different systems.

This is where the standard starts diverging from NIST. Performing the test several times across multiple systems serves several functions. It reduces the risk that errors or outliers skewer the results and allows the tester to cross-reference the data they produce and get a better understanding for what each tool is capable of and how they handle on different systems with different performance capabilities. The more data the testers have to work with the better. However, it obviously comes with the drawback that testing several times across multiple systems is a much lengthier process.

The fourth point is:

Pre-made, legitimate forensic images from trusted sources should be used to perform tests with.

Creating your own data sets or images does have benefits over using a pre-made image. It allows for much greater customization and as a result, gives the person many more options to explore. Features and filetypes that are missing from pre-made images simply would not be evaluated going by the proposed methodology, whereas with a custom image they creator would have the ability to add the things most relevant and interesting for them. An additional benefit is that the person would, likely, have possession of copies of the files and can thus more easily verify the results of the test.

However, there are two major drawbacks to creating your own images. The biggest one is the time requirement. It is significantly quicker and easier to download an image with known content that's been verified than it is to create something new. It also eliminates the risk that the person creating the image make a mistake and negatively impact the test.

Perhaps most importantly, using a pre-made image helps with reproducibility and verification of the results, something integral to the scientific method. Having everyone working with the same data means any result can be easily examined. If everyone uses their own image, a perfect reproduction would have to be made, or the image would have to be shared by the tester.

The biggest problem with pre-made images is that their content has to be agreed upon and they have to be maintained and updated by someone. An image would need to be as similar as possible to real cases without being direct copies of real images. A risk exists that images would be intentionally created to perform exceptionally in tool tests and mask the weaknesses in the tools used.

With both options weighed, pre-made images are still recommended because it eliminates potential human errors, save time and help facilitate better reproducibility and verification.

The fifth and final point is:

Limit the scope of the essential metrics. For carving those are accuracy rate, precision rate and performance.

This step is one that would change depending on what aspect of a tool is being tested and needs to be decided on by the forensic community, not a single paper. In the case of carving these metrics are pretty straightforward. The metrics for carving have already been discussed. However, performance is a general metric that applies to any tool test. The performance metric is useful for the obvious reasons if two tools can perform the same task, but one of them is significantly faster it is clearly superior. As hard drives and test cases continue to grow in sizes at a rapid pace, time efficiency becomes a more and more valued metric.

The reason for recommending a limited scope is to avoid bloating and getting bogged down in minuscule details. Forensic tools are constantly becoming more advanced and capable of more things, which makes performing full tests complicated and time-consuming. Determining the most important metrics to test that covers every aspect of the tool is therefore proposed to help with efficiency.

Determining all those metrics is a big task that shouldn't be left to any one person or group. A collaborative effort amongst the forensic community would be required and more research on the matter performed.

One of the primary aims of this thesis was to discover if it was possible to produce a new methodology for evaluating forensic tools, and the answer is obviously yes. Despite the limited resources to gather information and inspiration from, there is enough previous work and a solid enough foundation of knowledge to build upon. As for whether or not the proposed methodology is a significant improvement over the existing literature, that is for the community to decide. The thought process and reasoning behind every decision have been laid out, but it will be up to others to determine if it is of value.

7.2 Experiment

The result of the experiment was as expected and predicted in the hypothesis. Due to the fact that the forensic images used are small, old and well known it comes as no surprise that there was only one file that didn't get carved during the experiment. When comparing the result to that of Batten and Pan it's clear that the tools have become better since their paper was published as their accuracy and precision rates were lower.

The actual results of the experiment itself, while interesting and certainly not without use, are not the real results of interest in this experiment. The real result is not something that can be explicitly grasped; it's not learning which tool outperformed the other, but rather if the methodology was useful and effective for the testing process. It's a difficult thing to put one's finger on but overall this study should be considered a success. The methodology clearly establishes a safe environment free of tampering and interference, which means any result, any data gathered, can be considered legitimate and trustworthy.

Performing the tests multiple times on different systems also proved to be a worthy inclusion in the methodology. While it did increase the amount of time the test took, having more data to reference and compare gave a clearer picture of the results and made the process feel like it had more weight to it. Restraint does need to be exercised, however, as excessive testing quickly diminishes the return on time investment. Perhaps it would be wise to include a limit on the number of systems to perform the tests on, but ultimately the evaluator should be the one responsible for this decision.

Using pre-made forensic images was both a tremendous time saver and helped with the verification and recreation of the results, as expected. Finding and downloading a forensic image complete with descriptions and hashes of its content was both easy and makes it possible for others to follow along and recreate the experiment. The ability to customize the experiment in great detail was lost, and that was a clear negative factor. However, in the opinion of the author, the benefits do outweigh the costs.

Limiting the scope to only measure accuracy rate, precision rate and performance was sufficient to produce high enough quality results. It gave a very clear picture of each tools capabilities and the success rate of the experiment. In order to differentiate between tools in greater detail, more variables would have to be measured, but it didn't feel necessary. For those very detail oriented it is, of course, viable to do so, but it is not required by any means.

Overall, the experiment, while small in scale and not very technically advanced, gave a strong impression that the methodology is sound and fit for use for the specific task it was designed for. However, if a tester wanted to evaluate the tools based on other capabilities the measurements specified in the standard might not be so well suited.

8 Conclusion

At the beginning of this thesis, three questions were posed to be answered.

The first question was:

What are the current methodologies and standards for evaluating forensic tools?

The short answer is that there aren't many. The field of tool evaluation is a largely unexplored one, at least academically speaking. There exists very little research, and the proposed standards and methodologies are few. Batten and Pans paper, "An Effective and Efficient Testing Methodology for Correctness Testing for File Recovery Tools," makes one of the biggest contributions to the field but despite that, it is small in both scale and scope. Perhaps most important is the NIST proposed standards. The standards are thorough and well developed, informed and shaped by both professionals and vendors, and go into great detail. However, NIST's standards don't cover every function of the tools and have a long, slow development cycle.

As such, the conclusion is that there simply aren't many, and the few that exist are not developed enough to the point where they can be used universally in any test.

The second question was:

Is it possible to improve on the state of the art?

Simply put, yes. As discussed, there are flaws in existing methodologies that can be addressed. An example discussed in this thesis is the use of standardized premade forensic images in evaluations to help facilitate reproducibility and verification. Taking a stronger scientific approach to methodologies is emphasized in this thesis and is what sets it apart from previous work.

The third question was:

Can a useful result be produced in a practical experiment using the new proposed methodology?

Yes, the methodology is designed in such a way that the results are legitimate and can be trusted. However, the usefulness depends on the experiment and the arrangement of it. The bigger the scope, the better the results. Larger forensic images with more data and variation are going to produce better results than a small image that only contains a few items. The methodology also needs to be expanded upon to support more functions than just carving.

In conclusion, the question has been answered to a satisfactory degree. However, there is much more work to be done in order to produce a full standard. This thesis is a stepping stone to launch of off, not the last word on the subject.

9 Future Work

This thesis offers a methodology that's a solid foundation to build upon further. It is based on scientific principles and incorporates several aspects of important work done in the field. It should be easy to follow and produce results that can be easily recreated and verified.

It's going to be up to the forensic community to build on the foundation and help complete the methodology. The area most lacking and in the greatest need of input is the fifth step. Right now, the methodology only supports the testing of a tools carving capabilities. Coming to an agreement on which metrics are most valuable when evaluating a specific aspect of a tool is of paramount importance in order to create a solid methodology, and thus it's there the focus of future work should lie. This will require cooperation between researchers, institutes and government agencies worldwide, as the creation of a standard is a lengthy and intricate process that demands major community involvement.

Reference List

[1] Leigland, R. and Krings, A.W., 2004. A formalization of digital forensics. *International Journal of Digital Evidence, 3*(2), pp.1-32

[2] Slay J., Lin YC., Turnbull B., Beckett J., Lin P. (2009) Towards a Formalization of Digital Forensics. In: Peterson G., Shenoi S. (eds) Advances in Digital Forensics V. DigitalForensics 2009. IFIP Advances in Information and Communication Technology, vol 306. Springer, Berlin, Heidelberg

[3] Transparencymarketresearch.com. (2018). *Global Digital Forensics Market to Grow Substantially at 12.50% CAGR from 2015 to 2021*. [online] Available at: https://www.transparencymarketresearch.com/pressrelease/global-digital-forensics-market.htm [Accessed 28 Jan. 2018].

[4] Interpol.int. (2018). *Cybercrime / Cybercrime / Crime areas / Internet / Home - INTERPOL*. [online] Available at: https://www.interpol.int/Crime-areas/Cybercrime/Cybercrime [Accessed 28 Jan. 2018].

[5] Brorsson, P. and Wernebjer Cervinus, D. (2010). *En Jämförelse mellan EnCase och BackTrack*. Undergraduate. Halmstad University.

[6] NIST (2005). *Digital Data Acquisition Tool Test Assertions and Test Plan*. Gaithersburg: NIST, pp.6-47.

[7] NIST (2009). *Active File Identification & Deleted File Recovery Tool Specification*. Gaithersburg, pp.1-5.

[8] NIST (2008). *Forensic String Searching Tool Requirements Specification*. Gaithersburg, pp.1-5.

[9] Pan, L. and Batten, L. (2007). An Effective and Efficient Testing Methodology for Correctness Testing for File Recovery Tools. *Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP 2007)*. [online] Available at: http://ieeexplore.ieee.org/document/4457663/ [Accessed 1 Mar. 2018].

[10] Dftt.sourceforge.net. (2018). Digital (Computer) Forensics Tool Testing Images. [online] Available at: http://dftt.sourceforge.net/ [Accessed 1 May 2018].

[11] dfrws. (2018). DFRWS. [online] Available at: https://www.dfrws.org/ [Accessed 1 May 2018].

[12] Kröger, K. and Creutzburg, R. (2013). A practical overview and comparison of certain commercial forensic software tools for processing large-scale digital investigations. *Mobile Multimedia/Image Processing, Security, and Applications 2013*. [online] Available at: https://www.researchgate.net/profile/Knut_Bellin/publication/258332973_A_practical_overview_and_comparison_of_certain_commercial_forensic_software_tools_for_processing_large-scale_digital_investigations/links/00b495293485547b83000000/A-practical-overview-and-comparison-of-certain-commercial-forensic-software-tools-for-processing-large-scale-digital-investigations.pdf [Accessed 29 Jun. 2018].

[13] Maher, H. (2000). *Online and Out of Line Why: Is Cybercrime on the Rise, and Who's Responsible?*. [online] Available at: https://web.archive.org/web/20001115032300/http://abcnews.go.com/sections/us/DailyNews/cybercrime_000117.html [Accessed 22 Feb. 2018].

[14] Reith, M., Carr, C. and Gunsch, G., 2002. An examination of digital forensic models. *International Journal of Digital Evidence, 1*(3), pp.1-12.

[15] US-cert (2008). *Computer Forensics*. US-Cert, p.1.

[16] Collective work of all DFRWS attendees (2001). *A Road Map for Digital Forensic Research*. [online] Utica, New York: Digital Forensic Research Workshop, p.16. Available at: http://dfrws.org/sites/default/files/session-files/a_road_map_for_digital_forensic_research.pdf [Accessed 23 Feb. 2018].

[17] Philipp, A., Cowen, D. and Davis, C. (2009). *Hacking Exposed*. New York, USA: McGraw-Hill Professional Publishing.

[18] Casey, E. (2004). *Digital evidence and computer crime*. Amsterdam: Academic Press.

[19] Mohay, G. (2006). *Computer And Intrusion Forensics*. Norwood: Artech House.

[20] Champagne, N., Dumont, C., Johnson, A., Castro, J., Leonard, A., Palmer, J. and Craig, T. (2016). Forensic Tool Comparison. [ebook] Burlington, VT: Champlain College. Available at: https://lcdiblog.champlain.edu/wp-content/uploads/sites/11/2016/05/EDITED_Forensic-Tool-Comparison-Report.pdf [Accessed 1 May 2018].

[21] Garfinkel, S. (2010). Digital forensics research: The next 10 years. Digital Investigation, [online] 7, pp.S64-S73. Available at: https://doi.org/10.1016/j.diin.2010.05.009 [Accessed 2 May 2018].

[22] Strengthening forensic science in the United States. (2009). Washington, DC: National Academies Press.

[23] Beebe N. (2009) Digital Forensic Research: The Good, the Bad and the Unaddressed. In: Peterson G., Shenoi S. (eds) Advances in Digital Forensics V. DigitalForensics 2009. IFIP Advances in Information and Communication Technology, vol 306. Springer, Berlin, Heidelberg

[24] Garfinkel, S., Farrell, P., Roussev, V. and Dinolt, G. (2009). Bringing science to digital forensics with standardized forensic corpora. Digital Investigation, [online] 6, pp.S2-S11. Available at: https://doi.org/10.1016/j.diin.2012.05.001 [Accessed 2 May 2018].

[25] Cfreds.nist.gov. (2018). The CFReDS Project. [online] Available at: https://www.cfreds.nist.gov/ [Accessed 1 May 2018].

[26] Dykstra, J. and Sherman, A. (2012). Acquiring forensic evidence from infrastructure-as-a-service cloud computing: Exploring and evaluating tools, trust, and techniques. Digital Investigation, [online] 9, pp.S90-S98. Available at: https://doi.org/10.1016/j.diin.2012.05.001 [Accessed 2 May 2018].

[27] Kiper, R. (2018). Pick a Tool, the Right Tool: Developing a Practical Typology for Selecting Digital Forensics Tools. [ebook] SANS Institute InfoSec Reading Room. Available at: https://www.sans.org/reading-room/whitepapers/forensics/pick-tool-tool-developing-practical-typology-selecting-digital-forensics-tools-38345 [Accessed 2 May 2018]

Victor Andersson

Besöksadress: Kristian IV:s väg 3
Postadress: Box 823, 301 18 Halmstad
Telefon: 035-16 71 00
E-mail: registrator@hh.se
www.hh.se