Towards an Automated Digital Data Forensic Model with specific reference to Investigation Processes

A Survey of Actual and Desirable Practice

By Johan Scholtz

Student Number: 0651823

Department of Computing and Information Science

Auckland University of Technology

Submitted in partial fulfilment of the requirements

For the degree of

Master of Computing and Information Science

Primary Supervisor: Professor Ajit Narayanan,
Head of School of Computing and Mathematical Sciences,

Secondary Supervisor: Krassie Petrova
Programme Leader – Master's Programmes

Contents

Tables referenced

Figures referenced

“I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.”

Acknowledgements

I would like to thank my supervisor, Professor Ajit Narayanan, for his supervision, patience and good interpretation during this thesis. Although we did not follow a normal research thesis structure due to the topic, this thesis still formed well under your guidance.

Professor Ajit's secretary Saide Lo, thank you for helping with conference bookings and other administrative arrangements.

My proof-reader, Diana Kassabova thanks for making sense of my work and adding more readability to my thoughts.

I would also like to thank my wife Karien and my children who had to listen to my endless explanations of research concepts. Just bear with me for a few more years when starting my PhD. research, I will try keeping it brief and only discuss the headings!

Thank you for keeping both the house and I together.

Ethical Approval

MEMORANDUM

Auckland University of Technology Ethics Committee (AUTEC)

To: Ajit Narayanan
From: **Madeline Banda** Executive Secretary, AUTEC
Date: 11 June 2009
Subject: Ethics Application Number 09/98 **Digital data forensics - a critical study of the Investigative Examination Process - towards an Automated Digital Forensic Model.**

Dear Ajit

Thank you for providing written evidence as requested. I am pleased to advise that it satisfies the points raised by the Auckland University of Technology Ethics Committee (AUTEC) at their meeting on 11 May 2009 and that I have approved your ethics application. This delegated approval is made in accordance with section 5.3.2.3 of AUTEC's *Applying for Ethics Approval: Guidelines and Procedures* and is subject to endorsement at AUTEC's meeting on 13 July 2009.

Your ethics application is approved for a period of three years until 11 June 2012.

I advise that as part of the ethics approval process, you are required to submit the following to AUTEC:

- A brief annual progress report using form EA2, which is available online through http://www.aut.ac.nz/about/ethics. When necessary this form may also be used to request an extension of the approval at least one month prior to its expiry on 11 June 2012;
- A brief report on the status of the project using form EA3, which is available online through http://www.aut.ac.nz/about/ethics. This report is to be submitted either when the approval expires on 11 June 2012 or on completion of the project, whichever comes sooner;

It is a condition of approval that AUTEC is notified of any adverse events or if the research does not commence. AUTEC approval needs to be sought for any alteration to the research, including any alteration of or addition to any documents that are provided to participants. You are reminded that, as applicant, you are responsible for ensuring that research undertaken under this approval occurs within the parameters outlined in the approved application.

Please note that AUTEC grants ethical approval only. If you require management approval from an institution or organisation for your research, then you will need to make the arrangements necessary to obtain this. Also, if your research is undertaken within a jurisdiction outside New Zealand, you will need to make the arrangements necessary to meet the legal and ethical requirements that apply within that jurisdiction.

When communicating with us about this application, we ask that you use the application number and study title to enable us to provide you with prompt service. Should you have any further enquiries regarding this matter, you are welcome to contact Charles Grinter, Ethics Coordinator, by email at charles.grinter@aut.ac.nz or by telephone on 921 9999 at extension 8860.

On behalf of the AUTEC and myself, I wish you success with your research and look forward to reading about it in your reports.

Yours sincerely

Madeline Banda
Executive Secretary
Auckland University of Technology Ethics Committee

Cc: Johan Scholtz fzw0309@aut.ac.nz

Abstract

Digital Data Forensics is constantly under scrutiny to standardize processes. Previous researchers moved between various frameworks without presenting a firm platform or solution, addressing standardization. Only a few researchers referred to automated investigation processes. Established data banks do not exist. We investigate whether investigators use forensic frameworks in their investigations. We question if these frameworks are guiding the investigation and the feasibility of an automated investigation model. We also investigate if a prediction based on a global digital forensic data bank is possible. Investigation processes with regard to the readiness of automated investigation is also investigated. Problems encountered are primarily linked to privacy is a major concern. The lack or willingness to address privacy up front, place obstacles in the way of would be researchers. The term automated forensics and automated tools are misunderstood, some participants regard automation as automated software tools and address this as: “Forensic automation is already becoming a problem by giving untrained examiners a false sense of security when in reality, they are not conducting an examination at all” Investigations using software that reflects a click and drag scenario, does not promote an academic research platform.

We suggests automated forensics to be the process of investigation where the investigator make use of previous data based on predictive analysis of data bank from previous data and make use of forensic software in a lesser part. We suggest changing the mindset from “automated software”, to “automated analysis” whereby investigators could sift through the first level of classification and determine sub levels of the investigation with optimal running of scripts, suitable for level comparison and prediction.

(Beebe, 2009) suggests using an Intelligent Analytical Approach extending artificial intelligence and other intelligent search enabling successful retrieval, making use of algorithms. This supports our point of view as well; using a stronger reflection to a semantic vs. literal searching technique should set a base platform, substituting the traditional literal searches.

This also fits well with our vision of having a structured, relational data structure in place thereby improving data indexing. This would ultimately present a match based on “fuzzy hashing” which require a complete paradigm shift. This shift would step away from the overwhelming traditional search patterns and move to prediction of similar cases. We suggest using predictive Markov models, analyzing data for predictive similarity in events. We will also move to a fuzzy re-classification of data models. Since each case differs substantially, a model built from a generic level to predictive sub levels is suggested. This research did not cover relational database creation and classification of variables, further research will be conducted.

In other words, we form predictions, irrespective of the investigation model followed. Further research is required in classifying variables and groups. It is questionable whether forensic investigators would follow standardized procedures at all—considering they are following their own customized methods to date. This presents a problem for standardization and ultimately automation.

Research Problem

Present research does not sufficiently cover how existing forensic frameworks prescribe how investigation procedures should be conducted. I consider this to have a negative impact on the investigation processes and in particular how data verifications conducted during and after investigation. This also challenges data consistency testing and the legality of investigation processes used for models testing.

Based on my preliminary investigation, I will explore the following:

- Whether existing digital forensics frameworks are sufficient to conduct investigations covering all aspects of interdisciplinary associations that reflect a full spectrum of associated dependencies.
- The lack of consistency in reusing investigation data.
- Whether regulation of certification and training lends itself to standardisation.
- Standardisation of frameworks for forensic investigation.
- Confirming and binding of a framework to the investigation.
- Automated investigation procedures.

Assumptions

Existing digital forensics has a vague and confusing terminology and is not always clear about the terms. For instance, the terms Framework Model, Method, Procedure and Process are often misunderstood. The digital forensics discipline lacks clear terminology descriptors. From my initial perspective I assumed enough data would be available from participants in the forensics field. In this thesis I selected the following descriptors.

Terminology descriptors

a. Framework: a hypothetical description of a complex entity or process. It is a basic conceptual structure to allow the homogenous handling of different objects.
b. Model: Abstraction of a real-life system used to simplify understanding and to aid in decision making.
c. Method: Systematic arrangement and design, also an orderly procedure or process, or a regular manner of doing anything.
d. Procedure: a particular course of actions intended to achieve a result
e. Process: perform mathematical and logical operations on data according to programmed instructions in order to get the required information.

In this thesis I use these descriptors as follows.

- *Framework:* defines the conceptual understanding that a basis exist with defining terms describing a wide platform of disciplines, assuming linking between these are possible based on a theoretical perspective. Therefore, in a digital forensics framework, we need to consider influences from similar disciplines that play a subjective role covering points related to digital forensics, specifically relating to interdependencies. Researchers presently need to study this to set up a reliable platform of cross-referenced frameworks.
- *Model:* has a stronger definition than a framework. In essence it represents a more defined descriptor of something specific, a system that is composed of entities that can be grouped, classified or interrelated.
- *Procedures: these* are as the name suggests, i.e. from a model we can build procedures defining actions that run in a model, whereby categories or groups are assessed.
- A *process:* this is the finer detail of a procedure, thus we note specific actions assigned to specific operations which are carried out with specific instructions resulting in a logical output.

Research Initiative

On completion of a few readings it became clear that the existing digital forensics frameworks do not provide clear guidelines for conducting digital forensics investigation. However, had a framework existed, an investigation based on known standards and procedures would prescribe a standardised investigation platform. On the other hand, this points to all investigations following a set method for comparisons; ensuring future investigation is following one standard.

A few frameworks currently exist that address basic investigation processes, but as digital forensics is a *relative young* discipline they do not show enough adaptability to the ever changing digital forensics science. This raises further questions when proposing a framework that functions according to strict and consistent procedures and processes. It has to be noted that “relative young discipline” when referring to digital forensics is an obsolete term and although used by most researchers, it has been in use for about 15 years now.

Chapter 1 - Typical Digital Forensics Case – Embezzlement

The following section sketches a fictitious digital forensics case. I present a case that has digital forensics grounds for investigation. I also point to some difficulty in ordering questions against the defendant reflecting whether the investigator is following correct standardised procedures. I will then show how types of questions vary depending on the investigators' experience; this again emphasises inconsistent investigation procedures.

1.1 Fictitious Digital Forensics Case

"Alleged stolen company secrets"

James asked you to defend his case against Keith. Keith is a digital forensic expert. This is a summary of the events leading to the investigation.

High Trust Investments (HTI) had a strong investment portfolio for the last few years, up to December, when there was a sudden decline in their stock value. A few major investors cancelled their investments and subsequently HTI had great losses. The directors of HTI became suspicious that something was wrong and asked a forensics accounting firm to examine the situation in the company. Keith, an upcoming forensics investigator will conduct this forensics audit. HTI's human resource department confirmed that three staff members were made redundant in December following a staffing review. One employee that draws attention is James. James is an American citizen working in New Zealand (in this case we assume all servers, personal computers and laptops are domiciled in New Zealand, therefore local laws of New Zealand applies) for the past five years and had access to high-level investment information. Although James received a three-month redundancy packet, it also became clear that he was unhappy with the redundancy arrangements and threatened to take steps against HTI. Keith was asked to study the digital footprints of this employee in the company. When James went overseas for meetings he had remote access to HTI's servers. James is computer-literate and was well-known for his skills and knowledge of the company

network. He would occasionally help the network administrator and work with him in the server room.

1.1.1. Data gathering

Keith started gathering data of potential activity from James' office desktop computer. When Keith collected the desktop computer at HTI's head office, neither James nor an office manager was present that could give permission to collect the computer. However, Keith spoke to another staff member and got permission to access James' office. Keith found the desktop computer switched on and followed procedures shutting it down and after that unplugged the computer from other peripherals. Keith decided to extract data from James' desktop's hard drive and made a bit-stream image. He also noted that an external hard drive was still running while connected to the desktop computer. Keith decided to unplug the external drive and take it to his office for examination. Keith only then released he did not record his progress, neither did he take photos of the computer setup; he then took a few photos. Keith also followed procedures securing James' laptop. Before Keith left the office he also made a bit-level image copy of the laptop's disk, crypto graphical hashing and processed the disk as a whole with all directories, files and disk sectors.

1.1.2. Investigation

Back in his office Keith started examining both images from the desktop and laptop hard drives. He used proven forensic software finding hidden, deleted and encrypted files. He also examined backup data on James's laptop and examined password protected files. Keith noted a high-level of network activity while James was in Germany before the financial meltdown of HTI. Keith discovered from logs and IP address scans that James clearly had access to a data bank of a European investor group. Keith also noted a suspicious plug-in placed in a hidden folder that triggers weekly data queries from his laptop to a remote IP address. It was also determined that the IP address was from the overseas investor group and data was also sent to James's work desktop computer. James had stored the external data in a hidden password-protected folder on his work desktop.

Keith connected the external hard drive obtained from James' office desktop directly to a test computer, but forgot to add a data blocker between the external hard drive and the test computer. However, he managed to make a full mirror of the drive before he started with other investigation processes.

Keith accessed the external hard drive contents and found a password-protected file in a hidden directory. Using his own (untested) scripting code, he tried to break the password and eventually got access to the file after a few hours. He found out the file contained backup data of most investor accounts. Time stamps of the file creation properties suggested data was backed-up daily. Keith also found file logs, showing activity that suggested data files had been opened from remote site for extensive periods, and showed unique IP addresses and data packets were sent. This matched the hidden data sets from both the laptop and desktop.

Based on these findings, Keith wrote a report for HTI which implied that James had access to confidential investment information and he leaked this to an external investors group in Europe. James was interviewed by senior HTI staff and was told HTI was taking a legal action against him.

1.1.3. Defense Preparation

In order to recognize specific reference to a series of escalating questions, I compiled a "Level Guide", explaining how the different levels would monitor Keith's' competency, and explore deeper investigation procedures.

Fist Level

First Level guidelines would give us an initial overview of actions taken by Keith and his level of competency to undertake this investigation. Typically responses to these questions are factual and might require clarification

If we were to defend James in court, we would have to determine whether this case adhered to first level guidelines. This would test Keith's investigation methods, for instance:

- the investigators' background;
- assessment of the case - review of procedures and processes used;
- recovery processes, imaging - acquisition processes;
- analysis of the authentication process;
- Findings and documentation.

You might recall from the case study that Keith made a few mistakes while gathering data. This might later present an opportunity for James to argue the correctness of the procedures.

Second Level

Second Level questions have been taken verbatim from participant feedback. It is noteworthy that the same range of questions could be put forward in the fictitious case as well. In other words, these are real topics of the days from real investigator data. The results of the investigation are presented in Appendix A. I will show the relevance of these questions to participants' feedback, standardisation and certification in Chapter 8.

On a second level, to test the reliability of Keith's evidence, James' defender could also ask questions about the appropriateness and relevance of actions during the investigation, for instance:

- Did you use a clearly described framework or model in your investigation process to ensure a systematic approach to your investigation?
- Did you use any known theory as a guideline for the case investigation or did you just followed a hunch?
- Did you use previous investigation data as a guideline to conduct new investigations or you extracted the data as you advanced?

- Do you think the type of software you use to conduct forensic IT investigations plays a decisive role in extracting evidence that is critical to your investigation and error-free is that software?
- Do you think that having a database (corpora) of previous cases can help in digital forensics investigation and case analysis and if so why did you not use it?
- Do you think an automated digital forensics investigative process is possible so every case is treated equitably, fairly and consistently?

Third Level

On the Third Level, I include an even wider range of questions; issues can be raised about the repeatability and reliability of digital forensics investigation processes.

- Please describe your training and processes you followed in gathering evidence.
- Which procedures did you take to safeguard the evidence from external interference?
- Have you conducted typical investigations like this before?
- Describe the acquisition process and chain of custody in this case.
- Did you use off-the shelve recovery software?
- Did you correlate the different time zones to verify the time of the alleged transfer of data from James' laptop?
- Did you check server logs from the local HTI server indicating the actual length of connection duration could successfully transfer data in the available time to the overseas link?
- Did you follow recognized procedures to backup the data from James' personal computer and external hard drive?
- Did you analyse the system with minimal invasiveness?
- Since James is a US citizen, would James be on trail under New Zealand law or US law?

- The Fourth Amendment to the United States Constitution states: “The Fourth Amendment requires that all warrants particularly describe the place to be searched and the items to be seized. To pass constitutional muster, a warrant (1) must provide enough specific information to guide the officer's judgment in selecting what to seize, and (2) the warrant's breadth must be enough narrow to avoid seizure of purely unrelated items.” How does this protect James against unreasonable searches and seizures?
- Did Keith follow the correct shut down procedures?
- Did the self-written (untested) scripts for to accessing the external hard drive risk James’ chances for a fair trial?
- Does probable cause play a role? Did the court instructions only stipulate the seizure of the desktop and laptop computer?

  It might be argued the seizure of the external hard drive was improper. *(On the other hand, Keith had probable cause to believe the external hard drive might be useful in this investigation, since it was switched on and connected to the main computer).*
- Did Keith act within the scope of the warrant?

Disclaimer: The chosen case scenario is an example only and any association to an actual case and litigation is purely coincidental. Names and locations in this example are fictitious and are not intended to reflect actual people or places.

To put the **Level Guide** in perspective, I presented some verbatim questions to participants at the Second Level. Although questions in this section appear to be fabricated, we will see from my research how these questions reflect answers with varied responses by participants. When I relate these (albeit generic) questions to the real data collected from participants, none are presenting the same answer. This shows that participants are not using the same investigation approach in the same situation. We can make the assumption that the Level described in the Level Guide act in similar manner in the fictitious case as would the real data presented by the participants'' show as well.

I also tried to match questions and corresponding participants' responses to existing investigation platforms, expecting a match to a prescribed standard. However, results were different from participants since most participants had diverse interpretation of cases that were based on non-existing frameworks. *Arguably, if all investigators were adhering to a rigorous technical and academic certification; investigation results would produce consistent data interpretation. To the contrary, no clear pattern emerged from this as shown in the research findings in Appendix A, Section A.*

From Section 2, as stated in the Chief Police Officers (ACPO, 2007) publication, guidelines are set in place explaining the importance of investigating procedures. When I compare this guideline to the investigation conducted by Keith, it shows obvious mistakes made by Keith while conducting the investigation. We may question Keith's competence and affirm that his investigation lacks systematic and rigorous investigation processes. Keith's current answers would differ from the "ideal" set of answers he provided, had he followed a standard or acceptable guideline. Moreover, Keith would have done better if he followed an automated investigation process. He might have obtained suitable answers if he had used an automated investigation model based on an initial template approach.

I am aware that core variables for any investigation differ, but given the correct data structure, data entry would be possible in future when designing a dynamic relational data structure. This is another topic for future research.

1.2 Investigation Principles.

Buskirk (2006) mentioned that a computer expert should be able to present suitable answers when questions about reliability and software are disputed. In a recent court case Kumbo Tire Co. vs. Carmichael, (1998) emphasis was placed on technical reliance that extended from previous cases referenced. This showed guidelines for digital forensic evidence.

These guidelines are:

- whether the theory or technique has been reliably tested;
- whether it has been subjected to peer review;
- the known or potential rate of error of the theory or technique;
- Whether the technique is generally accepted.

From the above, similarity exist to our Level Guide in the previous section. *Keith's competency would be questioned based on procedures followed in this investigation.*

If reasonable doubt is caused by the reliability of the software or testing procedures, then this might lead to the acquittal of the defendant. Since there are literally hundreds of scenarios, procedures for each case would also change. Forensics expert witnesses must ensure a high level of responsibility and experienced in presenting a firm case.

Four principles about the recovery and investigation of computer-based evidence are described in (ACPO, 2007). *These principles intended to guarantee the integrity of evidence and to allow accurate reproduction of results that would remove any doubt and would not or present an opportunity for challenge in court.*

Principle 1 - No action taken by Police or their agents should change data held on a computer or other media which may subsequently be relied on in Court.

Principle 2 - In exceptional circumstances where a person finds it necessary to access original data held on a target computer, that person must be competent to do so and to give evidence explaining the relevance and implications of their actions.

Principle 3 - An audit trail or other record of all processes applied to computer-based evidence should be created and preserved. An independent third party should be able to examine those processes and achieve the same result. This record must be repeatable to an independent third party.

Principle 4 - The Officer in charge of the case is responsible for ensuring the law and these principles are adhered to. This applies to possessing, and access to, information contained in a computer. They must be satisfied that anyone accessing the computer, or any use of a copying device, complies with these laws and principles.

These principles are widely used as (Kennedy, 2006) suggest. These principles are best practice guidelines on the recovery of digital evidence. One of the essential ingredients of a witness statement is to directly state that guidelines had been followed. Any changes and all reasoning or actions that fall outside the prescribed principles should be documented. Consistency in reporting should be present and should reflect the trail of the investigation.

Another question addressed in this thesis is as follows: What is the relationship between the “principles” and the three levels of questions asked by James’ defender in the fictitious case study? We expect that a digital forensics investigator has integrity and a good professional reputation. But can these expectations be fulfilled by frameworks, guidelines and principles adopted by current digital forensics investigators?

I address these questions again in Chapter 2 and Section 8, showing the importance of standardisation, certification and quest for automated investigation.

1.3 An Overview of Existing Digital Forensics

Previous research covered several approaches to setting a forensics framework, which are adaptations of previous models. I found that only a few models present a framework that defines or delivers qualified similarity between the different disciplines. From this possible pattern analysis from different disciplines is possible. (Kohn, 2007)

I also explored whether existing forensics frameworks could act as a sufficient platform to create automated forensic data procedures. I researched several frameworks and expected to find a model that would make data extraction from corpora possible and to group data for clear investigation analysis (Aanya-Isijola, 2009). This would allow analysis of expected clustered data. I also examined the feasibility of an automated digital forensics template that presents investigation processes from different perspectives, this is especially presenting alternatives under a different search patterns as (Beebe, 2009) suggests. I then discuss existing models and attempt to find possible features that might be useful for the development of an automated framework. Discussions in Chapter 7 and 8 cover automated investigation. Binding features of alternative procedures to investigations and presenting a new model of standardisation was expected, although a drawback in existing digital data corpora influences recommendations in establishing a framework that handles automated predictions.

This however proves to be difficult as results from participants' responses show.

A few interesting research activities:

- Existing forensic investigative frameworks do not allow automated fast tracking of digital data investigations. Finding associated data clusters of specific digital forensics activity is still being researched. (Olivier, 2009). Oliver also suggests using a multidimensional model compared to one-dimensional file systems, making alternative relational structures possible.
- It is the author's view that a completely different approach regarding investigation should be taken. I suggest a bold step into automated

investigation procedures, linked to a digital forensics data base where keyword searches are possible.

- Automated searches could only be established if a data bank exist and searches per keywords are possible. This would enable other disciplines to find touch points to the case scenario associated to their own investigation procedures. For instance, creating a multi-disciplinary relational structure that includes networking and real-time mobile disciplines might create links to other processes, linking to similar case scenarios. From this perspective, digital forensics researchers can create specific scripts capturing mismatched or irregular pattern analysis of digital data that allows a customised automated investigation framework.
- I also recorded whether investigators use a specific framework and if changes to their frameworks, or hypotheses would influence the way they conduct future investigations. It is assumed that investigators should at least follow a methodology for archiving their investigation processes while they are performing their investigations.
- I explored if investigators conducted-assessable and verification based investigations, based on their initial hypothesis.
- I questioned investigation process that did not set a standardised framework in place through detailed procedures and processes.
- I promoted the importance of creating forensics corpora that forms a reproducible forensics platform with accessible real data.
- I researched the feasibility of a framework that could potentially lead to the development of new procedures and processes, deploying an automated forensic investigation. It is envisaged that a workable automated template would support a scientific investigative process by setting up a firm scientific research platform. This framework would test digital data forensic processes. Comparing the different forensic software tools to obtain, verify and test digital data, would also confirm processes which are reliant on software for gathering data.

1.4 Hypothesis

From my initial research I established that existing digital forensics frameworks do not allow automated investigation procedures. In order to test the hypothesis, I question the establishing of an automated digital forensics framework and prescribe specific processes that could evolve into standardisation, guiding investigators to use the same model for validating all cases. I also suggest that global forensic data contribution is necessary to establishing a reference data base (or corpora) which will only become feasible if technicians moved towards an academic approach for investigation analysis.

Since the hypothesis is related to a few basic indicators, I describe these as follows:

- This research tests the feasibility of a framework (based on a template approach) that would enhance investigation processes. It takes a tiered approach of investigating the responses that classifies criminal activity or suspected interference. *These assumptions were difficult to confirm from the collected data, because only a few participants had used a framework before or during their investigations.*
- From the above I established whether existing digital forensics frameworks met forensic investigators' expectations about completeness and if creating a standardised automated template is possible. I also paid particular attention to resolving the apparent tension between investigation groups that appears to have a negative impact on the forensics discipline cohesion.

Additional tasks that support the hypothesis

- From the questionnaire, I tested whether participants' were willing to participate in the creation of a feasible forensic data bank and whether gathered case information would present options to fast track investigation procedures.
- I also investigated the feasibility of an improved framework, formulating data extraction and grouping data from the investigation process as this could present a predicative analysis of clustered data.

- Discussed existing models and possible features that might be useful for the creation of an automated framework and suggested alternatives to the traditional classification of digital forensics crime, recommending a framework that handles automated predictions.
- Attempted to prove whether existing forensic frameworks could act as a sufficient basis for establishing automated data forensic procedures. This suggests the creation of forensic corpora that forms a reproducible forensic platform with accessible realistic real data.
- Investigated the feasibility of an automated framework that could potentially lead to new standardised procedures and processes. It is envisaged that a workable template assisting investigators would also assist scientific investigation processes thereby establishing a firm scientific research platform. This framework would validate previous or existing digital forensics processes by comparing different forensic tools.
- Suggested a template, based on existing forensic data and created a framework for the first stages of the investigation process. This template would assist investigators conducting scientific analysis and by verifying if the investigation results match their initial hypothesis. From this I propose a new method for conducting scientific investigative processes.
- This research proposes a framework that improves the Investigative Process, consisting of stepped levels of responses to enable classification and confirmation of criminal activity or suspected interference.

1.5 Methodology and Rationale for Research

I tested whether digital forensics investigators used a set method to investigate their cases, based upon a valid framework. I also investigated if frameworks are used at all by investigators. A classical research methodology was followed that involved, setting a hypothesis, gathering data, evaluation and discussion of results conclusion of findings.

However, the actual research process is more complicated and interwoven than that, with preliminary data gathering leading to refinements of the hypothesis and certain aspects of the discussion leading to further research. Never the less, for the sake of presenting a structured thesis, the classical method is adopted.

Test data collected shows that existing frameworks using specific processes in practice while conducting digital data forensics investigations. I did not research whether the processes adhere to specific laws from different countries and how investigation guidelines show differences in interpretation. Furthermore, I did not check if these were used in court proceedings. From the participants' contributions, we assumed a high level of legal interpretations, since a few participants are from law enforcement backgrounds.

I discussed details of various research findings and required steps as well as other alternatives as suggested by researchers. In addition, I attempt to match a layered approach to participants' methods to determine if the investigation processes they followed could be standardised. These processes are based on particular procedures, which are grounded in a broad investigation platform. I expected to obtain results from participants showing their preference for investigation processes and to match their outcomes to a specific investigation methodology.

Analysing, defining and comparing these steps form part of the expectation to investigate the feasibility of an automated framework. Once these steps are broken down to a layered approach we might see how other disciplines could interact and influence the way investigations should be conducted. In the rationale we consider what alternative methodological tools might have been employed (particularly those employed by related studies), with advantages and limitations. This indicates that an established data bank would allow comparisons between existing case data previous final findings, thus allowing some form of prediction.

Using a custom designed questionnaire, I expected to position investigation processes based on final investigation results and then compare methods used for obtaining these findings in relation to their success ratio.

I then attempt to match their investigation methods to a template thereby building a method that incorporates procedures. This would then be used to establish a procedure sampling method that acts as a template to conduct investigation in the future. I am aware that this might not be achieved, since digital data would vary from one case to the next. For instance, using a questionnaire in a structured interview might be different from designing a pilot questionnaire used in the actual survey. Furthermore, a structured interview is a quantitative research method commonly employed in survey research to ensure that each interview is presented with exactly the same questions in the same order. In other words, answers can be reliably combined and comparisons can be made with confidence between sample subgroups or between different survey periods.

In my research I did not use sample groups as the sample groups would not have sufficient experience in this specialised field to present meaningful results.

Chapter 2 - Forensic Background

Digital forensics is lacking confirmed and tested methods; this is clear from substantiating digital forensics procedures. This underlines the urgency to standardise processes, to ensure proven and consistent results. Digital Forensics Science needs to take a bold step towards a new approach for defining and standardising investigation processes and to make this a confirmed platform for adaptive globalisation. This thesis identifies a major current issue for digital forensics, namely the need for consistency versus current practice. I also discuss current expert expectations of available frameworks. When I compare this to academic research that prescribes a standardised platform, I notice an obvious mismatch between the expectations of forensic technicians. This highlights the problems for setting up a standardised platform.

I also explored some existing frameworks and suggest initiatives related to investigation methods. For instance;"are automated investigation methods possible?" If such investigation method was followed it would inherently support a standardised framework. Investigations would then be carried out faster based on standardisation that allows recognition of both forensic technicians and academic perspectives. I address this issue further below when analysing the responses and interpretations from different participants.

In this chapter, I reflect on various aspects of the digital forensics that make up the forensics discipline. In order to understand the complexities of previous research covering digital forensic history, I covered this topic under the following headings:

- Digital Forensics Background
- The Need for Forensics
- Investigation guidelines
- Phases of Forensics Comparison
- Presentation of Digital Evidence

2.1 Digital Forensics Background

Computer Forensics only became recognised in the later part of 1999 and more in the last 5-8 years when workshops tried to establish a definition for forensics as a whole (Garfinkel, 2009). Some issues, such as establishing a framework to investigate digital forensic processes were raised. Since there was not a clear framework in place, it became a challenge to position forensic science in relation to other disciplines. Investigators at the time suggested that data should be collected, preserved, validated, identified, analysed, interpreted, documented and presented. Suggestions also included the importance of tracing events back to the user or originator thus allowing preventative actions.

Since 2001, when a major initiative was launched by the first Digital Forensic Research Workshop held in Utica New York, no particular process has been accepted as the ultimate answer in this area. Discussions have mostly been held around the establishment of a framework to understand the methodologies used to guide scientific processes in this discipline. Issues about the trustworthiness of evidence as well as detection and recovery of hidden data play a major role as well the importance of networked environments. (Palmer, 2001).

Digital data forensics as a discipline addresses crimes committed in a digital environment. Problems associated with digital data gathering and its validation means the discipline is continuously changing and adapting to new cases and technologies. This also occurs because new challenges are presented by illegal or fraudulent actions aiming to hide or destroy data tracking. Data hiding and anti-forensics processes are constantly evolving in order to derail any proof of evidence or existence.

According to (Carrier, 2003) “A forensic Investigation is a process that uses science and technology to develop and test theories, which can be entered into a court of law, to answer questions about events that occurred”. In particular digital forensics is about the examination of digital objects, which entails specific reference to any potential of developing new standards and testing procedures to challenge existing and accepted methods.

One argument has been that there is not an established taxonomy in forensic computing, and the discipline has failed to combine and leverage the strengths of individual disciplinary investigations of particular forensic issues. Hannan, (2003) proposed a taxonomy that included multiple dimensions and sub-categories.

2.2 Can we standardize Forensics?

The *Research DFRWS Technical Report*, (Palmer, 2001) refers to the core function of data forensics as to detect and extract data based on exact processes. Methods used by forensic experts might influence the validity of the obtained data and might be disputed in court. Forensic analysis can only be successful if data is presented in such a way that no doubt exists about the link between the user and the actual data. Investigators experience uncertainty when presenting data in court, this highlight that standardised forensics processes are not in place (Hall, 2005). This leads to further uncertainty in court when deviations from previously accepted standards or frameworks do not adhere to standardised forensics criteria of classification. If investigators use a standardised platform, this would enable them to find and verify data according to set guidelines and would help them make informed and tested decisions about valid case data.

Since every case is different from the other, common references might only be found in a generic investigation model or framework. If investigators could follow one standard or guideline as starting point and break this down to sub-levels in the investigation processes, one might see standardisation on the same entities. Investigators should have experience in positioning or aligning investigation guidelines as required for the particular scenario.

We also cover how the digital forensic market defines “forensic experts” and how this portrays representation in courts when investigators do not share the same level of the expertise. When we consider frameworks that integrate methodology processes, principles and resulting evidence, we note that assumptions can confuse investigation teams when terminology is vague and only constructed for a top-level framework.

It is suggested that an initial framework would only suit a broader generic context. (Selamat, 2008). When deeper investigation levels are required, generic frameworks does not present a clear indicator of all levels. Therefore, precise verification based on the detailed features of each case becomes vague as the investigation develops. We also cover some of the industry requirements for forensic experts and how this term can be potentially misused in the forensic industry. When we compare integrated frameworks between methodologies, processes, principles and resulting evidence, we note that a generic description of digital forensics only presents coverage in the top level forensic framework. Deeper investigation requires a comprehensive understanding at all levels. Thus, escalating processes presents a defined way of verification based on the specifics and levels of each case

2.3 Investigation Guidelines

When a general formulation of a standardised methodology or framework exists, these models would assist the non-forensic expert to conduct investigations according to specific guidelines.

We see this issue emerging from research by Garfinkel (2009) who states that digital forensic science is not yet a true science because the research community has not as yet adopted understanding and rigour of reproducible test results. This is clear from a wider understanding when perspectives on re-creating forensic corpora are considered. Although a few frameworks try to establish a model to which certain procedures could be linked, most procedures still do not deal with the core issue – the gap between technical aspects of digital forensics and the judicial process. This is an unqualified problem as the argument between technical specialists and legal practitioners' boils down to who knows best.

The underlying concepts of reconnaissance, reliability and relevance are beyond the ability of the existing framework the most important to establish a link between factual information and judicial review. (Leong, 2006). We note that at least one generic framework exists that, defines the phases of investigations. This is a basic framework and does not allow for deeper investigation processes.

Data from our participants (Figure 1) suggested that seven are fairly satisfied with the existing investigation process or are borderline satisfied with the investigation processes. Two participants are not satisfied at all and only three are very satisfied with the existing investigation model. It is worth noting that a good match exists between participant's B answer to this question, question 6, in Section A. "Are there any aspects of current investigative processes that you would like to see changed?" to which participant B also answered NO. This indicates that no changes to the existing investigation process are needed (Detailed results are provided in Appendix A) (This refers to the last section and findings)

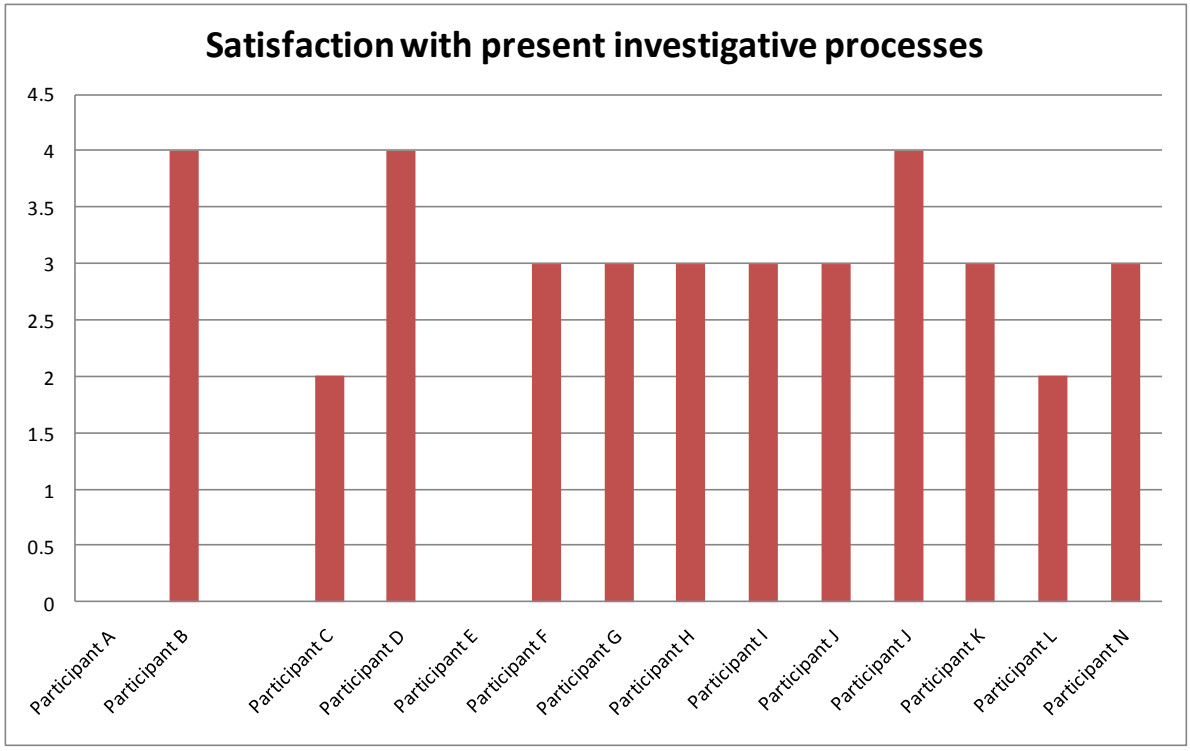


Figure 1 - Satisfaction with present investigation processes.
A higher score indicate satisfaction with the existing Investigation Processes

Adding additional perspective to the previous question, we observed an interesting answer to the next question:

"Do you use an explicitly described framework or model in your investigation process?"

Response: Yes: 4 No: 9

Nine participants out of 15 indicated that they do not use a framework, one had no response and four said YES. Although nine said NO, they indicated that

they use one or another kind of procedure. When we compare this to the question related to processes used, we note the same participants never had clear steps or any steps at all in their investigation processes.

2.4 Phases of Forensic Comparison

We note that as a common digital forensics framework does not exist some wider ad hoc procedures develop as the investigators attempt to fit processes to the investigation. Although this seems to be a customised approach to a specific scenario, this method could force the investigation process in a direction preferred by the investigator and might not cover all aspects of the case.

From the lack of information and limitations of procedural, technical, social and legal interpretation, it is clear that an investigation model should provide alternative but adaptive investigation processes. Investigators should at least have a minimum investigation process in place, thus adhering to a required standardisation of this particular model. All these processes should be conducted in the ever present legal aspect of validating assumptions with proof of evidence and a chain of custody that reflects procedures and processes in place that can be tested for authenticity. Investigators and forensic researchers could find it almost impossible to prove authenticity of cases if data is not made available for investigative purposes or when no data bank exists from which researchers can pull information for validation and research purposes.

Leong (2006) suggested a better-defined forensic framework. To reduce the incidence of incorrect conclusions based on unreliable or incorrect data, it is necessary to quantify uncertainty and correct it whenever possible. In addition to using corroborating data from multiple, independent sources, forensic examiners should also attempt to rate their level of confidence of the relevant digital evidence. Using this systematic method to research conclusions would help decision makers assess the reliability of the information they are investigating.

This investigator could anticipate the challenges that would be presented in courts as attorneys would become more familiar with digital evidence describing measures taken to document and minimise loss of data.

Leigland (2004) promotes a mathematical description of a forensic model that changes the way traditional models firstly construct an informal model of procedures that might influence the effectiveness or integrity of the investigation as a starting point. Leigland mentioned Procedural, Technical, Social and Legal issues as major deficiencies of the research conducted by Spafford (2001). This came as a surprise, since Spafford addressed these issues from a wider perspective as being influential to the investigation processes. Spafford also directs challenges in the procedural, social, and legal realms, in order to craft solutions that begin to fully "heal" rather than constantly "treat" our digital ills. Spafford (2001).

From these procedures we could simplify the term procedural as the compliance of data gathering. Since data could be many gigabytes, each stage has to be clear and procedures have to be followed. This presents challenges at all stages, from gathering the data to storing and finally analyzing the data. One common approach is to extract only relevant information while the system is still running, which limits the amount of data gathered. From a social perspective, the lack of standardization of procedures has led to uncertainties about effectiveness of current investigation techniques. Additionally, privacy concerns about investigation suspects can hamper the forensic process. Legal interpretations of digital evidence in legal proceedings are continuously challenged. This is because the methods of gathering evidence have been informally compiled by forensic investigators with different interpretations and personal experience.

Uncertainty in a set framework or model matches my research, indicating that some investigators do not follow a standard at all and likely follow own protocols. Some investigators' techniques may not be rigorous enough to use in the courtroom and does not portray consistency and trust in procedures. In other words some procedures are not repeatable in similar investigations, or are not testable if the same test would be run again.

In contrast to Spafford – Kent mentioned the following; Kent (2006) describes the basic phases for performing digital forensics as follows:

- Collection: Identifying, labelling, recording, and acquiring data from the possible sources of relevant data, while following procedures that preserve the integrity of the data.
- Examination: Forensically processing collected data using a combination of automated and manual methods, and assessing and extracting data of particular interest while preserving the integrity of the data.
- Analysis: Analysing the results of the examination, using legally justifiable methods and techniques, to derive useful information that addresses the questions that were the impetus for performing the collection and examination.
- Reporting: Reporting the results of the analysis which may include describing the actions used, explaining how tools and procedures were selected, determining what other actions need to be performed (e.g. forensic examination of additional data sources, securing identified vulnerabilities, improving existing security controls) and providing recommendations for improvement to policies, procedures tools, and other aspects of the forensic process

Use of a "traditional forensic framework"

Figure 2 presents marginal differences between participants' views on the importance of using a traditional framework model. The meaning of Reporting might also be considered by participants similar to preparation and awareness. Participants consider Data Comparison as the last activity, or the least important in the investigation process.

This supports our suggestion that digital forensics does not conduct data comparisons because no corpora exists to compare it to.

Traditional data comparisons to test findings have never been conducted because it never has been established. We observed a good combination of traditional framework processes, for instance, Reporting followed by Data Gathering and Analysis.

When we look at Data Comparison as the least important, this fits results from Section B, question 5, indicating that if participants are not contributing to a database; they have nothing to compare their results against. It is clear that data comparison is not important to all investigators. Noteworthy is the importance of Post-Investigative analysis. This response is in fact contradictory to the actual benefit of collecting data, since only one participant uses the data effectively after collection for analysis.

From a post-investigative perspective, this fits well with 50% of participants who indicate they collect previous data.

Do you collect data of previous cases currently?

Response: Yes: 6 No: 6

Six participants collect data, but only Participant C uses the data to test examination techniques – post investigative analysis.

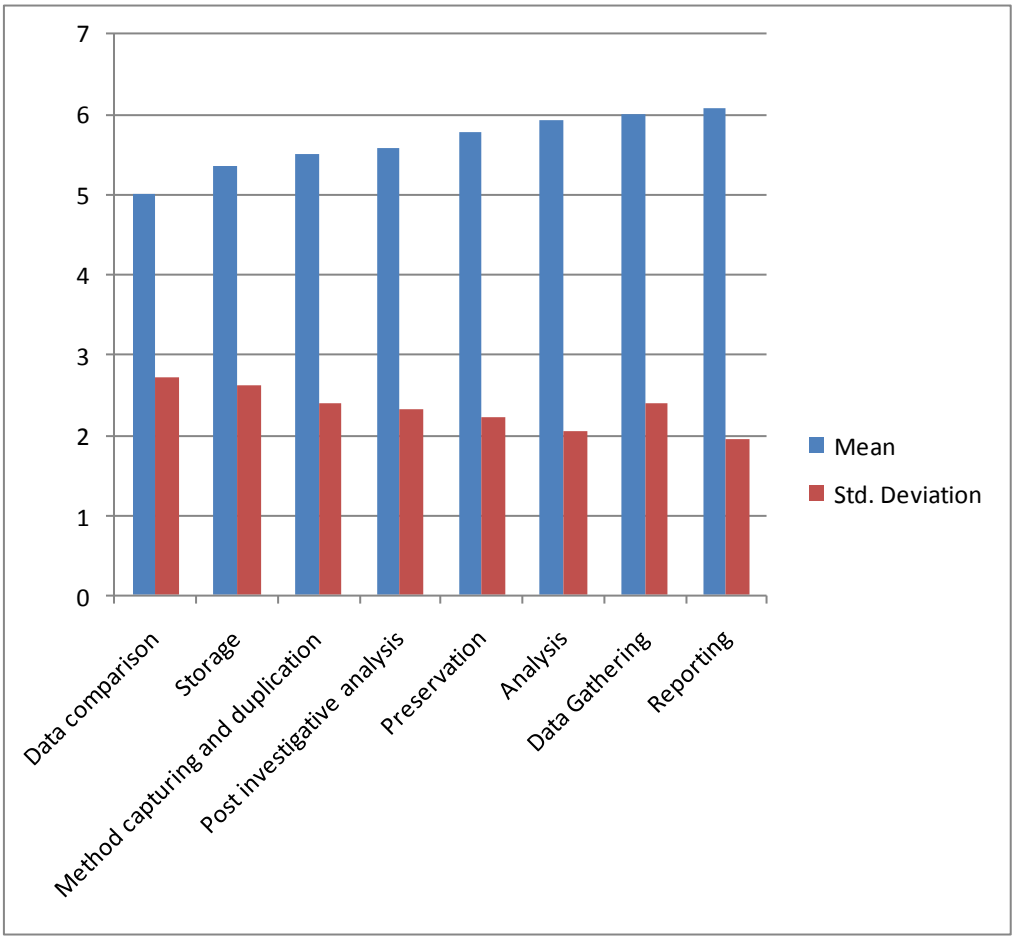


Figure 2 - Importance of using a traditional framework model.
A higher score indicates higher importance.

2.5 Presentation of Digital Evidence

A Comprehensive Approach to Digital Incident Investigation by Peter Stephenson, describes a method called End-to-End Digital Investigation or EEDI (Stephenson, 2003). This process allows investigators to apply a structured investigation technique that combines the use of computer technology with traditional investigative methods. In trials during actual investigations as well as in presentations to law enforcement and civilian practitioners EEDI has received a positive response. The processes of collecting and presenting data would be researched under these guidelines.

The National Association of Criminal Defense Lawyers suggests a Four Phase Forensic Comparison model (Tobin, 2006). In order to draw meaningful conclusions from a comparison between physical items of evidence, a forensic scientist generally must engage in a process consisting of the following four phases:

Phase 1:	Sample evaluations and analyses.
Phase 2:	Comparison of samples (“matching” or “grouping”).
Phase 3:	Assessing the likelihood (relative frequency) of the matching features in the relevant population from which the evidence originated.
Phase 4:	Conclusions about whether the samples have (or are likely to have) a common source.

Weakness at any phase of this process can undermine the reliability of the expert’s ultimate conclusions.

Ibrahim Baggili and Matthew Kiley (Baggili, 2007) state that the current approach of expert determination relies heavily on forensic software certification and places very little emphasis on education in Digital Forensics. This impetus should drive international efforts towards the standardisation of international forensic laws for the international community.

Chapter 3 - Models in (a Complex) Digital Forensics Frameworks.

This chapter makes an attempt to match existing frameworks to a process model as found under a forensic lifecycle. We also try to simplify the forensic lifecycle by breaking it down into the following sub headings:

- Establishing and Organising a Forensics Capability
- Classification Problems
- Existing Framework Similarity

3.1 Establishing and Organising Forensics Capability

Although this thesis covers a vast range of existing theories, it provides by no means a complete list of forensic frameworks, since there are other researchers continuously trying to re-invent the wheel when it comes to forensic framework discussions. In Table 1, a wide range of basic frameworks had been developed over the years; these frameworks covered varied interpretations by researchers.

Digital data forensics has been on a fast track in terms of setting up a defined framework and is still a young science (as Ahmad (2006) shows), compared to other sciences, with many positioning and shuffling proposed methods and trying to get their convincing statements across. We also note this from a study by Brinson (2006) that refers to the infancy of cyber forensics. The lack of awareness about various related disciplines hamper the forensic investigation process because no framework exists so far that caters for investigation processes and combines academic development of specialist skills but still covers other disciplines as well.

We note several emerging frameworks; Carrier (2004) focuses more on an event-based digital framework, whereas Beebe (2005) proposes a framework that is objective-based. Both these frameworks in essence still refer to the traditional model initially proposed by Palmer (2001).

In contrast, we note that Baryamureeba (2004) suggests another approach that separates the investigations at the primary and secondary crime scene while

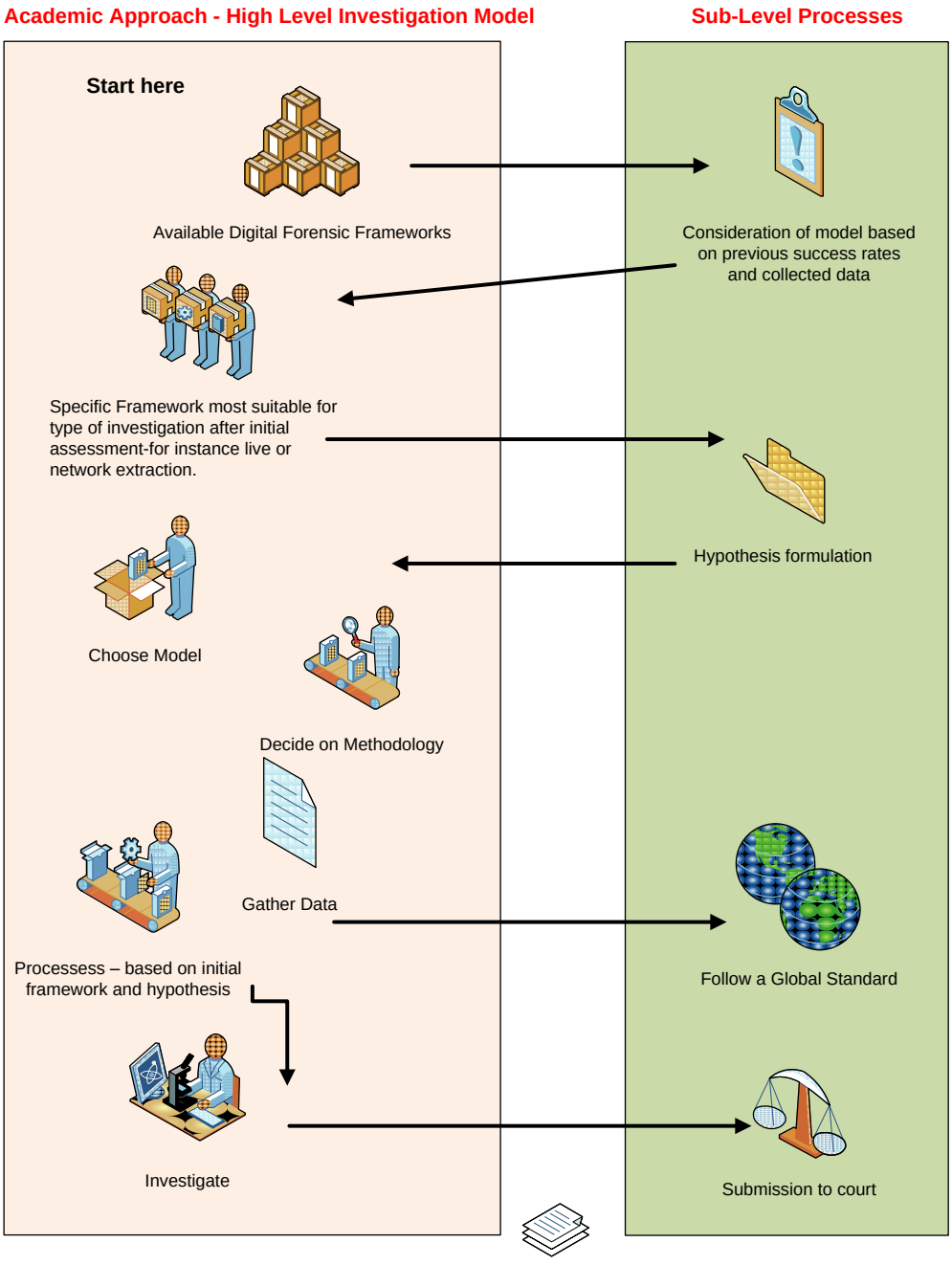
describing the phases as iterative rather than linear. This further suggests that reconstruction is only made after all investigations have taken place instead of having two reconstructions that might be inconsistent.

Another framework suggested by Barbara (2007) covers network forensic readiness. If all these frameworks could in fact integrate and build on a readiness cross-platform, we might have a much better prediction for the expected forensic investigation results. When comparing various frameworks, we note that recently researchers have moved away from the traditional framework as initially suggested (Palmer, 2001). This paper was the first to propose a framework for Digital Forensic Science. It suggested that the processes of Identification, Preservation, Collection, Examination, Analysis, Preservation and Decision should be followed.

Phases and sub-phases function in a sequential manner and have distinct steps. Principles cover procedures and guidelines as well as methodological approaches; this might overlap with some or all other phases and sub phases. Besides principles, goals and objectives are also defined to remind the investigator of early scoping of the investigation. Beebe refers to forensic investigation as grounded on phases, sub-phases, principles and objectives. Investigators express that it is easier to meet objectives-based steps than task-based steps because of the uniqueness of each situation and presentation of results after a digital crime scene is investigated. (Beebe, 2005)

In order to establish a clear investigation process, investigators should differentiate between an academic approach and a technical approach. From a technical perspective, I suggest a new classification of sub-level processes as “influencers” that guide academic approaches or “functioners”. Note from Figure 3, that Sub-Level Processes are the core supportive structure guiding the main academic investigation process. In other words, these (theoretical) processes should be considered that shows a success obtaining results or a likelihood to succeed from previous investigation models.

Influencers are setting a platform – with its own sub-categorisation models (not shown here), guiding and determining how “functioners” present data that has been built on previous assumptions and/or predictions. Figure 3 is a concise version of these activities.



- Documentation in all phases for peer review and random Testing.
- Before court submission consider: model/methodology/data capture/verification based on global standard/acceptance
- Expert witness reputation and references.

Figure 3 - Simplified Framework

3.1.1 Classification Problem

We see an immediate problem of broad classification just by naming these disciplines and clearly interrelationships between these disciplines should exist to strengthen the way forensic analysis is conducted on digital data. In other words, we have to define achievable objectives for each case as well as common grounds with other disciplines that are objectively orientated towards finding solutions. On the other hand, we also have to design secure systems that can gather data without other interferences. For example, when we put a security protocol in place, we would do this from a technical security perspective and would not design a system to track and trace evidence acquisition processes.

Beckett (2007) recalls why computer forensics changed to the level at which it is at present. Forensic science and in particular computer forensics changed due to the intensifying demand from law enforcement to verify if and how a digital crime has taken place. It also questions how and when these criminal activities have taken place which originated from a specific person. We note in this research that major streams of thought defines frameworks that cross-reference existing platforms, but no suggestion is made of alternative approaches that might deal with the changing forensic investigation environment.

We find another recommendation from Carrier (2004) and supported by Palmer (2001) indicating a simplified investigation framework. Both researchers refer to the basic format of the existing framework, although they do not elaborate on the specifics at any stage. If a new approach to an improved framework is required, we then would have to change the model to which processes and procedures apply and investigators then would carry out their investigations based on an inter-disciplined framework. Forensic researchers differ when selecting phases of the current digital forensic framework. This changes their perspectives on solution finding, instead of sticking to a basic platform in data and file acquisition processes. Further research is required to find a method streamlining the existing framework.

Most researchers start with a reference to the most basic investigation frameworks in use, as Perumal (2009) suggested. Perumal reflects on the wide range of choices, when these stages are compared and re-considered, one should be able to find a generic Model Name, presenting a selection of descriptors covering a particular model range.

Model name	Inventers	Year	Number of Stages
Computer Forensics Process	M.Pollit	1995	4
Generic Investigation Process	Palmer	2001	7
Abstract Model of the Digital Forensic Procedures	Reith, Carr & Gunsh	2002	9
An Integrated Digital Investigation Process	Carrier & Spafford	2003	17
End to End Digital Investigation	Stephenson	2003	9
Enhanced Model of Cyber Crime Investigation	Baryamureeba & Tushabe	2004	21
Extended Model of Cyber Crime Investigation	Ciardhuain	2004	13
Hierarchical Objective Based Framework	Beebe & Clark	2004	6
Event based Digital Forensics Investigation Framework	Carrier & Spafford	2004	16
Forensic Process	Kent K, Chevalier, Grance, & Dang	2006	4
Investigation Framework	Kohn, Eloff, & Olivier	2006	3
Computer Forensic Field Triage Process Model	Roger, Goldman, Mislam, Wedge & Detota	2006	4
Investigation Process Model	Freiling & Schwittay	2007	4

Table 1 - Complete Digital Forensic Investigation Model

When we compare old frameworks, we first note similarity of concepts. We would use the traditional and almost generic framework as described by Carrier (2004). (Beebe, 2005) maintained that the two most important phases in the investigation process are evidence preservation and documentation. This view differs from the findings of the research reported by Casey (2002) who addresses some other key processes. When we examine the minimum requirements of digital investigation processes we note that a few researchers cover the same two distinctive processes.

Documentation is an example of a principle; these principles represent goals and objectives sought through the processes themselves, thus producing a completed objective of a phase. Casey (2002) refers to six phases in the investigation process.

He refers to linear processes and that is based on particular phases that follow one after the other in a sequential fashion. We note that most models concentrate on processing digital evidence and they do not automatically include issues like the chain of custody, requirements and needs of different groups of users, as in Palmer (2001). Although this framework has some drawbacks it is more descriptive which compares it well to the one suggested by Warren (2001) where only four phases are included, namely: Assess Acquire, Analyse and Report.

We need to consider the possibility that investigators might try to create their own frameworks, depending on the specific environment. This phenomenon is discussed by Beebe (2009) who researched and proposed alternative models to investigate separate phases of examination and analysis as Palmer (2001) suggested. This shows that the examination phase is preceded by search and extraction activities. Even though the analysis phase is more focused on following activities that produce useful information from the extracted data, Beebe (2009) proposed a model to improve sub phases of data analysis. They support the idea that each sub-phase can be broken down into activities which aids iterative interaction between investigation phases. This enables iterative data transfer of information and leads to activities of Survey, Extract and Examine in each sub-phase. When we consider each of these sub-phases, they do not only complement each other when detail is lacking from previous inspections or investigations, but also aids to the layer investigation of each abstraction interpretation about activities required to the investigation.

Another reference to (Beebe, 2005) adds strong focus on transparency of phases by introducing a hierarchical structure by which sub-phases may be related to fewer higher level phases. She also addresses some new class principles, which gives a firm guide on evidence in preservation and documentation processes. One might therefore argue that any other framework would work as well or as poorly as the previous traditional models if we do not consider proposed alternative methods to conduct investigations with more clarity on different levels and therefore being more descriptive in all the phases.

All depends on the number of layer activities required for specific forensic issues that might arise. We can therefore only attempt to streamline the investigation processes. As soon as we start to confirm or require a bigger subset of sub phases we start to develop a new framework which is not necessarily consistent to the investigation process, but this might influence the abstraction layers of investigation processes for instance when other digital data is gathered and tested. This framework might then aid a new approach requiring useful information of a specific field of research such as a tool in creating objective outcomes in investigation processes. This might help setting a clearer understanding of objectives based task activities.

Foster (2004) maintains that the forensic disciplines lack sub-processes and that a more defined forensic framework should be put in place to address issues related to digital data extraction which belongs to a specific characteristic in the evidence collection process. Foster uses the term “process forensics” to describe the processes involved in evidence collection that are obtained by setting checkpoints in researching the activity while gathering of this information takes place. Stephenson (2003) supports this idea and discusses the importance of having a proper chain-of custody procedure while conducting an investigation. Process check pointing will aid these types of investigations. Chain-of-custody procedures are one of the cornerstones in court to prove connections of time and place of the digital forensic criminal activity.

We note that although a framework suggested by Barbara (2007) has been developed for the network environment, there is some similarity to normal computer framework analysis. This is because, as with any framework, some limitations exist, and a framework is a mere tool that aids the investigator in the investigation process. Investigators would call on a framework which describes certain processes to follow in the specific situation and guide the investigator based on the assumed forensic issue at hand. Allocated forensic procedures are then used to investigate the case. The investigator also seeks to investigate the real applicability of the framework to the case and draws conclusions that are guided by specific procedures suitable to solve the case.

Whilst using this new framework the investigator is incapable of drawing immediate conclusions; however, the framework ensures the investigator follows a pathway in procedural investigation principles and would compile a report reflecting on the success of this approach.

Forensic primitives as described by Carrier (2003) refer to the basic forensic investigation based on the core platform of the computer system that covers all aspects of root channelling and instantiate the dependencies to be followed in further investigations. It is therefore important to follow suggested paradigms for the particular issue at hand that might change according to the specifics of the investigation itself. Suitable investigation processes guides specific computer investigation processes. Therefore, the investigation of a computer which is running and the investigation of a computer which has been switched off will differ in procedures followed in the investigation process. Naturally, for each of these scenarios alternative frameworks might be considered in the investigation. When we consider a relatively recent study reported in by (Gladyshev, 2004) we see that this research has bound algorithm interpretation to the basics of digital forensic investigation. Gladyshev (2004) refer to the investigative process being divided in several stages which mirrors the initial study by Palmer(2001), Carrier (2004) and Kohn (2007).

The major stages of preservation, collection, examination and analysis play a role in all forensic investigations. As a guideline, a non-forensic person would have to consider how these stages — albeit very generic, fit into the framework of a basic investigation and how these stages are really nothing more than a very broad guideline to consider when they enter the digital forensic arena.

Comparing research reported in (Hall, 2005) and (Landman, 2002) we note that although similarity between the frameworks is evident, the investigation processes lack the detail needed to address a specific forensic criminal activity. The reason for this might be that a lot has been researched in the past in terms of the theoretical understanding of the “assumed framework” instead of making sure that an adaptive framework exists from which elements from other possible criminal scenarios has been eliminated.

This would ensure that a pool of possible criminal activities from these scenarios had been scaled down to a framework of most likely predicted outcomes.

3.1.2 Existing Frameworks Similarity

Comparing existing frameworks, in the first instance we note a similarity between of concepts. We use the "traditional forensic framework" – a generic framework as described by (Carrier, 2004). Note the similarity of the concepts in the Table 2, adapted from Aanya-Isijola (2009).

Kruse and Heiser (2001)	*DOJ*	*Lee*	*Casey (2002)*	*DFRW, Palmer (2001)*	*Reith(2002)*	*Ciardhuain (2004)*
Acquire evidence	**Collection**	**Recognition**	**Recognition**	**Identification**	**Identification**	**Awareness**
Authenticate Evidence	**Examination**	**Identification**	**Preservation**	**Preservation**	**Preparation**	**Authorisation**
Analyse Data	**Analysis**	**Individualisation**	**Classification**	**Collection**	**Approach**	**Planning**
	Report	**Reconstruction**	**Reconstruction**	**Examination**	**Strategy**	**Notification**
				Analysis	**Preservation**	**Search and Identify Evidence**
				Presentation	**Collection**	**Collection**
				Decision	**Examination**	**Transportation**
					Analysis	**Storage**
					Presentation	**Examination**
					Returning Evidence	**Hypothesis**
						Presentation
						Proof/Defense
						Dissemination

Table 2 - Framework Comparison

Although phases from different researchers' are not reflecting a particular order matching a series of phases from all researchers, we note that relative placement of these phases varies according to importance while conducting investigations.

For example, consider the occurrences of “analysis”, we note that it has been used only four times. Perhaps this also reflects the importance participants’ put on analysis in their investigations. Since predictions are difficult to formulate, we need to make assumptions to a higher level and gradually break this down to a tiered level at which the framework matches findings to a methodology in place.

Establishing a framework which adapts to the forensic scenario would provide the ultimate tool assisting forensic investigations. From this, a question-set guiding the investigation should allow detailed scenarios explanations of similar crime investigations. Although Armstrong (2003) did an in-depth research of the forensic model used at the time, he did not establish a clear framework while testing forensic models, he only rephrased finding by other authors, e.g. (Noblett, 2000), who focused more on the procedures and techniques under discussion with regard for the process in gathering the data. On the other hand, Broucek (2005) voiced some problems with the developing of taxonomy of forensics computing that might enable the development of a clearer structure that in turn would ensure some consistency in further research. We have not reached yet a point where we can promote an integration model that combines procedures and techniques without influencing potential case interpretation. Four major disciplines would benefit from a framework that refers to a standardised terminology namely computer science, law, information systems and social science.

As we found out from participants, most companies tend to develop their own procedures for examining gathered data. These basic processes of acquisition, identification, evaluation and admission as evidence form the foundation of the forensic framework. However, we need to analyse these procedures in depth to show a more defined structure. This is because researchers need to distinguish between investigations frameworks that define the physical investigation steps compared to a more abstract framework that defines abstract entities as described in (Carrier, 2003).

For example, “Abstraction layers occur in multiple levels. The file system itself is a layer of abstraction for the stream of bytes from the disk media.

Within the file system are extra layers of abstraction and the result is a smaller stream of bytes that represents a file, which is then applied to an application level of abstraction and it is processed further."(Carrier, 2003). Given the complexities of abstraction, there is room for error when the meaning of phase, activities, components, processes, stages, steps and classes are misinterpreted.

A more recent framework suggested by Freiling (2007) focuses on analysis as a means to improve the investigation. It involves pre-incident preparation, pre-analysis, analysis and post-analysis. As with most other framework, these are mere guidelines aiding the investigators in assessing data and compiling these into unique groupings. We state this because no study as yet has delivered a framework that prescribes digital gathering procedures for the different interdisciplinary procedures.

In other words, we propose a new framework that starts at a higher level covering complex forms originating from other disciplines but still consider additional similar references as the investigation processes unfolds. As soon as a higher level has been proved, we could look at the specifics, but still with good interrelation with other disciplines which might assist in the investigation process.

Carrier (2004) started by using some very basic ideas in the investigation process. He described how an event was triggered by a cause with finality in effect because of the cause which transgresses into an event. This further led to a change of the state of the object, which means the digital object's characteristics which may be traceable through investigations. Digital forensic investigations are considered to be on the same high-level as normal digital investigation but one distinct and important aspect is that digital forensic investigations must be presented in court with enough evidence to prove guilt. Each process in the investigation is subject to other phases, for instance in a crime scene preservation and documentation, evidence search and documentation and the event reconstruction and documentation.

Another study (Carrier, 2006) points to the difficulty of using only one forensic model in investigations, since no framework has sufficient levels catering for different case scenarios.

To address this problem, Carrier proposed a model based on the history of a computer to define categories and classes of analysis techniques. Subjective analysis of this framework would suggest a starting platform for classification of potential digital related crime personas. Clearly this model could be improved by grouping and classification of these individuals and further used in predictive analysis. When the emphasis is on the classification process and forensic corpora are freely distributable, such analysis would become possible.

Meyers (2005) suggests five almost similar steps conducting a computer forensic examination and suggests the order in which they should be conducted. Investigators should conduct documentation as a continuous process and note down all details from the first introduction of the investigation.

- Policy and Procedure Development;
- Evidence Assessment;
- Evidence Acquisition;
- Evidence Examination;
- Documenting and Reporting.

3.2 Digital Forensics Life Cycle

Digital forensic processes are inherently bound to a System Development Life Cycle (SDLC), which therefore shows alignment to phases in the investigation process. From responses by our participants we note that some participants are using their own investigation “lifecycle” which shows their broader understanding and expertise to solve cases.

For example, we could say technical “button investigators” (chapter 7 and 8) merely find answers to a problem, whereas “solution finders” are the investigators with an insight into broader aspect of interpretational issues, relating to other disciplines and confirming their findings with academic precision. Had a phased SDLC investigation being followed, (planning, analysis, design, implementation and maintenance) then matching appropriate phases of investigation to the digital forensics investigation process would have presented a clearer progressive phased process and present a clearer investigation model.

Another approach to positioning tasks and processes for investigators is built on a forensic SDLC model. Each activity or stage is an important process in the forensic investigation model. As the investigation progresses, investigators might encounter some cross-discipline references, e.g. from networking or encrypted system files. Influences from these disciplines might present a misguided interpretation. This is because no clear transgression exists between progresses of procedures and processes in cross platform disciplines. From this perspective, we present an awareness of a forensic model built upon a SDLC framework which indicate a progression of phases. This model transformed into a proven and reliable framework for investigation of forensic cases with the establishment of a forensic databank.

Although certain solutions or explanations might satisfy the goal of the investigation on a case by case basis, investigators should attempt finding solutions. Answers are presented from the context of the problem and a solution forms an in-depth analysis. This is because solutions present an unbound rule with options of comparisons that does not bind itself to a correct answer, but rather on interpretation. Therefore, a solution covers various methods and processes that show higher analytical perceptions in context of the problem.

It is important to keep a chain-of-custody in place – thereby ensuring that a forensic lifecycle reflects how the investigation was conducted. An integrated documentation plan that describes documentation processes as well as a description of the software and hardware used in the investigation must be provided too. Proposed changes to the proven methods followed before must be presented in the documentation, thus ensuring a particular guideline is followed in future cases. If this indicates that there are means for automation of processes that improve efficiency by freeing investigators 'time for valuable analytical and investigatory work, these processes must be documented as matching the SDLC model and changes should reflect this as well. From the altered SDLC processes we may establish a process flow that allows for improvement of the way case investigations are conducted and a better understanding of future assessments of similar type.

An expected prototype model reflecting a faster approach to possibilities might increase effective communication and decrease development time. These benefits are further enhanced by any decrease of costly mistakes. As the investigation progresses from broad expectation to possible findings, we suggest a pathway or a platform built on requested investigation requirements. It is common for new ideas to develop from a rough estimate after conducting initial prototype modelling. Applying this approach in forensics would force the investigator to find out at least what type of case and environment is required thereby planning specific processes for the investigation. We see this at an early stage when redundant features or data are eliminated in contrast to feature-rich data characteristics. Kissel (2008) mentioned that a typical SDLC includes five phases: initiation, development/acquisition, implementation/assessment, operations/maintenance, and disposal. Investigators have to incorporate sufficient processes that allow repeating a life cycle without destroying the case or drastically having to change their hypothesis.

When we put this into practice, we again note that models and frameworks can be a burden if excessively used and ultimately scaling down the investigation process to an achievable click-box environment. Using a SDLC model in digital forensic investigations should rather act as guide for trainee investigators establishing a sound basis for investigation. This would ensure these investigators' understand the basic steps before deeper analysis is conducted.

From our participants' responses we note that companies should create and maintain procedures and guidelines for performing forensic tasks, based on the organisation's policies and all applicable laws and regulations. It also became clear that only a few companies briefly mentioned this in their investigations. A good starting point might be to put I place procedures and guidelines focusing on general methodologies for investigating incidents using forensic techniques, since it is not feasible to develop comprehensive procedures tailored to every possible situation. Consideration should be given to developing step-by-step procedures for performing routine tasks. The procedures and guidelines should facilitate consistent, effective, and accurate actions. This is especially important for handling incidents that may lead to prosecution.

When decision makers handle forensic data evidence in a sound, thorough manner, they are in a position to take necessary follow-up actions with confidence.

The importance of investigation varies considerably as participants' responses show in Figure 4. From participants' feedback, we established that the lowest ranking in this graph indicates the order of importance or actions taken by participants during investigation. In other words, most investigators started with collection as the first action and only at a later stage did they consider duplication, hypothesis and prediction. Figure 4, also supports the idea that investigators do not consider prediction as an option at this stage. Since reasonable prediction is only possible after data analysis from a data bank, we might not see this happening soon since the digital forensic discipline does not yet have corpora to test against. This also affects automated prediction as a vast number of data is required to analyse pattern occurrences.

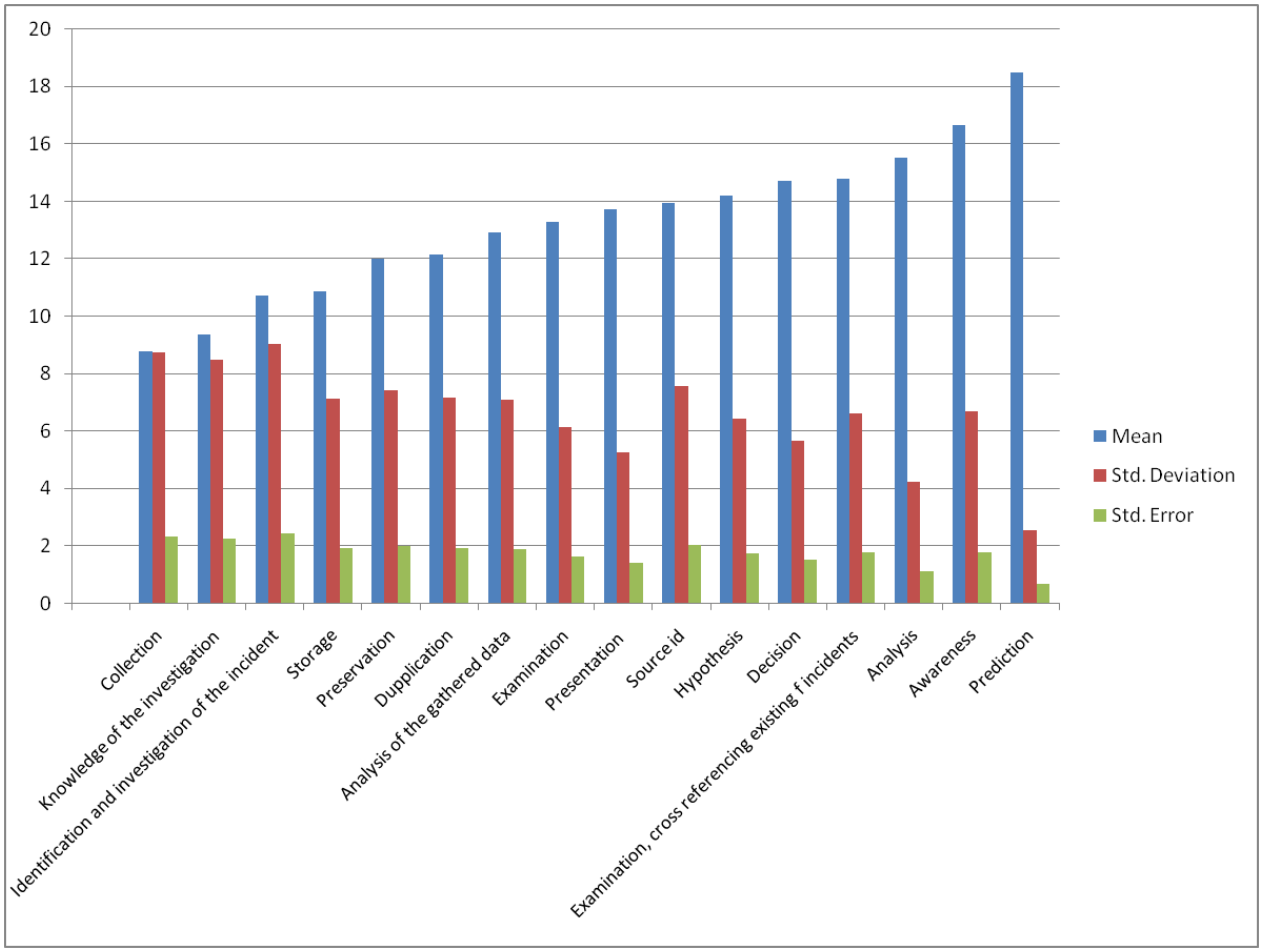


Figure 4 - Investigation Process
A low score is more important or occurs first in the investigation process

Radack (2009) indicates that a basic forensic investigation process consists of collection, examination, analysis, and reporting. These processes are better understood from in a wider procedural basis of:

- Preparation: case briefings, engagement terms, interrogatories, spoliation, prevention, disclosure and discovery planning, discovery requests;
- Record: Drive imaging, indexing, profiling, search plans, cost estimates, risk analysis;
- Investigate: Triage images, data recovery, keyword searches, and hidden data review, communicate, iterate;
- Report: Oral vs. written, relevant document production, search statistic reports, chain of custody reporting, case log reporting;
- Testify: Testimony preparation, presentation preparation, testimony.

Challenging positioning of the investigation processes come to light as soon as different frameworks are compared. This is because of the interrelation between different processes from each framework which in turn describe different processes in the investigation. It therefore becomes increasingly difficult to establish a platform in the second and third tier of the analysis that cannot become grounded in existing models. This might lead to intuitive solutions and guesswork from other disciplines which relate to similar processes in forensic investigations. Investigators might accordingly experience predictions of closely related and misleading guidelines while conducting investigations.

We also note from participants' responses that different roles in an organisation are meant to achieve optimum results; however, not every participant works on every activity within a phase. This is in line with research conducted by Kissel (2008) who states that the determination of which participants need to be consulted in each phase is as unique to the organisation as the development.

With any development project, it is important to involve appropriate information security personnel as early as possible, preferably in the initiation phase. In some organisations, a single individual may hold multiple roles.

3.3 Processes in Existing Frameworks

We note that several frameworks have emerged. Carrier (2004) focuses more on an event-based digital framework whereas Beebe (2005) proposes a framework that is objective-based. Both these frameworks essentially still refer to the traditional model initially proposed in (Palmer, 2001). In contrast, we note that Baryamureeba (2004) suggests another approach that separates the investigations at the primary and the secondary crime scene while depicting the phases as iterative instead of linear. It further suggests that reconstruction is only made after all investigations have been taken place instead of having two reconstructions that might be inconsistent. Another framework suggested by Barbara (2007) covers network forensic readiness. If all these frameworks could in fact be integrated and built on a readiness cross-platform, we might have a much clearer idea about expected forensic investigation results.

When comparing the various frameworks suggested for digital forensic investigation, we note that more recently researchers have moved away from the traditional framework as initially suggested (Palmer, 2001). This report was a leading paper in establishing a framework for Digital Forensic Science. It suggested that the processes of Identification, Preservation, Collection, Examination, Analysis, Preservation and Decision should be followed.

Chapter 4 - Framework in Action

The digital crime scene may consist of a number of computing and storage devices, as well as the network connecting them. We specifically consider the digital crime scene that consists of a number of computer systems or standalone home based personal computers. Digital evidence is any digital data that contains reliable information that supports or refutes a hypothesis indicating the validation of digital evidence. Digital evidence may be found on the hard drives or in the volatile memory of all the involved hosts, as well as in captured network traffic, referred to as network dumps.

A central issue in evidence dynamics is to identify the causes and effects of events. The evidence dynamics of different digital media varies. A file can be modified or deleted, and timestamps can be updated. Unallocated data on a hard drive can be overwritten, and volatile memory can be overwritten or moved to page files. Data transmitted on a network may leave traces in log files and monitoring systems.

Our approach to event construction and testing starts with a hypothesis, based on the present way digital data is accessed. We need to proof that these assumptions in the hypothesis are correct and fit for court proceedings. We will take real-life case studies and test the way data is verified. While conducting these tests, we might perhaps find a particular analysis is not suitable based on the reconstructed case and determine whether other methods are possible.

A chain of event corresponding with processes gathering data is then replayed on the test bed. The virtual environment is analysed to find the effects of the events. These effects are in turn compared to the actual digital evidence. The purpose is to replay the suspected attacks in a controlled environment in order to study the causes and effects of the events involved in the attack. This allows us to replay the attack in a forensically sound manner without compromising the integrity of the original evidence or relying on files that have been compromised by the attacker.

As noted above, a multi-step attack can be studied as a series of interconnected events, where the effects of an event are the causes of the subsequent event. Although the digital forensic reconstruction framework separates causes and effects, differentiating between these may be difficult in practice, as it may require exhaustive testing. Using the present framework to determine how forensic processes should work, we can analyse how the processes can be hampered if one or another part of the framework is missing. In some cases, there may be several theories about the chain of events leading to the digital evidence found in a digital crime scene. In this case, each hypothesis is formulated and tested separately.

Although research by Elsaesser (2001), showed a proof-of-concept implementation of a system that might assist researchers with pattern matching and knowledge engineering, their findings did not present prove of a range of attacks without sufficient modelling of templates. In (Nolan, 2005) we note an advanced technical approach that discusses procedures and process characterisation and uses software tools in forensic investigation. The researchers also cover an automated process that prompts that the same processes could be used to automate script and batch file calls enabling first hand extraction of forensic data from the compromised unit. This would enable a faster extraction of data, further minimising potential corruption

Leong (2006) proposed Forensics Zachman framework (FORZA), which forms the foundation of that research and also proposed some legal requirements. It also suggested a multilayer approach in solving the different activities associated with each layer. These layers are; Contextual investigation layer, Contextual layer, Legal advisory layer, Conceptual security layer, Technical presentation layer, Data acquisition layer, Data analysis layer and Legal presentation layer.

Leong also suggest a forensic plan conducting investigation under these layers using methods defined under:

- What (the data attributes) defines a plan to conduct the investigation, for instance strategy objectives.
- Why (the motivation) hypothesis of the incident occurrence.
- How (the procedures) forensics strategy outline the data acquisition and analysis procedures and requirement.
- Who (the people) relationships with external third party or file originator.
- Where (the location) source of incident, for instance IP address.
- When (the time) hypothetical forensics, event timeline, confirmation of time line.

There had been earlier suggestions for a multi-level approach in automation. Armstrong (2003) proposed a method in for the development of a framework for evaluation of the appropriateness of computer forensic tools.

This research suggests the following steps of evaluation:

- Configure a test bed with appropriate software according to a hypothesis;
- Replay attack according to the hypothesis and save snapshots for each state;
- Acquire and verify images of all snapshots;
- Perform analysis through the comparison of states;
- Compare images to digital evidence to support or refute the hypothesis;
- The process can be repeated to find alternatives to the initial hypothesises;
- Experimental Measurements;
- Anti-forensic measures by users;
- Define and improve methods used.

In order to obtain research information from forensic experts, the following steps would enhance data gathering:

- Interview forensic data experts who are experienced in delivering court proceedings; Question and investigate their processes and software they use;
- Test the international interpretation of the existing frameworks in question;
- Verify if the test results from the hard drive matched similar occurrences from a control group or administrator who monitored all changes to the hard drive;
- Obtain the required results;
- Present the data;
- Diagnose your system after results were obtained without jeopardising the kernel processes.

Digital Data Capture (DDC) has been evolving over the last years with a considerable progress towards a methodologically proven framework. However, it has not reached a satisfactory and consistent performance consistent with an error-free model.

Had (DDC) developed to an adaptable model that evolves with variations to the basic investigation platform, one would expect steps improving the procedures, since DDC is performed as a core procedure in digital data forensic investigations. This research considers how any part of the existing framework complements a few mainstream findings by other researchers in establishing a framework that guides towards an automated process on case by case basis. Technical aspects of forensic data will always be present because data has to pass through a hardware component at least once in its life cycle. This covers time stamp of MAC properties of the data such as time of creation, deletion and alteration in any form. Data leaves a digital footprint that can be traced back to the originator.

Digital criminals try to cover their tracks and it is up to the forensics experts to find, validate, store and present the data in such a way that a connection is established between the originator and the data.

Forensic science is unavailable to the average computer user and only a handful of well trained forensic experts are able to present findings fit for court procedures. This situation has led to a selected group of investigators being able to use these software packages and thus only conducting inspection based on proprietary packages. An advantage perhaps is that the investigators get a greater success ratio in resolving cases. There is a shortage of real-time forensic investigators, we would prefer to call them forensic scientist and not just "button-experts" which indicates the average investigator's ability to do only "drag-and-click"

4.1 Data Storage

Computer storage capacity has been steadily increasing. This phenomenon puts strain on resources available to forensic investigators. This presents investigators with less time available and strains their personal skilled approach in finding data. Ultimately these investigators start using a "button-forensic" approach. Kovar, (2009).

From a data storage perspective, digital data is stored on hard drives with ever increasing capacity and finding forensic data is likely to become more difficult in future. We see the expansion of hard disk capacity as a real threat to the securing of a standardised platform, since new processes must be developed to maintain integrity of the large hard disks. For example: according to The International Data Corporation (IDC) media available to store the newly created and replicated bits and bytes of the digital universe will grow 35% a year from 2006 to 2010, or from 185 exabytes to 601 exabytes. (Gantz, 2007).

Such a vast amount of data is possible if we consider that the average home personal computer has a hard drive size of not less than 750 GB. (Villars, 2009). Investigators will have their hands full in future, when we expect the growth in the world's digital data to expand. Hard drive manufactures are constantly pushing the limits of their designs. Seagate is about to release 1.5-Terabyte Desktop PC. This storage capacity opens a massive gap for digital data validation, in other words for data forensics (Seagate, 2008)

4.2 Generic Investigation Guideline

As quite a few software packages cater for specific type of investigations, we cannot cover all here. As an example we present a generic Encase investigation guideline. Adapted from (Kovar, 2009)

From a very general perspective we would start an investigation as follows:

- Turn the computer off. This is arguably the most important issue and varied responses are returned from participants when asked whether to keep the system alive or turn it off immediately. Investigators should have sufficient proven experience to determine this.
- Photograph the exterior and attached devices.
- Inspect and document the exterior.
- Inspect and document the interior.
- Create case: Ensure that you have all relevant information - custodians, clients, case name, etc.
- Add evidence - E01, LEFs, loose files, etc.
- Document the details of the hard drive(s).
- Confirm disk geometry, sector count, partitions.
- Connect the drive(s) to a write-blocking device, if available.
- Run Partition Finder if indicated.
- Run Recover Deleted Folders.
- Search case - hash and signature analysis.
- Run File Mounter - recursive, not persistent, create LEF, add LEF to case
- Copy the data from the hard drive(s) to the newly formatted hard drive. Alternatively create an evidence file with the forensic software Run Case Processor- File Finder. Export results, add back in as LEF.
- After the copy is complete, verify the integrity of the copied data. This is done by taking before and after MD5 hash values. Forensic software does this automatically.
- Search case - hash and signature analysis.
- Search for encrypted or protected files. Address as appropriate.
- Extract registry hives
- Index case.

Chapter 5 – Standardising Digital Forensics, is it Possible?

It appears that data collected by participants in this thesis did not use a preferred order in their investigation processes. This opens the issue of standardisation since all tests should be verifiable by the forensic discipline as valid and conducted following standardised processes.

The competence of investigators seems to be questionable as a study reported in (Hannan, 2003) found. Accordingly, on the job training was valued higher than tertiary education in forensics. The study also covers competence requirements for the investigators and reflects frequency computer science is higher than investigations skills. In contrast, under mean of importance, investigations skills are higher than computer science. An interesting finding from this study indicated that from a sub-competence level, knowledge of the operating systems and application software was regarded higher than investigation skills. Lack in trained personnel has been an issue for some time, as we see from the responses by the participants. Although funding availability for training plays a role, a recent article in Forensic Focus, (Kovar, 2009) underlines the importance of a more scientific approach, although in most cases examiners tend to fall into the “button forensics” group.

Stallard (2003) also suggests that a limiting factor in developing an evidence process (also considered to be a framework *per se)* is the limited number of qualified technicians. This statement is contrary to statements from participants surveyed for this thesis, who stated that there is a distinct rivalry between technicians and forensic scientists. While this is a valid argument, and groups the forensic scientists together, we still need to develop more proven methods for the investigation process, since a lot has been spent on methods for gathering evidence and not as much on the investigation process. This situation had been addressed by the DFRWS as far back as 2001 as stated in (Stallard 2003).

At another meeting of the Digital Forensic Research Workshop (DFRWS) in 2005, it was mentioned that forensic examiners often make statements that are never verified for their veracity by judges or counsel.

The lack of progress in this body of knowledge is of concern as it has not been developed due to a number of reasons: lack of experienced individuals in the digital forensics community, lack of collaboration amongst forensics professionals, and lack of high standard peer-reviewed journals. Furthermore, these reasons can be traced back to the short supply of qualified professionals in the field because of the shortage of educational curricula and certifications. Technological progress is only exacerbating the situation, since the knowledge base must be constantly updated to prevent stale and outdated information. With proper standardisation these issues would most likely be solved and accredited tools and methods could be used for validation and process confirmation.

(Beebe, 2009) states that digital forensics is lacking a common body of knowledge. When we consider that the digital forensics industry is ever changing to new technology, it is easy to see that an existing body of knowledge is ever evolving. This might lead to miss-interpretations or confusion of complex forensic concepts and ultimately leads to difficulty in obtaining accredited results. It is clear that time and resources for solving cases would be constrained.

Beebe also mentions that the industry is expecting more tools to be developed for forensic work. We should instead rather consider which comes first – the tools or the discipline, in order to maintain a scientific forensics rigour and stability. To the contrary, it is clear from published papers, and duplicative research from researchers that try to reinvent the wheel. Beebe also suggests that a clear research agenda should be specifying the development of a new body of knowledge in contrast to the industry's expectations. This would counteract stepping through tools without having full understanding of the reasoning and accumulative assessment of well-established findings used for skill and knowledge generation. The forensics discipline should take care not to neglect other non-digital aspects such as environment, legal and accreditation issues that indirectly influence the building blocks of a digital investigation.

The author of this thesis observed that training in software tools as well as scientific analysis of digital data is required to maintain growth in this discipline.

5.1 International Standardisation of Computer Forensics

Although results from participants were collected globally, we did not encounter specific reference to difficulties to conduct investigations. However, if we would like to standardise cross-border digital forensics, through a classification model, we might encounter difficulty in doing so. This is because digital crime in one country might not be regarded as a crime in another.

Investigations that leap from server to server, from country to country, crossing many borders on the way, are complicated not only by differences in handling evidence, but also by political and legal differences. Privacy laws regarding digital data transfers seem vary between countries as well. This is why efforts are being made to bring some standardisation to procedures regarding digital evidence. The G8 group has recommended six principles for digital evidence gathering:

- All standard forensic and procedural principles must be applied.
- Upon seizing digital evidence, actions taken should not alter the evidence.
- People accessing the original digital evidence should be trained to do so.
- All activities relating to the seizure, access, storage, or transfer of digital evidence must be completely documented.
- Individuals are responsible for all actions taken while the digital evidence is in their possession.
- Any agency that is responsible for seizing, accessing, storing, or transferring digital evidence is responsible for complying with these principles. (Wilson, 2009)

Although these principles are wide and seem almost generic, forensic academics should take a lead in setting standardisation protocols in conjunction with the industry – thereby touching on all sub-level procedures and processes of investigation. Future research should find a method to streamline the existing frameworks. Most researchers start with a generic reference to the most often used framework as suggested by Carrier (2004) and supported by Palmer (2001).

Both refer to the basic format of the existing framework although they do not elaborate on the specifics of each stage. If we consider that a new approach to an improved framework would modify the model to which a process and procedures might apply, investigators might apply a complete package of investigation based on an interrelation framework.

Forensic researchers differ to argue in their views on the selection of phases in the current digital forensics framework and how new methods in these phases are followed to find solutions. Instead keeping to the basic platform in data and file acquisition processes, they follow untested methods. A new framework should start at a high level covering complex forms related to other disciplines and still maintain touch points as the investigation processes unfolds. From this established high level, we could look at specifics, but still maintain the interrelationships with other disciplines which might assist in the investigation process.

(Ciardhuain, 2004) proposes a comprehensive framework where each stage describes the processes involved better than the previous researcher. However, a better understanding of the processes at earlier stages of creating a hypothesis might assist investigators and enable them to assume expected investigation procedures, before procedures of search and evidence identification take place. Recent research reported in (Selamat, 2008) covers 13 forensic investigation frameworks. Distinguishing between these frameworks shows how several phases cover specific output as a result of their activity or processes. Each researcher used their own interpretation of the framework and did not agree to use a standardised forensic platform to conduct their investigations. An observation of this is that the syntax describing each framework differs; it is therefore easy to see how standardisation becomes almost impossible to achieve.

It is questionable if digital forensics standardisation would contribute to the improvement of the investigation process since most investigators follow their own methods to obtain results. Standardisation appears to be only addressing academic concerns and would likely not set a standardised testing platform for the industry.

Chapter 6 - Preservation and Presentation of forensic data

Since one of the phases in digital forensics investigation is data preservation, it is of utmost importance to follow specific guidelines to ensure data consistency for further analysis. It is important to distinguish between forensics terminology used in layman's terms for defining computer forensics and a forensic framework. The term "Introduction to Computer Forensics" is often used to describe how a hard disk is functions and has nothing to do with regard to the forensic process at all. We see this inconsistency of meaning in the research reported in (Hall, 2005) where software and computer forensics definitions are compared.

Digital data presented in any digital format as evidence in court presents an investigation process by itself. The meaning of the investigation processes enables opposing court parties to understand how these processes are interlinked, allowing access to more documentation, e.g. technical reports. This would also allow them access to data results, even if specific software is required.

(Armstrong, 2003) indicates that forensic investigations are perhaps widely defined as preservation, identification, extraction, documentation and interpretation of computer data. This description does not cover the exact expectation that digital data should be presented as evidence nor does it explain how the digital data should be preserved for future reference, or how stored data could be accessed.

Most researchers suggest using the term "digital data" in their assumptions. We note from the research findings that only a few participants mentioned using a theory as a foundation for conducting their investigations. Participants refer to the importance of keeping track of data originators and changes initiated by the data originator. This process is in essence what digital forensics is about, i.e. to track the data originator and match changes to the date and properties describing time.

Research by Hall et al. quotes other researchers, such as ADDIN (Yasrnsac, 2003), who refer to digital forensics as having many synonyms and different meanings in the context of the investigation. This term is sometimes wrongly referred to as computer forensics media analysis.

6.1 Legal

Since court presence reflects an investigator's knowledge, they should know their discipline and we assume all presentations would follow a specific framework. To the contrary, legal presentations by investigators tend to differ from country to country, allowing international digital criminals a loophole for conducting their activities.

Data forensics becomes a science in its own right when evidence of a scenario is presented in court. Investigations should follow a mythology to obtain results and a clear scientific process ensures proven and standardised methods are followed. Testing theories would present success ratios and failures thereby validating undisputed evidence. Validation and testing should be conducted based on previous reviews and acceptance from the digital forensic scientific community. Digital data forensics make use of specific methods to obtain data and apply processes that have to be tested to ensure evidence is beyond reasonable doubt and is uniquely associated with an alleged criminal act. Only then can the court succeed in proving the guilt of a perpetrator.

Evidence about the connection between the alleged criminal activity and the alleged perpetrator has to be presented about how the alleged criminal activity is linked. International crime prevention has to meet the challenges of finding an acceptable across the border standard for applying methods and procedures to forensically test criminal activity. Comparing the frameworks and suggested processes must be carried out irrespective of the software used. For instance, is the software used dependent on the suggested framework or would any framework give the same results if the same software is used for different scenarios? We suggested further research in this field in order to improve the understanding of the processes and requirements and to streamline the investigation over a shorter time span and also to deliver consistent data results.

This issue had been addressed by Meyers (2004) where concerns were expressed about the qualifications of expert witnesses.

Since computer forensics has no defined credentials or a formal educational process in place, except some training courses, we note that a few lower courts accept qualifications based on the skills and previous work experience of the experts. This might have been sufficient at the time of Meyers' writing, but it is expected that contesting the expertise and qualifications of expert witnesses will become more common in the future. Therefore, there is a need to standardise the industry by introducing a national and internationally recognised certification and standardisation for computer forensics.

6.2 Presenting digital evidence to court

Presenting valid evidence in court must be supported by sound investigation processes. Unfortunately valid evidence can be rejected if correct processes were not followed.

According to (Ami-Narh, 2008) problems exist when determining if evidence is admissible or corrupted. In their study they reflect on the existence of unreasonable search and seizure protocols during the investigation process. Therefore, it is difficult to determine if authorisation and approval of the investigation must be decided before it begins. This issue opens concerns that investigations are conducted without using a standardised method that defines exact processes.

The importance of the investigation process is built on methods used to gather digital evidence. This also depends on the procedures followed. It appears that technicians are unaware of the impact complete and verified test processes have on clearing the field for legal investigations. A *"lazes faire"* approach seems to be a practice as most technical processes applied to digital evidence do not have to pass any formal test for it to be placed before a court.

6.3 Adding More Convincing Reasoning

According to (Ryan, 2005) it is not enough to simply produce an unbiased and technically accurate document describing the outcome of a forensic examination. The primary purpose of the statement is to assist the court in evaluating the admissibility and weight of any evidence found on the digital devices examined for

the case. This statement confirms that an understanding of the examiner's findings is required in order to decide the strategy and the legal points to prove.

The law that stipulates how to prove the existence of criminal activity varies from country to country. The UK law stipulates that each offence needs to have what are known as 'points to prove'. For example, under Section 3 of the Computer Misuse Act 1990, a person is guilty of unauthorised change of computer material if it can be proven that he or she: (a) does any act which causes an unauthorised change of the contents of any computer; and (b) when the act was performed he or she had the necessary intent and the essential knowledge to do so. These two points explain what in legal terms is called the *actus reus* (guilty act) and the *mens rea* (intent/knowledge) of the individual. (Ryan, 2005)

These requirements point out that we need to consider how the following phases are linked together, covering at least:

- Procedures and principles;
- Technical preparations – before submission;
- Guidelines;
- Precautions.

Clearly digital forensics evidence for admission in court should satisfy at least two conditions: it must be relevant, and it must be "derived by the scientific method" and "supported by correct confirmation." (Ryan, 2005)

When a novice legal forensic expert is to give evidence in court, their knowledge are questioned and a basic forensic framework should be part of their preparation. Their knowledge of the basic file structures and how these differ from active computers where time sampling occurs is of utmost importance. For instance Linux operating systems and Windows operating systems have different interpretations of the file directory structure.

Procedures must follow a set protocol defining the chain of custody. This show the evidence most likely would fit according to prescribed procedures, which ensure data integrity.

Forensic science and evidence is about the processes to apply an application of scientific procedures or techniques in illegal investigations. An example was set in the *Daubert vs Merrill Dow Pharmaceuticals Inc* case (1993).

This case mentioned several non-mandatory and non-exclusive criteria for discovering scientific truth. Digital evidence is mostly about stored data that has being transferred or changed through variations in digital form but still reflects to the file originator existence. When we use the *Daubert* case as an example for testing the validity of a framework we start to ask questions like:

- Is the framework used testable, can it set a foundation for recurring investigations.
- If this framework is testable, would other investigators use the same methodology employed to build this case and if this investigation is setting the ground rules, would investigators suggest the same processes in similar situations?
- Is it possible to base alternative investigations on the initial framework? If this is not the case, we need to offer alternatives within the second tier of possible solutions and framework functionality. What is the margin of error?
- Is the used methodology accepted in the forensic science community?
- Is the expert's testimony based on the expert witness' skill?
- How applicable is the average framework to all forensic case investigations, without losing its core focus?

Building more questions around these thoughts, we start to distinguish between categorisation and groupings in the framework. In today's changing digital forensics environment, it is important to have an adaptable framework that would act as a platform allowing automation of investigation processes. It is important to allow subtask scaling of processes when a testable automated framework exists, thus adding more features. It is therefore essential to set up a framework which presents digital data in a format that ensures the consistency of the data gathered data. This would ensure those at least minimum specification guidelines are met.

Conducting tests have to be interpreted properly, because it sets a reference point when compared with other areas of forensic identification science. Therefore, digital forensic investigations should cover at least all aspects of file creation, file encryption and file concealment. We note that log file analysis, file system analysis and file attributes all play major roles in interpreting how the files had been tampered with or did in fact change. (Gladyshev, 2004)

If we base this on a digital forensic framework and model how investigators interpret a case based on their initial hypothesis and own experience, it is clear that "one works for all" modelling would not work for digital forensics investigation. This is because each situation differs from the next and researchers have not yet developed a framework that caters for some interchangeable sectors of the framework. We suggest a template design based on a combined first tier that develops into sub tiers up to five levels.

This approach does not seem to be possible anymore. As the research progresses, we note that an improved investigation model would perform better through indexing. This allows improved template design that forces investigations into a specific channel. We need to control too strict template designs, as this might prescribe pre-determined guidelines, which might force the investigation into a one-channel approach trying to find a solution. I believe that given the separate building blocks of the suggested framework most interactivity would be possible in the proposed framework.

The framework should also be dynamic enough to allow for chunks of sector analysis to interact with other aspects of the research. Add-ons to the main framework still maintain their own contribution as enhancements to the main framework which needs to adapt and make changes to the outcome of the investigation. The key therefore is in the design of a verifiable framework which is comprehensive but also adaptable to cover and allow changes within the major chunks of the mainstream investigation. This would allow enhancements to existing investigation processes.

Because the complexity of computer forensics makes it hard for interpretation, it is most important that investigators should at least understand that within their own interpretation of the discipline, various other sub-processes exist which are not always easy to interpret based on the frameworks used at present.

In order to succeed in a legal case, presenting digital data evidence has to be based and proven the following; the facts in issue, the facts on which the disputing parties disagree, the circumstantial facts whose existence can be used to prove or disprove, the facts that must be proven in order for appropriate law to be applied or evidence to be admitted into court proceedings. In the digital environment evidence has to adhere to the basic characteristics of real data and that the evidence must be admissible. It also depends on the type of dispute and how the evidence relates to the facts being proven.

Evidence also has to pass evidential integrity testing. Evidence has to be proven as probable with regard to the source of the evidence. All digital data is created by someone and should be linked back to where it started. A piece of evidence that is possible to originate from tampering has no weight in proving the fact. If evidence has been tampered with it would not be admissible in court. Although digital evidence is as yet not clearly grouped in different classes we note that evidence can be grouped based on the probability of possession of the real evidence.

Real evidence can be the data before or after it has been tampered with. It also can refer to e-mail messages or other ways of documentation that can have imbedded text or images in the content, which can also be verified as being from the originator. Circumstantial evidence also refers to log files as a source of system information to reconstruct a sequence of events or file handling processes., While the investigation is conducted all documentation would be regarded as evidence. All these situations add to a problematic interpretation of the digital data. For instance there might be problems with the verification of anonymity of digital information.

Chapter 7 - Recommendations

7.1.1 Challenges for the creation of a Forensic Corpus

Recent research (Kahvedzic, 2009) proposed model to describe an investigation at different levels of detail. This suggests that independent vocabulary can be used to describe the investigating process in more detail. In similar manner, we could use this notion to present a data structure that defines specific groupings of similar concepts and their attributes, thereby ensuring representation of variables in a relational data structure. This database should be scalable ensuring new entities are related to the existing structure. Global forensic researchers would input data accordingly into this database. We envisage a noticeable benefit to the forensic community if members contribute and share their resources. This would provide realistic data sets that would assist in establishing a platform in digital forensics whereby an automated framework might evolve.

Garfinkel (2007) outlines the difficulties for research caused by the little interest in the creation of digital forensic corpora. One reason appears to be concerns about privacy. Because of privacy concern and the lack of a sponsor willing to address them upfront, many institutions placed roadblocks in the way of would-be researchers. Fundamental questions surrounding evidence in the digital world begin with identity, providing some digital link between binary data we collect and analyse, and the human being we call a suspect

Two papers presented at the 2006 Digital Forensics Research Workshop approached this issue from opposite perspectives. Garfinkel (2007) referred to the difficulty obtaining enough data from existing corpora. This is because forensic research might favour a certain type of investigation or they might place their preferences on the framework suggested as the ultimate model. According to his research most scientists are divided according to the kind of data being analysed, rather than the kind of analysis to be performed.

This finding highlights the problems in creating algorithms of prediction because of data insufficiency. Garfinkel et al. also had problems obtaining large scale corpora; he used 750 disk images which he bought on the second hand market.

Garfinkel's paper "Forensic Feature Extraction and Cross-Drive Analysis" presented a new techniques or automatically determining the owners of hard drives and for finding hard drives that are used by various social networks. What made this work possible was the possession of a corpus of 750 drive images that Garfinkel had purchased on the second hand market. Even so, that entire corpus is tiny compared to the number of hard drives seized on a regular basis by US intelligence operations. Currently there is no way to know if these techniques will work on a large scale corpus because such a corpus of hard drive images is unavailable. Garfinkel(2009) reports the latest developments in the effort to bring science to digital forensics, with reflection on the need for establishing a standardised forensic corpus. It is clear that once a data bank is created, classification could be started and data analysed. We support this point of view, as did some of the participants in this research.

When we consider that Gantz (2007) predicted an increase of the digital universe of 35 % a year from 2008 to 2010, or from 185 exabytes to 501 exabytes, questions arise about how all this digital data can be effectively used, when we need to make correct judgments after detailed investigations. Predictions would become more difficult to manage while a proper automated system is not in place. It is clear that small corpora when compared to the massive hard disk usage of the real world would not perform to expectations and would not be able to support well data predictions.

Although Kornblum (2005) showed promising result when applying a rolling hash algorithm for the prediction of forensic similarity, he encountered problems using the small corpora available to make predictions. From his research he showed how false positives or negatives could be used as a rolling hash algorithm to develop antispam which could be adapted to forensic purposes. A typical application include finding and matching altered documents, and determining if similar fragments of a document is present on a suspect's hard drive.

Kornblum was not able to report on false positives or negatives when the algorithm was run against a standard corpus of Microsoft Word files, because no such corpus exists. Digital Forensic investigations are in dire need for a reliable data set corpora. This data set should be globalised enabling participation and population of the database by members of the forensic community. Most investigation results are obtained as a singular event, that is, researchers are using specific investigation techniques suitable for a unique case and cannot reproduce the same investigation under different circumstances. This indicates the unlikely reproducibility of exact data for validation by other researchers' results. Scientifically, these results are hard to reproduce.

Forensic investigation software tends to gather only results by default and has few customisation options. We note from Garfinkel (2009) that minimal options exist to compare results after investigations took place. Comparing results from a forensic investigation, gathering investigated data from one case and thereafter attempts to validate these results testing if the same results are repeatable, find it difficult to produce the same results.

7.1.2 Software preferences

In order to create a data bank, investigators mainly use existing software. Arguably we could challenge if the software results, presents the best extracted results, delivered in the shortest possible time. Figure 5 presents participants' preference to five major software packages. Other open source packages are also presented. The figure supports our suggestion that investigators are technically orientated and prefer "click and drag" or "button-investigating" processes. Although open source packages have more functionality, they require investigators to use scripting instead of a visual interface to extract data. This is another major question mark in establishing a standardised investigation platform that is testable for consistency by peers. This is reflected in Figure 5, indicating that the use of user interface software packages is a top choice over script or command prompt investigation software.

Do you think the type of software you use to conduct forensic IT investigations plays a decisive role in extracting evidence that is critical to your investigation?

Response: Yes: 12 No: 1

From a number of positive responses to this question we note that software plays a major role when investigations are conducted. When we compare this to the section Software Used, question 1 (Appendix A) we note that participants prefer three types of software packages above the rest. Participants B, C and G also prefer using Encase (Appendix A). While this is a preferred software package, it does not allow complete customised script creation. This makes customised investigations difficult and closely matches earlier reference to "button investigators".

This point to the inability of investigators to comprehend the full spectrum of the investigation and only make use of the software packet features instead of scripting their own commands in a customisable manner to gather specific data. Only Participant H and J indicate that they use customisable script to access data.

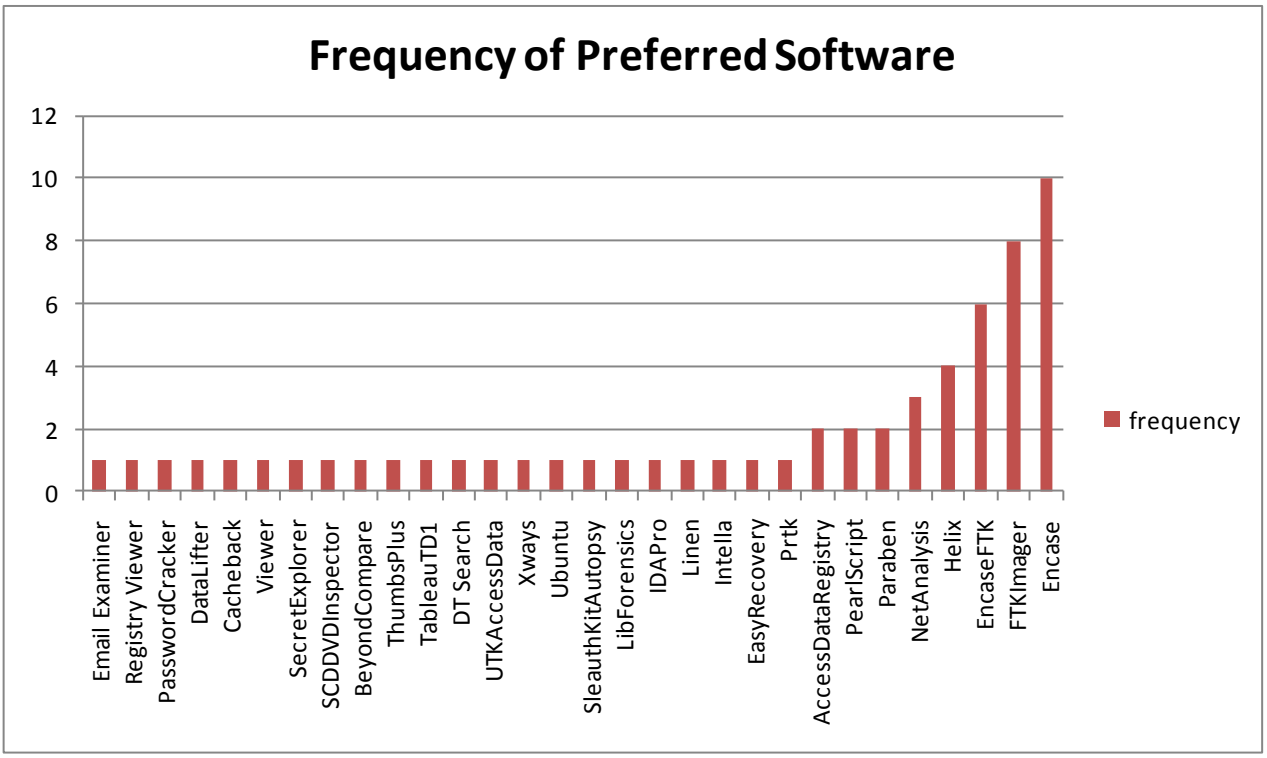


Figure 5 - Software preferred by Survey Participants

7.1.3 Databank Creation

Digital Forensic Automation (DFA) is not yet possible because of the diverse data bank structure. Data banks will have to be created with a vast number of tables and covering many, although not all possible scenario variables. According to (Garfinkel, 2007) such a database does not exist.

If we had sufficient data for building a structured taxonomy in forensic modelling which defines how we would conduct investigations, and produce groups of similar clusters, we might succeed in getting a higher accuracy level. This would also lead to a platform whereby associations among predicted data are more defined from a procedural point of view.

Data must be gathered before investigation can take place; this is not always possible since modern computers hold many gigabytes of information. Data now has to be grouped and associated in clusters for analysis. Extracting live data while the computer is running is a challenge because data corruption is a risk. When we look at technical issues, computers are evolving fast enough to allow solutions for data gathering with the help of toolkits. Unfortunately changes to toolkits are not in pace with technical advancements. Social impact plays on the lack of standardisation which eventually is reflected in a non-standardised investigation process. Data gathering then would be unproductive because not all of it is usable; we also have to consider how privacy has been ensured all these years in investigations that took place in the safe haven of private and protected data.

7.2 Enhanced Framework

We would like to create an automated databank that is established by a global forensic group of contributors. However, we foresee difficulty creating a databank that allows automated methodologies which proves digital forensics can be automated. Creating a new predictive automated model, as presented from the research findings has a fairly low-level of approval from participants.

Nevertheless, in equal comparison, creating a model based on the existing frameworks did not look promising. This is because a vast number of variables play a role and it is difficult to determine the likelihood of similar events in a digital crime. One of the drawbacks had been the small corpora of research in forensic data banks. This is because of the non-existence of a proper data bank to confirm case findings. Setting up a relational database should strengthen and reflect the reliability of these predictions.

At best digital forensics has always been associated with unpredictable data which were unexplained or unallocated when analysis took place. We note from various studies that it is nearly impossible to pinpoint a specific investigation methodology suitable for all digital crime scenes. Forensic investigators cannot guarantee how and when investigations are conducted based on existing frameworks.

Arguments from other forensics disciplines like law enforcement and social science, computer science and information systems are of thought that each discipline has their own expectations and interpretation of forensic data. It is therefore important to have a continuum of development in mind which will eventually interact on an interrelated platform or framework enabling satisfactory findings based on international standards. This would further provide the opportunity to present a template from which analysis can be conducted according to these standards.

As we note from various frameworks, they all have the same common principle – to identify, examine and report data. We could also add some other phases when we consider the importance of core phases, but after all it comes down to having a consistent standardised model. This is also obvious in the type of investigations conducted by skilled technicians or academics. Irrespective of internal tensions or clashes between the two investigation groups, a standardised platform must be explored.

We therefore propose a basic investigation platform and suggest phases of investigation, although the precise process is open for discussions as it is not clear from some participants if they require a fixed model.

Several previous frameworks are referenced in this thesis and we propose an alternative approach to initial framework reassessment. This allows the setup of a platform that forms the basis of analytical investigations that eventually sets a platform for reclassification purposes. Reclassification is required to ensure a model in which data is grouped and clustered in a relational structure.

In this thesis a framework is proposed, covering all higher level phases of the investigative process:

- Awareness/presence
- Find
- Duplicate/preserve
- Initial hypothesis
- Analysis
- Continuous reassessments
- Documentation
- Preserve/storage

Most of the core phases mentioned in previous frameworks can be incorporated into this framework. This framework also sets a legal base as foundation. Because of this a clear understanding of the legal requirements is proved right at the start of the investigation and informs each resulting step or phase. The most applicable framework and integral steps will become clear.

The Investigation stage should include at least the following:

- Searching for and identifying evidence on a computer;
- Collection of the evidence from the computer (original is duplicated);
- Transportation of the evidence to a secure environment;
- Storage of evidence collected at the scene;
- Examination of the evidence using the proper tools (finding incriminating evidence);
- Analysis (looks at the product of the examination to find out the significance and value of the evidence found) (Kohn, 2008).

The question is “what” to do with the data and related issues like sub classification and inter-relational dependencies. After the classification takes place the investigator move to the level of “how” to do it. The full spectrum of the investigation takes the form of *Methods (What) vs. Procedures (How).*

(Rurbin, 2005) proposed a framework that displays the benefits of computer intelligence technologies. It uses automatic evidence extraction and provides a basis to build more knowledge through reusability. This might result in great savings on human resources and creating more Real Data Corpus that would be supplemental to the creation of a large-scale unclassified corpus of real information from real computer users all over the world. It is the author’s view that such a data base then should be carefully modelled allowing sub-level classification.

7.3 Simplified Investigation Model

I suggest simplifying previous frameworks and presenting a model with four quadrants of major activities. I call this the *Quadrant Phased Investigation Model* (QPI) Figure 6. We use known phases but re-aligning these to produce a quadrant slicing of different sub-phases, showing clearly how the investigation is interlinked and progresses. This model can be further enhanced in the 3d and 4th quadrants, adding more detail to the processes.

The previous sections outlined several important forensic frameworks. In this section a new framework will be proposed. The aim of this thesis is to open discussion around a new proposed framework, presenting several infinite steps, and suggesting a framework that is guided by the type of investigation and specifies processes most suitable for the investigation. I also suggest a framework that has the ability to be adaptive. This shows changes to the Collective Phases but especially with consideration to the 3d quadrant when all other processes interact and are dependent in a relational context. Figure 6. This would enable inter-activity with other related phases where sub-processes are linked to a basis of similarity; this might present opportunities of possible relation to instances of the same type of forensic activity.

It is important now to determine how these top structures would look like and how they can be build in such a way that quick changes can be made within the framework, still preserving the basic flow of the investigation. In each quadrant data is collected according to a defined documentation process. This would allow development within the Collective Phases and accepted improvements as sub-phases are built on comparisons and lead to data mining options. Research by (Kohn, Eloff, Olivier, 2008) is considered here based on assumptions and suggestions and building a fuzzy analysis of likely outcomes of the investigation.

The QPI model I suggest, mainly focus on defining and integrating phases of initiation, preparation, data collection and extraction, hypothesis pre- and post investigation, investigation or data analysis and adding data to a data bank, final representation of data in court or reporting. Documentation and comparing data as confirmation is also required as is processing based on initial modelling and processes.

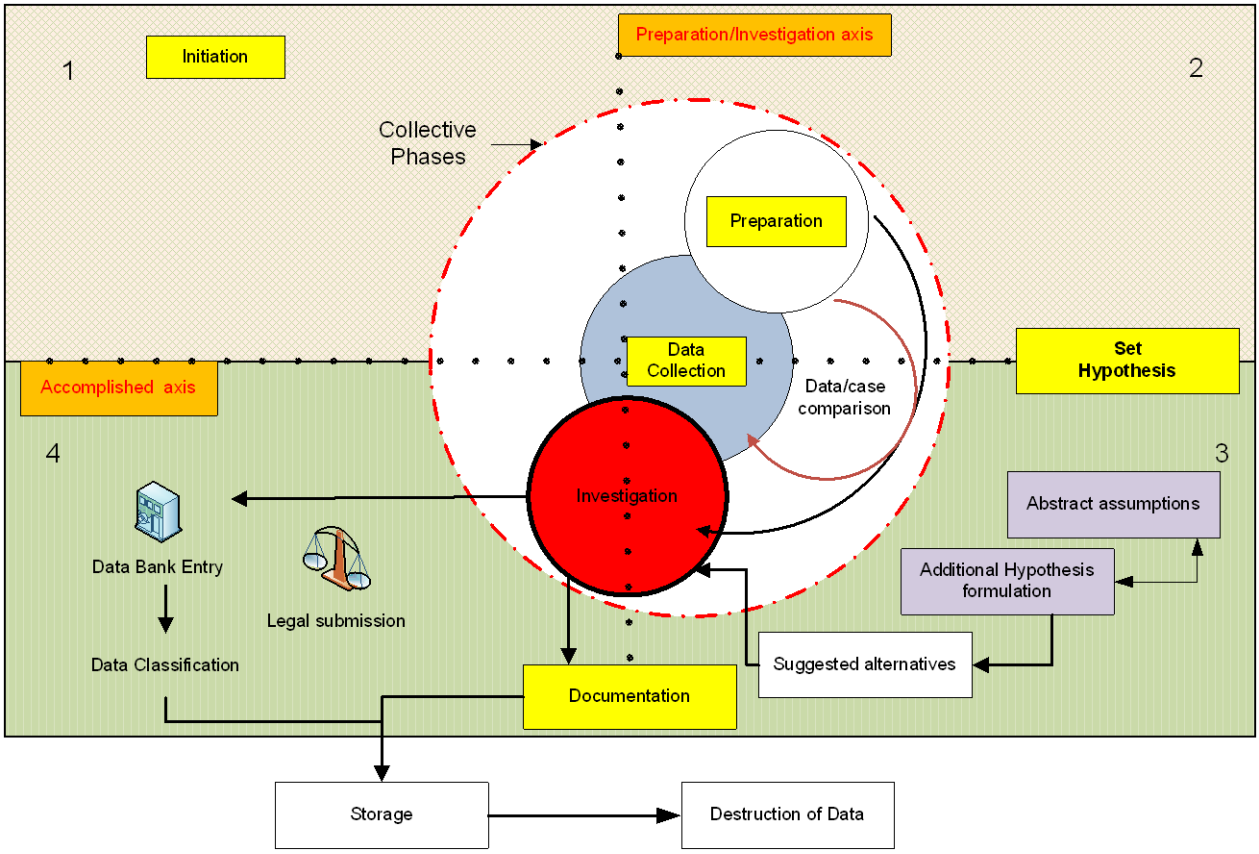


Figure 6 - Quadrant Phased Investigation (QPI)

1. Quadrant 1 - This investigation model starts with Initiation that only touches the Collective Phase, specifically Data collection. Initiation of the investigation indicates an awareness of potential disruptive activity and less than 18% of data is collected in this stage.

2. Quadrant 2 - Preparation and preliminary Investigation, more than 60% preparation and data collection; hypothesis formulation and start comparing data.

3. Quadrant 3 - Touch points between data comparison and investigation guided by abstract assumptions and additional hypothesis formulation as the investigation progresses as alternatives are considered or required. Setting a hypothesis takes place between activities placed in the 3th quadrant of the Collective Phases. This is after more information has been collected and initial data comparisons are conducted and related to Abstract assumptions. We can now form an Initial Hypothesis which interacts with Additional hypothesis formulation and Suggested alternatives. This phase is also influenced by other disciplines and involves more data associations. Data that borderlines between phases should still follow the objective of the initial investigation. One future objective of all investigation is collection and preparation of data for data bank entry.

   All investigation phases are documented. Presentation and Legal submission is performed just before data classification and storage. Data destruction should only be considered after final data analysis has been confirmed and data has been input in the data bank. More emphasis is on quadrants 2 and 3 because each successful accomplishment is setting the basis for the next phase and expected result. Most future research would focus on this quadrant, as this shows activities that have links to other phases proportional to the other quadrant dividers.

4. Quadrant 4 – Here we find important data bank updates once investigation is completed by following the bigger phases of the Collective Phases. Since the documentation process is evident in all quadrants, a

clear data structure of the calcification process would be evident and selected data is accepted into the data base. This will contribute to the building of the global forensic corpora. Once the data has been compared we could affirm whether alternatives to the hypothesis aided the final data acceptance.

5. Always consider if the data is to be archived or destroyed. Consider data destruction on a case by case basis.

The Collective Phases of Preparation, Data Collection and Investigation are integrated since the Investigation phase is dependent on Hypothesis formulation. The continuous processes of finding answers is determined by alternative approaches to solve irregularities as the investigation progresses and data comparisons are put in place to validate the initial hypothesis. This formulates progressive model of investigation. Data bank capturing is ever present but especially in the final stage when strict data classification models in place ensure data is correctly entered. Evidence witness reports are possible at any stage between data bank creation and destruction of data.

The final stage of any forensic investigation should include a Presentation stage. This stage is important because it satisfies the key requirement specified by the definition of the word “forensic”. This stage should include vital steps such as presenting the analysis, and proving the analysis matches either an initial assumed theory or the findings were accumulated as the investigations progresses and new data came to light. Therefore it is very important in the final steps to pay attention to presentation. Accordingly, it might be difficult to bind a theory to the findings; however, it might be easier if investigators make use of a tested framework, which might include a basic generic standardised reference for comparison of future data. This provides an initial starting point for the investigation hypothesis. The evidence presented should also hold up in court if the proposed framework and all previous steps were followed correctly.

The proposed framework quadrant approach has highlighted the importance of the Collective Phases. Knowledge of a relevant legal base prior to setting up the

framework is vital, as well as setting an initial hypothesis, since this will have a bearing on the whole investigative process.

7.4 Automated Future

As collected data is documented and structured in a classified manner, relational data structures would enable the creation of a table structure. This research suggests a possible automated framework based on which investigators with none or limited forensic experience could run a new investigation. Unfortunately, information collected from different digital crime scenes is never the same which creates problems for the design of a template for automated investigation methods. Another drawback is for new forensic investigators, who find it difficult to test software or write new script when working with big data sets.

If data from previous inspections are stored after investigation and analysis according to a set framework that is unique to the type of investigation, it would assist in finding similar patterns in data mining results. At present this does not seem to be the case. Noblett (2000).

Currently research is not geared towards developing new methods for finding probable data mining and knowledge engineering techniques to extract this data. Researchers also have to look at cross-border operations and the problematic inspections of hard disks in another language, for instance. A template design or automated process to predict the presence of certain types of data would allow solving cases faster. Garfinkel (2007) mentioned that problems about existing forensic corpora do not allow consistent data analysis because the data banks are too small.

It appears that investigators are handicapped in many occasions when software documentation does not allow for changes of standardised script made to the data. Investigators have no access to the internal processes or specific code and are therefore not allowed to customise new requests to the hard drive. The closed environment makes it difficult for investigators to understand the internal operation or interpretation of the program package; fortunately some of the closed code does

however present freedom to interpret the data on a case by case basis if the investigator does not start more complex analysis.

Since the ability to rewrite code script is essential for customised case investigations, it also adds a different perspective when differentiation to the normal code is allowed. This in turn would allow for automated calls over and above the normal call of standardised script. Automated inspection of data is time-intensive and automatic transcribes or data interpretation does not allow enough freedom to inspect the specifics of the case. Automated software might not cover all aspects of extraction required for unique cases.

Supporting the statements above, participants responded as follows:

Do you think an automated digital forensic investigative process is feasible?

Response: Yes: 6 No: 8

From this result we note a strong response for and against automated procedures. Most concerns are that no investigation is ever the same as before and automated procedures might miss unique or complex associations which only an investigator can interpret. This is arguable; since human investigators might not always associate similar cases from the databank with present investigations, when similarity of the cases exists.

However some participants presented a clear preference for automation, indicating the importance of having a basis of initial investigation processes. This notion is similar to the author of this thesis' perspective that a tiered approach from initial level investigation leads to rationalisation of redundant data, thereby allowing more detailed assessment of the factors at hand unique to the particular investigation scenario.

An earlier response from participant C, shows a contradictory response here.

Participant C indicated:

" Forensic automation is already becoming a problem by giving untrained examiners a false sense of security when in reality; they are not conducting an examination at all. When used properly some automation is good. However, it is not to the point where any

time in the near future, an automated tool can conduct a thorough enough examination to be trustworthy."

Based on the above response we need to change perceptions about Digital Forensics Science and move towards creating forensic corpora from global contributors following strict reporting and data structures. Privacy is a major concern. The lack or willingness to address privacy up front is an obstacle in the way of would-be researchers.

The term automated forensics and automated tools are misunderstood – as participant C above comments. In earlier sections we already addressed software that reflects a click and drag scenario, which does not promote an academic research platform. I suggest automated forensics be the process of investigation where the investigator makes use of previous data based on predictive analysis of data bank containing previous data and makes use of forensic software to a lesser extent.

I also suggest changing the mindset from "automated software", to "automated analysis" whereby investigators could sift through the first level of classification and determine sub-levels of the investigation with optimal running of scripts – suitable for level comparison and prediction.

Questions arise such as how many times does a hard disk have to be inspected to verify data? Or is it possible to extract sufficient array of data without having to go back in the investigation process? Answering this from an automated perspective, we suggest that varied templates in automated script would allow sifting of data faster, thereby minimising risks of data loss. Danger of damaged information comes to mind when data has been extracted and has to be transported or stored for delivery to the evidence room or to the court room. The use of check summing before and after data evidence is very important. In this light we note the importance of a proper high-level process to verify all data.

Questions arise whether data might get lost if tests had to be re-run based on different data verification processes, thus the data's integrity would still be maintained.

This argument would not be valid when real time inspections are conducted and special processes must be followed to ensure captured data does not interfere with streamed data packets.

Presenting a tested automated framework, based on a methodology with rigid processes should therefore presents findings that reflect assumed outcomes of a particular case. Investigators must take in account what role data vulnerability plays in the investigative process. Functional areas of vulnerability are important because if these are not securely met we cannot verify the data.

For instance:

- Identification and authentication
- Accountability
- Object reuse
- Object reuse
- Accuracy
- Reliability of Service

When detecting the source of the intrusion, the framework, we should verify if and when data tampering has taken place and verify that forensic evidence in fact does exists on the alleged computer. We also need to prove if and when data was stored on the computer. Thus, when I position “automated processes” mainly based on the findings obtained from a hard drive’s investigation, I note that various factors still play a role, pre-occupying process influencers in establishing a platform for automation.

7.5 Automated Results

To be effective, digital forensics research should consider focusing on the following areas:

- advances in automated malicious detection and legal reverse engineering technology;

- heightened focus on wireless technology, its vulnerabilities, and the forensic indicators that will assist operations personnel and investigators in identifying questionable activity; and
- Continue to work toward the establishment of approved standards and best practices to strengthen the foundation for Digital Forensics Science. (Palmer, 2001).

Automated responses would only be successful if a big enough corpora base had been established. This would enable the creation of a data set that supports forensic examiners' to search for the best combination of words or relevant case selection identifiers. Researchers are not pursuing automation because they do not have sufficiently large corpora of forensically interesting data to develop reliable automated algorithms and tools. Instead, much research in both the academic and corporate worlds has emphasised the development of interactive visualisation tools. Since they are designed to be operated by a trained individual, tool failures can be more readily tolerated. Questions about "forensics expert's" ability come to mind.

In some cases, there can be only one way to achieve some objectives, e.g. by observation, a structured interview or a questionnaire.

Chapter 8 – Practical Implications of this Research

To conclude the thesis, please re-consider some questions posed in Chapter 1.

I provided two sets of answers by Keith: the first set (A-type) presents his current answers and the second set (B-type), assumed he followed a method of standardisation. I also provided a summary (C-type), addressing some major developments in regulating training and certification in the future. Universally recognised training should set a platform for developing investigative skills, thereby minimising discrepancies when solutions are provided. These should match across the board if investigators are using the same high level professional training and certification model. Under this proposal, we could move closer to a standardised platform which motivates the goal of my research – to establish an automated investigation platform.

8.1 Examples of A-type questions.

A-type answers are easier to work with because investigators are not challenged or questioned before, during or after investigations. A-type questions and answers address issues from a general "level one" perspective and allow scope for variation. The accused could argue that these questions could have wide interpretations that are not fact-based. A guilty finding would be unfair because these questions types present arguments disputing the applicability of specific investigation procedures. We could literally bind Keith's answers to the questions, as his answers are based on his own perspective as well as on obvious occurrences in the particular case. However, questioning a defendant's involvement in this manner might indicate that investigation hardly goes out of a prescribed range of possibilities.

This indicates that investigators allow only small deviations from their initial starting point while conducting the investigation, which also reflects the constricted explanations in the case.

Cross-examination could upset Keith if his answers to these questions cannot be proven and instead show that correct procedures were not followed.

- Do you use previous investigation data as guideline to conduct new investigations?
  Yes – No / I am not using any data as each case differs
- Is there any specific part of the investigation that is more decisive than the other in the forensic investigation?
  I am not sure I understand the question
- Do you think the type of software you use to conduct forensic IT investigations plays a decisive role in extracting evidence that is critical to your investigation?
  No, I use my own software, and I obtain good enough results
- Do you collect data of previous cases currently?
  No, I do not see the use – each case is different
- Do you think that having a database of previous cases (corpora) can help in digital forensic investigation and case analysis?
  No – as said before each case is different.
- Please indicate your preferred order of the Sequence of Investigation Processes
  I am not bound to a procedure; I do not prefer a specific investigation order and normally change my approach as the case develops.

8.2 Examples of "B-type" questions.

B-type questions present additional coverage of Keith's (the investigator's) perspectives. His answers now reflect his experience as well as his moving towards deeper analytical interpretation. We assume prescribed methods are followed which reflects answers with firm reasoning, in other words, they are based on experience and tested methods. Comprehensive investigation techniques are suggested and investigators at this level deliver confident in-depth interpretations.

Detailed descriptions of the investigation processes are presented and references to frameworks in use setup a preliminary investigation methodology.

It would become increasingly difficult for the prosecution to find evidence of misconduct in the investigation process, since Keith would confirm reasons for his findings with proper examples taken from the investigation. Keith is now also in the position to convincingly argue his point of view that is based on a firm standardised method in the investigation procedures framework.

For instance:

- Please describe your training and processes you followed in gathering evidence.
  I am a qualified digital forensic investigator with CCCEE qualifications obtained through 3 months of rigorous practical case analysis.
  I followed prescribed methods as stipulated in the FBI's investigation guide for first offenders.
- Which procedures did you take to safeguard the evidence from external interference?
  As mentioned before, I followed procedures as prescribed in the FBI's guideline. I also followed prescribed procedures and implemented methods as supported by the Scottish Police investigation guideline – described in the Good Practice Guide for Computer-Based Electronic Evidence
- Did you conduct typical investigations like this before?
  I have more than 2 year's practical investigation exposure and maintain active involvement with fellow digital forensic investigators through discussion boards with investigation group meetings.
- Describe the acquisition process and chain of custody in this case.
  In addition to following guidelines as stated before, I also use prescribed procedures as found in The Guideline for First Offenders.
  I am always aware of the procedures described in these documents regarding detailed documentation of all processes.

- Did you use off-the-shelve recovery software?

  I am using standard tested software, for instance Encase and Helix software. I am also using Linux based software that enable me to make exact hard disk copies. I substitute these packages with tested software tools. Before using self written script to extract data, I would test the software on a case with expected results in order to test robustness and reliability of the script.

8.3 "C-type" Interpretation

C-type questions assumed a firm standardisation platform. Solving regulation of digital forensic investigation requirements is considered. This process starts with setting up training and investigation procedures that would ensure investigators are competent in digital forensics investigation disciplines. It appears that academic expectations are vastly different from the core field investigator experiences in the industry.

Aspiring investigators are drawn to short courses, often a day or a week long that would lead to potential misinterpretations of complex digital forensic cases. Existing certification training programs, for instance the Global Information Assurance Certification (GIAC) and the SANS Institute are forerunners in certification pathways. To become a GIAC Certified Forensic Analyst (GCFA), you are only required to pass one proctored exam (150 questions, with 4-hour time limit) and achieve 69.3% (104 of 150 questions). The SANS Computer Forensic Investigations and Incident Response certification, covers a 6 days training session. (SANS, 2010) and (GIAC, 2010)

Preparing the "professional Keiths" out there, we need to set an absolute minimum expectation or regulation addressing issues of skill versus training, thus producing capable investigators that have knowledge of all aspects of digital forensic investigation. Skill upgrading begins with initial training, which starts when training and additional aspects of research are strengthened under a structured and comprehensive training program.

To capitalise on this, training should follow progressive stages. Normally this training process allows minimum “certification” requirements after only 6 days.

Motivation for this is supported by recent research by Hom-anek (2009) who reports their findings on how information security training forms a basis for digital forensics qualification. Digital forensics professionals are regarded as having functional roles when implementing their skills and training. Hom-anek also notes that only 42% of digital forensic investigators have competency that matches their job roles. This figure confirms that training is a major shortfall in most investigators’ skills package.

8.4 Research Results

I found that even with the number of frameworks discussed in this thesis and the possibility of having only one framework that include all others, we still would not convince investigators to use a generic framework, since all investigations are different. The challenge is rather to establish a framework that produces a “gliding scale” of possibilities; this scale could then be used in conjunction with a data base that contains the same entities and variables of the most consistent occurrences of similar types of crimes, thereby matching possible case results with predicted results. This will open new a direction for predicted analysis of digital forensics to regulate a standardised new approach to investigation processes.

This thesis addressed the difficulty in establishing a framework that the digital forensics industry follows. As my research show participants hardly use a standardised framework during investigations.

I also note different participants’ approaches regarding perspectives on suitable digital forensics investigation experience and skills required in digital forensics. This became clear when creating a proven framework for a general investigation scenario. A reason for this is that investigations are never the same and investigators do not seem to mind other factors that might be related to other cases, thereby addressing similarity. If such a relational data structure exists, case similarities and predictions based on number of occurrences might emerge.

As discussed in this study, monitoring, logging and preservation of case data need to form a data bank, thereby establishing corpora for confirmed methods of fuzzy approaches and predications. It is questionable whether forensic investigators would follow a standardised procedure at all—considering they have been following their own customised methods to date. This presents a problem for standardisation and ultimately automation. With regard to training and skill enhancement, I noticed that a few participants regarded own experience higher than formal qualifications.

I support Beebe's point of view with regard to a complete re-alignment stepping away from the overwhelming traditional search patterns and moving to prediction of similar cases. We should consider all options when conducting investigations, where earlier exposure to similar cases and level of expertise is important. Enforcing a grading level that allows only experienced investigators to conduct investigations at a specific level presents a degree of professionalism in the digital forensics discipline. This means that representation in court should only be allowed for those that are skilled enough. Theoretical expertise sets the standard that reflects a comprehensive understanding of the specifics of any particular investigation case.

Defendants are often at the mercy of investigators with *"I am a professional forensic investigator"* attitude. This might not necessarily reflect their skill or experience. Setting a certification regulated dependency would ensure investigators are in fact as good as they say they are. In addition, various software packages were investigated in this research. It seems that only a few core packages were chosen for the sake of simplicity in conducting investigations. Are there any investigators who are writing their own tested script and conducting investigations to suit specific enquiries? High costs of obtaining off-the-shelve software and yearly maintenance fees make this very costly for novice investigators.

Automated software seems to guide investigation intensity instead of automated procedures.

Are we getting “click and drag” investigators, reflecting real “experience” in the market, or are these investigators trained to the minimum expected level? As the field study suggests, these investigators are also known as “button investigators” since they do not have proper training and insights of underlying investigation processes. It becomes questionable if these investigations reflect the investigators’ real experience and skills or we only get software-based solutions without interpretation?

It seems that a few core investigation software packages have been written and most investigators follow this trend regardless of the fact that the investigation does live up to new expectations geared towards demanding research or originator tracking. To the contrary, it seems that only a few investigators follow an academic approach by writing specific script for solving challenging cases. If the present prescribed and recognised methods, as suggested by industry, are the only model to work with, investigators might present incomplete solutions without recognising a standard requirement. This situation might be rectified with extensive training.

Therefore, if investigations are not based on a sound theoretical basis and do not utilise a hypothesis that formulates a framework describing the processes followed, we assume investigators would not find a connection point in solving the case or proving where these principles originated from, thus failing to present academic reasoning or comments in court.

When we re-consider Chapter 1 that presents academic and technical approaches to certification and research, we again find the need for certification. In other words, if there is no certification, then no knowledge of academic interpretation and prediction is derived – thus no standardised platform is present. We may expect un-classified findings from unskilled investigators that might not follow a standardised framework.

Digital Forensic Investigators should demonstrate an inquisitive approach to solving cases, backed by theoretical knowledge about finding and associating hidden data. This should at least be the expectation, since defendants should get the best protection allowing the benefit of doubt.

The development of a Digital Forensics Body of Knowledge (DFBK) has been slow for a number of reasons; including the lack of experienced participants and the lack of collaboration amongst digital forensics professionals.

I suggest we need to have high-level peer-reviews in a selected journal that is published for a selected audience, thereby identifying and guiding training and certification requirements, since technological progress demands constant retraining and updating to minimize redundant information. Given the responses we received and further assumptions of the digital forensics trade in the field, I cannot confirm that all aspects of responsible investigation processes are followed by the participants. Based on my findings to date, I cannot support whether investigators with limited academic knowledge should be able to defend a case in court – irrespective of field experience. If such an investigator makes it to the court, the process of maintaining a rigid standardisation procedure loses accreditation. Hannan (2003) also voiced a similar point of view.

From collected data and industry trends, I recommend a skill training pathway covering at least the following aspects towards certifying experts to testify in court:

- Digital Forensics Certification that covers thorough pre-course preparation and preferably basic file, network and system security;
- Advanced hard disk functions and digital storage principles;
- Advanced file hexing, registry/ root / hard drive functions;
- Open source, Linux/ Perl/ DOS scripting;
- Security and network access, penetration testing, reverse engineering;
- Understanding of advanced theoretical framework and hypothesis formulation based on similar investigations;
- Peer testing during and after investigation as well as archiving in a databank for future reference;
- Frequent up-skilling on the latest changes in software and testing – in addition to own developed script solving particular case requirements.

Setting a high standard of certification and academic research should bind this to an implementation model thus following specific recommended guidelines.

This would allow seamless updating of specific procedures when challenged by newly developed processes. Furthermore, this will ensure all investigators are bound to regulations and will set a standardised working discipline.

Once a standard of investigation is set, I firmly believe the creation of a global data bank is possible. Further to this, if relational data structures are established, bulk data generation is possible that allows for sifting of data through predicting and identifying future occurrences. I also foresee using semantic search functions, interlinked with predictive models whereby pattern recognition and re-occurrences of similar crimes are indentified.

As this thesis shows, future research in this particular field is required; however, getting sufficient data might hamper its progress. One could only hope for open contributions from major players allowing in-depth research in future that will lead to standardisation and databank creation that allows for automated investigation.

Chapter 9 – Conclusion

9.1 Research Summary

I started out researching traditional forensics frameworks by comparing different investigation models. Although I have studied various points of view presented by researchers, it appears impossible to find a broad framework that lends itself to automated procedures. Participants' experiences show that the existing investigation framework model does not present a suitable framework in real investigation scenarios. I received expert feedback in the field when exploring the possibility of a standardized framework. Some experts consider it as being desirable, possible and even necessary to look at alternatives. Although I expected a higher level of participation or suggestions to this topic, the interest among participants was relatively low. I also discussed the mismatch between my early assumptions and the actual data collected from participants.

I found that some digital forensics investigators with an agenda of "just-get-the-job-done" might argue the digital forensics discipline does not need to follow a standardised platform. It appears their success ratio is based on their own methodologies, if in fact they are conducting investigations based on any known framework at all. In contrast, from an academic point of view, some authors addressed constrains such as legal requirements or lifetime of volatile data in live forensic investigation, as issues when a standard is followed. This adds to the real issue of reliability of actual investigation processes and stresses the importance of proof and legality of data origin.

Digital forensics is also lacking confirmed and tested methods; this underlines the urgency to standardise procedures and processes, to ensure proven and consistent results. Digital Forensics Science needs to take a bold step towards a new approach for defining and standardising investigation processes and to make this a confirmed platform for adaptive globalisation. From our research findings, it appears that standardised methods and procedures had always been pushed aside as being the responsibility of bigger corporations.

In addition, it also appears that smaller digital forensics investigation companies do not conduct their investigations based on (any) digital forensics standards. This means forensic investigators could potentially follow their own procedures that are not bound by a standardised platform. If a globally accepted digital forensics investigation platform does not exist, then standardisation becomes impossible and investigation results might differ from one investigator to the next.

This thesis identifies a major current issue in digital forensics, namely the need for consistency versus current practice. I also discussed current expert expectations of available frameworks. When I compared this to academic research that prescribes a standardised platform, I noticed a mismatch between the expectations of forensic technicians. This highlights the problems for setting up a standardised platform. I also explored some existing frameworks and suggest initiatives related to investigation methods. For instance;"are automated investigation methods possible?" If suitable investigation methods were followed it would inherently support a standardised framework. Investigations would then be carried out faster based on a standard that allows recognition of both forensic technicians and academic perspectives.

I noticed underlying tension between investigators with an academic approach and forensic technician investigators that are only interested in solving cases. I also observed a division between investigators, with its basis for disputes grounded in the lack of a standardised forensics discipline. No firm agreement for setting standardised investigation and training methods has been reached yet among the professionals in this field. As a result, this research finds that participants reject the idea for a global data bank. One reason for this might be because privacy and secrecy of the trade is a protected field, divided between participants and contributors. This reflects the way academics conducts research compared to technicians. Academic researchers adopt an investigative or comparative analysis, whereas technicians mainly complete their investigations at result level, clearly – as my research shows – without deeper analysis. This group focuses on interpretation and prediction as well as on setting a range of standardised types of questions.

Academic researchers' shows an academic approach to solving cases based on earlier data. This further emphases the need for standardisation of investigation procedures and training. In order to resolve this, a practical approach to either retrain technicians or retrain academics trying to find touch points in similarity, might provide a standardised process of investigation. Currently these two opposites are extremely divided. This might even lead to an overhaul of the existing discipline and standardisation methods. It is especially noticeable from a few participants' responses, referring to inconsistency in using a recognised framework in their investigations. Only a few participants make use of a set framework, suggesting that only a small group follows a (standardised) method, while conducting investigations. More questions arise, such as:"Do we have to choose between a (standardised) platform suggested by technicians and one suggested by academics?" Is this in fact the industry standard to work with?

Based on participants' feedback, I attempted to discover whether investigators were consistent in producing repeatable results which other investigators should be able to confirm, if they used the same investigation procedures. Response from investigators shifts towards new methods of investigation to database creation. For instance, better classification could lead to better predictions based on a large number of case examples. Given that many inconsistencies and preferences exist in various data gathering methods, questions about the feasibility of an automated digital forensics method is raised. Setting up a standardised control for digital data forensics would potentially show a firm commitment to specific management controls, which could lead to a fully automated process.

I also discussed participant responses and attempt matching their procedures with some existing frameworks to predict similar cases. From these assumptions, I suggested creating automated forensic corpora, which includes relational data structures and automated script. Future automated processes were briefly covered as well thus *promoting automated investigation rather than automated software methodologies*. Although these processes are separate issues, setting up a digital forensic corpus should be a starting point. It is the author's view that such a corpus should be developed in a manner that allows sub-level classification and expansion.

Developing a Real Digital Forensics Data Corpus (RDFDC) would be useful in data analysis; this would allow possible automated investigation processes. Emphasis on strict privacy guidelines while gathering data would be in place. To my surprise this was not considered favourably by the participants. While listing participants' arguments for and against automated investigations, divided interest in creating automated procedures came to light. The *present participant group did not use any automated investigation processes and corpora, therefore making the creation of a dynamic* (RDFDC) *a separate issue for future research.*

An interesting phenomenon is noteworthy. *It became clear from the participants' feedback that standardisation borders on impossibility, since there are too many different cases and forming a generic automated platform would be nearly impossible. Two distinct groups emerged; about 25% of the participants said they were interested in a new model, while the rest pointed out that the present order of investigation was acceptable and required no changes.* Despite the need for standardising and confirmation of data consistency, this sequence is obviously missing from other participants' views.

Another dissimilarity between two distinct groups emerged - uncertified investigators in contrast to certified investigators. I point to this important issue in the research findings section and the effect this has on standardisation and digital forensics case interpretation.

I found that even with the number of frameworks discussed in this paper and the possibility of having only one framework that include all others, we still would not convince investigators to use a generic framework, since all investigations are different. The challenge is to set up a framework that produces a "gliding scale" of possibilities; this scale could then be used with a data base that contains the same entities and variables of the most consistent occurrences of similar types of crimes, by matching possible case results with predicted results. This will open new a direction for predicted analysis of digital forensics to regulate a standardised new approach to investigation processes.

This paper addressed the difficulty in settling a framework the digital forensics industry follows. As my research show participants hardly use a standardised framework during investigations. As discussed in this study, monitoring, logging and preservation of case data need to form a data bank, by setting up corpora for confirmed methods of fuzzy approaches and predications. It is questionable whether forensic investigators would follow a standardised procedure at all—considering they have been following their own customised methods so far. This presents a problem for standardisation and eventually automation. On training and skill improvement, we noticed that a few participants regarded own experience higher than formal qualifications.

Reasons why specific industry certification is required leading to a standardised platform should be clear. Once this issue is resolved, we might add automated investigation allowing for future prediction of digital forensic instances.

Digital Forensic Automation (DFA) is not yet possible because of the diverse data bank structure. Data banks will have to be created with a vast number of tables and covering many, although not all possible scenario variables. According to (Garfinkel, 2007) such a database does not exist. If we had sufficient data for building a structured taxonomy in forensic modelling which defines how we would conduct investigations, and produce groups of similar clusters, we might succeed in getting a higher accuracy level. This would also lead to a platform whereby associations among predicted data are more defined from a procedural point of view.

Chapter 10 – Future Research

10.1 Practical Implications of this Research

I propose further study in forensic profiling, particularly establishing a basis of interaction between automation and profiling, thus creating a stepping stone for initial time-saving when a typical investigation is conducted. (Rogers, 2003) reference a research by (Pethenck, 2002) showing how criminal profiling might be achieved if a broader guideline is used as it was earlier suggested by the FBI. If we use these FBIs typology which was criticised for not having enough empirical testing, we might come up with a higher socio-criminal identity, thereby classifying potential criminals according to characteristics of typical groupings.

(Elsaesser, 2001) as referenced by (Stallard, 2003) presented an approach to generating automated hypothesis of computer attacks. This framework then simulates the computer attack and assumes matches to a target configuration using recognition techniques through searching for unique supporting data or patterns of the investigation. By using this approach we could also broaden the use of data extracts from data bases, finding the relation between fields and tables thereby getting patterns of similarity when constrained items are sifted using redundancy validation. The table structure should be designed in such a way that the vast number of variables could be checked for redundancy to ensure validity. This process of linking characteristics of a specific crime through the tiered level descriptors would allow a gradual disqualification of redundant data/characteristics and would lead to a sub-level classification or grouping of the crimes. This would be an eventual tool for aiding forensic investigations.

Forensics investigators should follow an internal forensic policy or a guideline that specifies roles and responsibilities; this would ensure a clear and more efficient handling of cases, without possible discrepancies. Earlier research by Hall (2002) and Landman (2002) points to these differences and although similarities in the frameworks are evident, the investigation processes lack detail addressing a specific forensic criminal activity.

One reason for this might be that in the past much research has been carried out on the theoretical foundations understandings of the assumed framework instead of making sure that an adaptive framework exist from which variants from all the other possible criminal scenarios have been eliminated. This would ensure that a pool of possible criminal activities has been scaled down to a framework of likely outcomes.

We know that predictions are inherently impossible since every crime is different. We therefore need to make assumptions on a higher level and gradually break this down to a tiered level investigation that allows an interchangeable approach based on the initial framework. This process of linking characteristics of a specific crime through the tiered level descriptors would allow a gradual disqualification of redundant data/characteristics and would lead to a sub-level classification or grouping of the crimes. This would be an ultimate tool for assisting forensic investigations.

In addition to this, investigations are conducted with a set of questions which would guide the investigator when eliminating scenarios of similar crimes investigations. (Beebe, 2009) suggests using an Intelligent Analytical Approach where artificial intelligence and other intelligent search would enable successful retrieval by making use of algorithms. *This supports my point of view that higher emphasis should be placed on semantic rather than literal searching techniques that should substitute traditional literal searches. This allows for a structured, but still adaptive, relational data structure thereby improving data indexing.*

I suggest using predictive Markov models, analysing data for predictive similarity in events and consider a fuzzy re-classification of data models. Using a Fuzzy logic approach in data classification and clustering, presenting a new approach into re-classification that is not bound to factual rigour, but rather focuses on occurrences and predictability. We now need to determine how these top structures would look like and how they can be build in such a way that quick changes can be made within the framework, still preserving the basic flow of the investigation.

This would allow development within a collective phases and improvements as sub-phases are built on comparisons that lead to data mining options. Research by (Kohn, Eloff, Olivier, 2008) is considered here based on assumptions and suggestions with an end result of building a fuzzy analysis of likely outcomes of the investigation.

(Rurbin, 2005) proposed a framework that displays the benefits of computer intelligence technologies. It uses automatic evidence extraction and provides a basis to build more knowledge through reusability. It is the author's view that such a corpus should be precisely modelled allowing sub-level classification and expansion. This might result in a Real Digital Forensics Data Corpus (RDFDC) that would be beneficial in data analysis.

10.2 Enhanced Automated Investigation Framework

Creating a (RDFDC) from global forensic contributors would be the ideal. However, I foresee difficulty creating a databank that allows automated methodologies which proves digital forensics can be automated. Creating a new predictive automated model, as presented from the research findings has fairly low-level approval from participants. Nevertheless, in equal comparison, creating a model based on the existing frameworks did not look promising. This is because a vast number of variables play a role and it is difficult to determine the likelihood of similar events in a digital crime. One of the drawbacks had been the small corpora of research in forensic data banks. This is because of the non-existence of a proper data bank to confirm case findings. Setting up a relational database should strengthen and reflect the reliability of these predictions.

Responses from participants vary towards a shift from new methods of investigation to database creation. For instance, better classification could lead to better predictions based on a large number of case examples. Given that many inconsistencies and preferences exist in various data gathering methods, questions about the feasibility of an automated digital forensics method is raised. I note a strong response for and against automated procedures. A participant reflected on forensic automation as follows:"

Forensic automation is already becoming a problem by giving untrained examiners a false sense of security when in reality; they are not conducting an examination at all. When used properly some automation is good. However, it is not to the point where any time in the near future, an automated tool can conduct a thorough enough examination to be trustworthy."

Most concerns are that no investigation is ever the same as before and automated procedures might miss unique or complex associations which only an investigator can interpret. This is arguable; since human investigators might not always associate similar cases from the databank with present investigations, when similarity of the cases exists. However some participants presented a clear preference for automation, indicating the importance of having a basis of initial investigation processes. This notion is similar to the author's perspective that a tiered approach from initial level investigation allows a more detailed assessment of the factors at hand unique to the particular investigation scenario.

The term automated forensics and automated tools are misunderstood. I also showed on using software that reflects a click and drag scenario, which does not promote an academic research platform. I suggest automated forensics should prescribe the processes of investigation when the investigator makes use of previous data, based on predictive analysis from a data bank, which contains previous data and makes use of forensic software to a lesser extent.

Furthermore I suggest changing the mindset from "automated software", to "automated analysis" whereby investigators could sift through the first level of classification and determine sub-levels of the investigation with optimal running of scripts – suitable for level comparison and prediction. Automated responses would only be possible when a (RDFDC), with relevant case data had been established. This would enable the creation of a data set that supports forensic examiners' to search for the best combination of words or relevant case selection identifiers. Researchers are not pursuing automation because they do not have sufficiently large corpora of forensically interesting data to develop reliable automated algorithms and tools. Instead, much research in both the academic and corporate worlds has emphasised the development of interactive visualisation tools.

Since they are designed to be operated by a trained individual, tool failures can be more readily tolerated. Questions about “forensics expert’s” ability come to mind. This would ultimately present a match based on “fuzzy hashing” which requires a complete paradigm shift. This means we should step away from the overwhelming traditional search patterns and move to prediction of similar cases. I suggest using predictive Markov models for analysing data for predictive similarity in events. I would then move to a fuzzy re-classification of data models.

Most recent research by Kahvedzic (2009) suggested a model to describe an investigation at different levels of detail. This suggests that application of an independent vocabulary can be used to describe the investigating process in more detail. In similar manner, we could use this notion to present a clear data structure that defines specific groupings of similar concepts and their attributes, thereby ensuring representation of variables in a relational data structure. This would enable representation of hidden meaning of words with the same semantic intent and create variables in an automated relational data structure. Predictions are inherently impossible since every digital crime is different.

This database should be scalable ensuring new entities are related to the existing structure. Global forensic researchers would input data accordingly into this database. We envisage a noticeable benefit to the forensic community if members contribute and share their resources. This would provide realistic data sets that would assist in establishing a platform in digital forensics whereby an automated framework might evolve. Predictions are inherently impossible since every digital crime is different.

Linking characteristics of a specific crime would allow a gradual disqualification of redundant data/characteristics and would lead to a sub level classification or grouping of the crimes. Conducting investigations from available data in a databank and based on a classification framework presents methods that guide investigators in predicting case similarity.

Since each case differs substantially from any other, a model built from a generic level to predictive sub levels is suggested. This research did not cover relational database creation and classification of variables, further research will be

conducted. In other words, we form predictions, irrespectively of the investigation model followed. Further research is required in classifying variables and groups.

10.3 Additional research

Research to date provides evidence of forensic frameworks that only provide guidelines for major forensic occurrences. The field study shows regulation of training and certification might provide a basis for standardising academic requirements for this discipline. Extensive research based on a data base structure is required to enable predictions based on existing data. Proposed forensic scenario based on an initial generic platform would form the first stages of the research. Further development of a dynamic framework would enable sub-level associations/clusters. With hidden Markov and fuzzy logic implementation this would allow smaller data sets to be used with more certainty, and would also still allow for predictive assumptions.

The nature of the topic dictates the use of both a chronological and a comparative analysis. This opens discussions on whether alternatives to the traditional framework are workable. An alternative approach with focus on classification and automation would allow this. Literature reviews and discussions covering digital forensics should present options for a new approach whereby these are presented with confirmed case studies, providing case results are testable based on a framework. Consistent performance and confirmation would strengthen the new framework and enable automated data bank creation that would link to various other subfields of investigated processes. This would enable alternative studies of proper data analysis or fuzzy predictions. I suggest using a new approach into re-classification that is not bound to factual rigour, but rather focuses on occurrences and predictability.

Initially, on the first level this might seem disorganised, but while gathering data from experienced investigators we should be careful to prescribe methods that are not always best fit for resolving standardisation.

A complete new methodology that has a blueprint of procedures and processes in place might be able to build in an adaptive approach to the latest technological interventions/ developments. Doing this will hasten the processes of putting in place a global standardisation pattern which takes years to develop – just to be outdated again with a new development. However, once a standardised base is established, monitoring digital forensic investigations should become regulated in line with expectations and reality. Not only would systems and processes then be developed as required by the industry but also, as investigators suggest, that would support the impetus for Globalisation.

I came to the conclusion that these assumptions were difficult to confirm since participants showed varied interest in using alternative methods or frameworks either before or during conducting investigations.

References

(1998). KUMHO TIRE CO., LTD., et al. v. CARMICHAEL et al. CERTIORARI TO THE UNITED STATES COURT OF APPEALS FOR THE ELEVENTH CIRCUIT SUPREME COURT OF THE UNITED STATES. No. 97-1709. http://www.law.cornell.adu/supct/html/97-1709.ZS.html

Aanya-Isijola, A. (2009) *Models of Digital Forensic Investigation*. School of Computing & Technology. University of East London.

Ahmad, A. (2006). *The Forensic Chain-of-Evidence Model: Improving the process of Evidence Collection in Incident Handling Procedures.* Department of Information Systems. Parkville, VIC, University of Melbourne.

Ami-Narth, J.Y., Williams, P.A.H. (2008). Digital Forensics and the Legal System: A Dilemma of our times. Edith Cowan University.

Angelopoulou, O. (2007). *ID Theft: A Computer Forensics' Investigation Framework*, University of Glamorgan.

Armstrong, C. (2003). *Developing a Framework for Evaluating Computer Forensic Tools. Evaluation of Crime and Justice: Trends and Methods*, Canberra, Australian Bureau of Statistics.

Arnes, A., Haas, P., Vigna, G., and Kemmerer, R. A (2006). *Digital Forensic Reconstruction and the Virtual Security Testbed ViSe.* Department of Computer Science. Santa Barbara, CA, University of California Santa Barbara.

Baggili, I., Kiley, M. (2010). Digital Forensics: A Brief Overview of Critical Issues. Retrieved 04 June 2010 from http://www.forensicmag.com/articledigital-forensics-brief-overview-critical-issues?page=0,2

Barbara, E., Frincke, D.A., and Taylor, C.A. (2007). "A Theoretical Framework for Organizational Network Forensic Readiness." *Journal of Computers 2(3).*

Baryamureeba, V., Tushabe, F. (2004). *The Enhanced Digital Investigation Process Model*. Institute of Computer Science. Kampala, Uganda, Makerere University. http://www.dfrws.org/2004/day1/tushabe_EIDIP.pdf

Beckett, J., Slay, J. (2007). "*Digital Forensics: Validation and Verification in a Dynamic Work Environment.*" Proceedings of the 40th Annual Hawaii International Conference on System Sciences (HICSS'07).

Beebe, N., L. (2009). *Digital Forensics Research: The Good, the Bad, and the Unaddressed.* 5th Annual, IFIP, WG 11.9. The University of Texas at San Antonio. http://faculty.business.utsa.edu/nbeebe

Beebe, N. N., and Clark, J.G. (2005). A Hierarchical, Objectives-Based Framework for the Digital Investigations Process. Department of Information Systems and Technology Management. San Antonio. The University of Texas at San Antonio.

Brinson, A., Robinson, A., Rogers, M. (2006). *A cyber forensic ontology: Creating a new approach to studying cyber forensics*. Department of Computer & Information Technology. West Lafayette, Purdue University.

Broucek, V., Turner, P. (2005). *Developing a Conceptual Approach for Emerging Academic Discipline*. School of Information Systems. Hobart, University of Tasmania.

Broucek, V. T., P. (2004). *Computer Incident Investigations: e-forensic Insights on Evidence Acquisition*. EICAR Conference CD-Rom: Best Paper Proceedings, Copenhagen: EICAR e.V.

Broucek, V. T., P. (2006). *Winning the Battles, Losing the War? Rethinking Methodology for Forensic Computing Research*. School of Information Systems. University of Tasmania.

Buskirk, E. V., & Liu, V.T. (2006). "Digital Evidence: Challenging the Presumption of Reliability." *Journal of Digital Forensic Practice 1: 19-26*.

Carney, M., Rogers, M. (2004). "The Trojan made Me Do It: A First Step in Statistical Based Computer Forensics Event Reconstruction." *International Journal of Digital Evidence 2(4).*

Carrier, B. (2003). "Defining Digital Forensic Examination and Analysis Tools Using Abstraction Layers." *International Journal of Digital Evidence 1(4).*

Carrier, B., Spafford, E.H. (2006). "Categories of digital investigation Analysis techniques based on the computer history model." *Digital Investigation 3S: S121-S130.*

Carrier, B. D., and Spafford, E.H. (2004). *An Event-Based Digital Forensic Investigation Framework*. Center for Education and Research in Information Assurance and Security - Cerias. West Lafayette, IN 47907 USA, Purdue University.

Case, A., Cristina, A., Marziale, L., Richard, Golden G., Roussev, V. (2008). "FACE: Automated digital evidence discovery and correlation." *Digital Investigation 5: s65-s75*.

Casey, E. (2002). "Error, Uncertainty, and Loss in Digital Evidence." *International Journal of Digital Evidence 1(2).*

Casey, E. (2002). "Practical Approaches to Recovering Encrypted Digital Evidence." *International Journal of Digital Evidence 1(3).*

Cheong, K., W. (2006). *Analysis of hidden data in NTFS files system*, Edith Cowan University.

Ciardhuain, S. O. (2004). "An Extended Model of Cybercrime Investigations." *International Journal of Digital Evidence 3(1).*

Coursey, D. (2004) Security Hardware & IT Security Software Part Two: A Forensics Inquiry, Step by Step.

Elsaesser, C., Tanner, M. (2001). *Automated Diagnosis for computer forensics*. The Mitre Corporation.

Farmer, D., Venema, W. (1999). "Computer Forensics Analysis Class Handouts."

Fernandez, J. D., Smith, S., Garcia, M., Kar, D. (2005). *Computer Forensics - A critical need in computer science programs*, Texas A&M University - Corpus Christi.

Forrester, J., Irwin, B. (2007). *A Digital Forensic Investigation Model for Business Organisations*. Department of Computer Science. Grahamstown, Rhodes University.

Foster, M., Wilson, J, N. (2004). "Process Forensics: A Pilot Study on the Use of Check pointing Technology in Computer Forensics." *International Journal of Digital Evidence 3(1).*

Freiling, F. C., Schwittay, B. (2007). A common Process Model for Incident Response and Computer Forensics. *Proceedings of Conference on IT Incident Management and IT forensics, Germany.*

Gantz, J. F. (2007). A Forecast of Worldwide Information Growth Through 2010*, IDC - Analyze the Future: 24.*

Garfinkel, S., Farrell, P., Roussev, V., Dinolt, G. (2009). "Bringing science to digital forensics with standardized forensic corpora." *Science Direct 6(S2-SII).* www.dfrws.org/2009/proceedings/p2_garfinkel.pdf

Garfinkel, S. L. (2007). "*Forensic Corpora: A Challenge for Forensic Research."* http://www.simson.net/ref/2007/Forensic_corpora.pdf

GIAZ. (2010) GCFA Certification Bulletin. http://www.giac.org/certbulletin/gcfa.php

Gladyshev, P. (2004). *Formalising Event Reconstruction in Digital Investigations*. Department of Computer Science, University College Dublin.

Gluzinski, T., A., Kida, J. (2006). *"Managing your evidence. Problems associated with proper collection procedures."*

Grobler, C. P., Louwrens, C.P. (2009). *High-Level Integrated View of Digital Forensics*. ICSA, Johannesburg, University of Johannesburg. http://icasa.cs.up.ac.za/issa/2009/proceedings/full/28-paper.pdf

Haagman, D. (2007) Good Practice Guide for Computer-based Electronic Evidence. Association of Chief Police Officers (ACPO). http://www.7safe.com/electronic_evidence

Hall, G. A., Davis, W.P. (2005). “Towards Defining the Intersection of Forensics and Information Technology." *International Journal of Digital Evidence 4(1).* Texas State University. San Marcos.

Hannan, M., Turner, P. (2003). *Australian Forensic Computing Investigation Teams: Research on Competence*. 7th Pacific Asia Conference on I.S. 1-13 Adelaide, South Australia.

Hom-anek, P., Apiwathanokul, C., Nachin, N., Pamomchaisirikit, S. Sripeamlap, T. (2009) Career Opportunities and Development for Asia Information Security Professional with the IT Security Essential Body of Knowledge (EBK). http://www.tisa.or.th/downloads/6.TISA_TISET_Final_Presentation.pdf

Kahvedzic, D., Kechadi, T. (2009). "DIALOG: A framework for modelling, analysis and reuse of digital forensic knowledge." *Science Direct 6(S23-S33).*

Kenneally, E., E. (2005). Confluence of Digital Evidence and the Law: On the Forensic Soundness of Live-Remote *Digital Evidence Collection. U. J. L. Tech.5.*

Kennedy, I. (2006). "Presenting digital evidence in court." Retrieved 15 May 2010 from http://www.bcs.org/server.php?show=ConWebDoc.7372.

Kent, K., Chevalier, S., Grance, T., Dang, H. (2006). *Guide to Integrating Forensic Techniques into Incident Response.* Gaithersburg, MD 20899-8930 Computer Security Division, Information Technology Laboratory, National Institute of Standards and Technology. http://csrc.nist.gov/publications/nistpubs/800-86/SP800-86.pdf

King, G., L. (2006). *"Forensic Plan Guid*e." *SANS Institute.*

Kissel, R., Stine,K., Scholl,M., Rossman, H., Fahlsing,J., Gulick, J. (2008). "*Security Considerations in the System Development Life Cycle.*" Computer Security Division. Information Technology Laboratory, National Institute of Standards and Technology, Gaithersburg, MD 20899-8930. http://csrc.nist.gov/publications/nistpubs/800-64-Rev2/SP800-64_Revision2.pdf

Kohn, M., Eloff, JHP., Olivier, MS. (2006). *Framework for a Digital Forensic Investigation.* Department of Computer Science, Information and Computer Security Architectures Research Group (ICSA). Pretoria, University of Pretoria. http://mo.co.za/open/dfframe.pdf

Kohn, M., Eloff, JHP., Olivier, MS. (2007). *Framework for a Digital Forensic Investigation.* Department of Computer Science, Information and Computer Security Architectures Research Group (ICSA). Pretoria, University of Pretoria.

Kohn, M., Eloff, JHP., Olivier, MS. (2008). *UML Modelling of Digital Forensic Process Models* (DFPM's). Department of Computer Science. Pretoria, University of South Africa.

Kornblum, J. (2006). *Identifying Almost Identical Files Using Context Triggered Piecewise Hashing*. DFRWS.

Kovar, D. (2009) *Guideline for EnCase workflow*. Forensic Focus. http://www.forensicfocus.com/index.php?name=forums&file=viewtopic&t=4975

Kovar, D. (2009) *Push button forensics – managing the downsides*. http://inegriography.wordpress.com/?s=button+forensics

Kumho Tire Co., Ltd. v. Carmichael, 119 S. Ct. 1167 (1999) 57

Kruse, W, G., Heiser,. (2001) *Computer Forensics: Incident Response Essentials.*

Landman, J. (2002). "*Forensic Computing: An Introduction to the Principles and the practical applications*". http://www.scm.uws.edu.au/computerforensics/Online%20Materials/FC.pdf.

Lee HC, Palmbach TM, Miller MT. *Henry Lee's crime scene handbook*. San Diego: Academic Press; 2001.

Leigland, R., Krings, A.W. (2004). "A Formalization of Digital Forensics." *International Journal of Digital Evidence 3(3).* www.ijde.org

Leong, R., S.C (2006). "FORZA – Digital forensics investigation framework that incorporate legal issues." *Digital Investigation 3S: S29-36.*

Lucas, J., Moeller, B. (2003). *Computer Forensics: An Evolving Discipline*, Addison Wesley Professional.

Meyers, M., Rogers, M (2004). "Computer Forensics: The Need for Standardisation and Certification." *International Journal of Digital Evidence 3(2).*

Meyers, M., Rogers, M. (2005). *Forensic Examination of Digital Evidence: A Guide for Law Enforcement Expert witness training*. Computer Forensics: The Need for Standardization and Certification.

Mocas, S. (2004). "Building Theoretical Underpinnings for Digital Forensic Research." *Digital Investigation 1: 61-68*. Original publication by Portland State University. http://220.231.93.23:8000/collect/EN-digital/index/assoc/HASH01cc.dir/1c%282%29.pdf

Noblett, M. G., Pollitt, M.M. and Presley, L.A. (2000). *Cyber Forensics. A field manual for collecting, examining, and preserving Evidence of Computer Crime*, Auerbach Publications, Boca Raton.

Nolan, R., Baker, M., Branson, J., Hammerstein, J., Rush, K., Waits, C., Schweinsberg, E. (2005). *First Responders Guide to Computer Forensics: Advanced Topics.* c. Mellon. Pittsburgh, Software Engineering Institute.

Olivier, M., S. (2009). *On Metadata context in Database Forensics*. ICSA Research Group, Computer Science, Pretoria, South Africa, University of Pretoria.

Palmer, G. (2001). *A Road Map for Digital Forensic Research. DFRWS Technical Report*. Utica, New York, AFRL/IFGB Air Force Research Laboratory.

Palmer, G. (2001). DFRWS, Report from the first digital forensic research workshop. A roadmap for digital forensic research. The MITRE Corporation. Report DTR-T001-01 http://dfrws.org

Pan, L., Batten, L,M. (2005). *Reproducibility of Digital Evidence in Forensic Investigations*. School of Information Technology. Burwood, Victoria, Deakin University.

Perumal , S. (2009) *Digital Forensic Model Based On Malaysian Investigation Process*. IJCSNS International Journal of Computer Science and Network 38 Security, VOL.9 No.8, August 2009. Faculty Of Science & Technology Islamic Science University Of Malaysia

Petherick, W. (2002). "*Criminal profiling: How it got started and how it is used*." http://www.crimelibrary.com/criminology/criminalprofiling2

Radack, S. (2009) *Forensic Techniques: Helping Organizations Improve Their Responses To Information Security Incidents.* Computer Security Division, Information Technology

Laboratory, National Institute of Standards and Technology. Retrieved on 6 February 2010 from http://csrc.nist.gov/publications/nistpubs/800-64-Rev2/SP800-64-Revision2.pdf

Reith, M., Carr, C., Gunsch, G. (2002). "An Examination of Digital Forensic Models." *International Journal of Digital Evidence 1(3).* Retrieved on 4 December 2009 from http://www.utica.edu/academic/institutes/ecii/ijde/articles.cfm?action=issue&id=3

Rogers, M. (2003). *The role of criminal profiling in the computer forensics process*. Centre for Education and Research in Information Assurance and Security (CERIAS). Purdue University. www2.tech.purdue/cit/course/cit556/readings/profile-rogers.pdf

Rogers, M., K., Goldman, J., Mislan, R., Wedge, T. (2006). *Computer Forensics Field Triage Process Model*. Conference on Digital Forensic, Security and Law.

Roussev, V., Richard III, G. G. (2004). *Breaking the Performance Wall: The case for distributed digital forensics*. Department of Computer Science. New Orleans, University of New Orleans.

Rowlingson, R. (2004). "A Ten step Process for Forensic Readiness." *International Journal of Digital Evidence 2(3).*

Ruibin, G., Gaertner, M. (2005). "Case-Relevance Information Investigation: Binding Computer Intelligence to Current Computer Forensic Framework." *International Journal of Digital Evidence 4(1).*

Ryan, D. J., Shpantzer, G. (2005). "*Legal Aspects of Digital Forensics*." The George Washington University. Washington, D. C.

SANS (2010) Focus: Learning how to discover new artefacts using application forensics. Retrieved on 04 June 2010 from http://www.sans.org/selfstudy/description.php?cid=13822

Sansurooah, K. (2006). *Taxonomy of computer forensics methodologies and procedures for digital evidence seizure*. School of Computer and Information Sciences (SCIS). Perth, Edith Cowan University Perth.

Schwittay, B. (2006). *Towards Automating Analysis in Computer Forensics*. Department of Computer Science, RWTH Aachen University Diploma Thesis in Computer Science.

Seagate. (2008) *Seagate Powers Next Generation of Computing with three new Hard Drives*. Retrieved on 10 May 2010 from http://www.seagate.com

Selamat, S., R., Yusof, R., Sahib, S. (2008). *Mapping Process of Digital Forensic Investigation Framework*. Faculty of Information Technology and Communication. Melaka, Malaysia, Universiti Teknikal Malaysia. http://paper.ijcsns.org/07_book/200810/20081025.pdf

Smith, F., C. (2002). *A guide to Forensic Testimony: The Art and Practice of Presenting Testimony as an Expert*, Addison Wesley Professional.

Spafford, E. H., Carrier, B.D. (2001). "A Road Map for Digital Forensic Research 2001." *Digital Forensics Workshop.*

Stallard, T., Levitt, K. (2003). *Automated Analysis for Digital Forensic Science: Semantic Integrity Checking*. Department of Computer Science. Davis, University of California. http://www.asac.org/2003/papers/89.pdf

Stephenson, P. (2003). *A Comprehensive Approach to Digital Incident Investigation*. Elsevier Information Security Technical Report.

Thuen, C. (2007). *Understanding Counter-Forensics to Ensure a Successful Investigation*, University of Idaho.

Tobin, W. A., Thompson, W.C. Evaluating and Challenging Forensic Identification Evidence. Journal, *National Association of Criminal Defense Lawyers (NACDL*).

United States, National, Institute. (2001). *Electronic Crime Scene Investigation: a guide for first responders.* United States National Institute of Justice Technical Working Group for Electronic Crime Scene Investigation.

Velasco, J. (2007). *A guide to Electronic Evidence Collection Methodologies, Renew Data*.

Villars, R.L., Reinsel, D., Woo, B. *Worldwide Storage 2009 – Top 10 Predictions: Grappling with Content Growth in a Contracting Economy*. http://www.idc.com

Waits, C., Akinyele, J,A., Rogers, L. (2008). "*Computer Forensics: Results of Live Response Inquiry vs. Memory image Analysis.*" Software Engineering Institute. Electronic resource number, CMU/SEI-2008-TN-017.

Walker, C. (2005). "*Computer Forensics: Bringing the Evidence to Court*."

Warren, G., Kruse, II., Heiser, J,G,. (2001). *Computer Forensics: Incident Response Essentials*, Addison-Wesley. ISBN, 13:9780201707199.

Weise, J., Powell, B. (2005). "*Using Computer Forensics When Investigating System Attacks*." Sun Microsystems.

Whitcomb, C. M. (2002). "An Historical Perspective of Digital Evidence: A Forensic Scientist's View." *International Journal of Digital Evidence 1(1).* https://www.utica.edu/academic/institutes/ecii/publicatiojns/articles/9C4E695B-0B78-1059-3432402909E27BB4.pdf

Wilson, S. (2009). "*7 Safe Information Security. Good practice guide for Computer-based digital evidence.*" http://www.7safe.com/electronic_evidence/ACPO_guidelines_computer_evidence.pdf

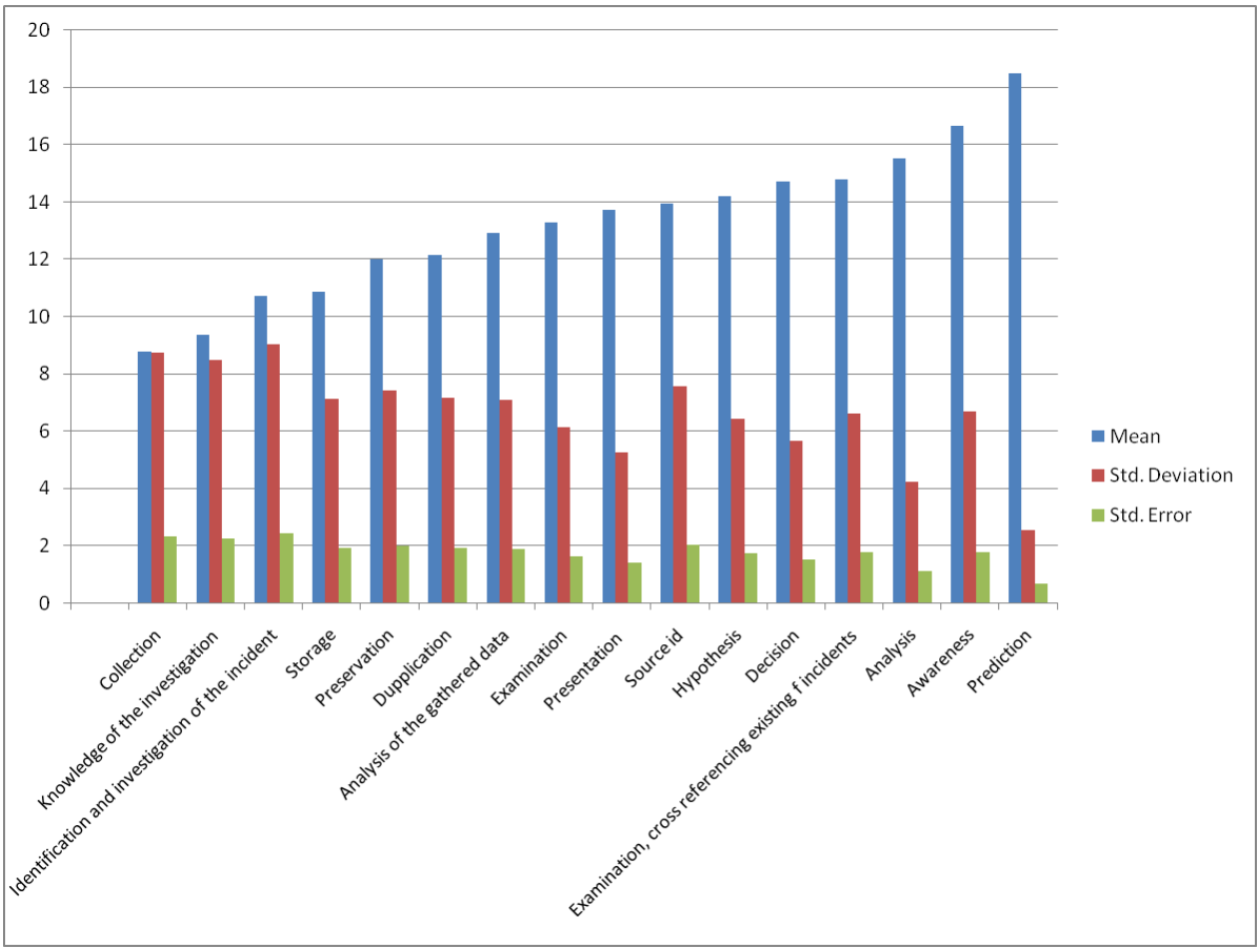
Wright, T. E. (2001). *The Field Guide for Investigating Computer Crime,* Part Eight: Information Discovery - Searching and Processing. R. Lemos, SecurityFocus. http://www.security focus.com/infocus/1245

Yasinsac, A., Erbacher, R.F., Marks, D.G., Pollitt, M.M., and Sommer, P.M. (2003). "*Computer Forensic Education.*" IEEE Computer Society 1540-7993(03). http://computer.org/secuirty/

Appendix A

Table Data from Section B

The sequence of the Existing investigation processes



A low score is more important or occurs first in the investigation process

Data from Section A

Present traditional framework and suggestions

Question 1.

Do you *use an explicitly described framework or model* in your investigation process? (Please circle or delete)

Response: Yes: 4

No: 9

Summary: Nine (9) participants indicated that they are not using a framework, one had no response and four (4) said YES. Although nine (9) said NO, they indicated that they use one or the other kind of procedure.

When we compare this to the question covering processes used, we note that the same participants never had clear steps or no steps at all in their investigation processes.

	Yes/No		Commend
Participant A		no	Every investigation is different and though some of the initial processes are always conducted, many of the process are specific to the type of investigation. Such as an Intellectual property matter you wouldn't worry about recovering picture files from unallocated space. These days drives are so big and often work is urgent, so you focus on finding the evidence based on a predetermined brief, hence a different model is required for each different type of job
Participant B		no	No comment
Participant C		no	We use an SOP to begin each case. Then move on to a specific model based on the type of examination.
Participant D		no	
Participant E		no	
Participant F	yes		We work to internationally accepted methodologies as published by numerous Institutes (DOJ, NIST, ACPO, SANS). We follow internal process models which are consistent across our organization globally. We follow the guidelines and methodologies published by the manufacturers / developers of the tools / applications we use.
Participant G		no	The investigation differs from case to case hence there is no definite model that follows an investigation. However, I follow closely the EDRM model on each of the investigations I work on.
Participant H	yes		We use DFF (Digital Forensic Framework) our own product as forensic software; we also have developed our own process for investigation.
Participant I		no	I do not see a framework where I currently work, however I know that frameworks, or best practices are employed. Usually these best practices are manifested in the preparation or triage of a system, not in the examination itself.
Participant J		no	The work I do is focused on providing specific answers to clients. During
Participant K	yes		A-E-A-R Model

Participant L		Not Completed	No commend
Participant M		no	No commend
Participant N	yes		No commend

Question 2.

Do you use any known theory as guideline before a case investigation is conducted? (Please circle or delete)

Response: Yes: 4

No: 9

Summary: Nine (9) participants indicated that they are not using a guideline. Four (4) said YES, they are using a guideline before they conduct their investigation. One (1) participant did not complete or answer the question.

From commends by Participant F, it is suggested that investigators might steer away from template use, since this might potentially lead them into pre-expected investigation modeling – which ultimately leads investigators away from conducting investigations with no pre-conceived expectations.

	Yes/No	Commend
Participant A	no	Again as above, if you specifically conducted the same type of investigations all the time then you may have a set theory and model. The Department of Internal Affairs for instance only do Child Pornography cases and nothing else, hence they could derive a single model that would work for most of their cases.
Participant B	no	No commend
Participant C	no	No commend
Participant D	no	I use a proven and internationally accepted mythology which has been discussed and adopted by the world's leading forensic consultants. This mythology is used in many qualifications that are set by the various software manufacturers.
Participant E	no	No commend.
Participant F	no	To apply a theory before a case can lead to templating of the analysis and missing potential data or out of the blue investigative leads
Participant G	no	No commend
Participant H	yes	No commend
Participant I	Not	No commend

		completed	
Participant J		no	
Participant K	yes		Lots in Law
Participant L	yes		No commend
Participant M		no	No commend
Participant N	yes		Uses in FBI acquisition guidelines

Question 3.

Do you use previous investigation data as guideline to conduct new investigations?

Response: Yes: 8

No: 6

Summary: It appears that the majority of participants use previous investigations as guideline. From participant D we note that "case initialization" might closely match to a template perspective as starting point before investigations are conducted.

Participant F is affirmative about using a methodology from previous investigations and applying this to new investigations, this reflects on using old data based on experience in similar environments and accessing the new case on the same models. In Section B, question 1, this participant answered also YES to the question:"Do you think that having a database of previous cases (corpora) can help in digital forensic investigation and case analysis"

Surprisingly, to the question: "Do you think an automated digital forensic investigative process is feasible?" the same participants answered NO, and mentioned that by automating the process you will remove the ability to change the process to address case specific exceptions.

	Yes/No	Commend	
Participant A		no	Not so much the data but you learn from each case and you known when you get a case that it may fit a similar scenario to a previous case and hence when you complete the report at the end, you may use the prior case's report as a template. The data from each case is always kept separate and never on the same computer at the same time to avoid contamination or confusion.
Participant B	yes		At times, I refer back to older cases for assistance with current case. An example would be a drug offense. I might look back on an older case for keywords, etc.
Participant C		no	No commend
Participant D		no	Each investigation is different and data from previous investigations would not be the same. The initial case initialization is the same for each case but after that, each case is different because of the individual

			requirements of each case.
Participant E		no	No commend
Participant F	yes		Always – your ability to operate as an effective analyst is based to some part on training but primarily on experience. Applying knowledge gained from previous work is how process and methodology are developed and refined to become more efficient.
Participant G	yes		If the previous data is related to the new investigation, then the previous data is used as part of the investigation not necessarily as guideline
Participant H	yes		We can't really use previous 'data' because the data must be destroyed after a case but we can reuse some specially developed tools or process
Participant I		Not completed	This would depend on the type of investigation. If there were multiple drives or multiple forms of media all related to the same case, an examiner would be remise in not using the previously discovered data
Participant J	yes		This depends. If the case is related, then (at least) some of the questions I attempt to answer will be focused on looking for similarities between the cases. Other than that, I may use knowledge from previous cases to help avoid (or create) specific situations. However, I don't use previous investigation data as a hard and fast rule.
Participant K	yes		Part of learning
Participant L	yes		Depends on cases
Participant M		no	No commend
Participant N	yes		No commend

Question 4.

Is there any specific part of the investigation that is more decisive than the other in the forensic investigation?

Response: Yes: 6

No: 6

Summary: An equal distribution of YES and NO, this reflects participants are not grouping their investigations methods as such, but rather try to initiate their investigations based on experience suited for the particular case, irrespective of prescribed methods.

	Yes/No	Commend	
Participant A		no	Again depends on the brief. If you can find facts as opposed to having to use opinion evidence it will always carries more weight in court. But often a mix of both is needed. The critical part is collection in accordance with the standards as if you don't

			collect the evidence correctly you are stuffed from the start.
Participant B	yes		This depends on the case. At times, the search warrant provides a wealth of information. At other times, the evaluation of all the data retrieved will be the deciding factor for case outcome.
Participant C	yes		Pre-investigation of non-digital evidence is critical to conducting a good computer forensics investigation. You have to know what you are looking for before starting to look; otherwise you will be wasting your time on a fishing expedition.
Participant D	yes		The initial Evidential Imaging process is probably more important than any other part of the case. If this is conducted correctly, the entire integrity of the case may be compromised.
Participant E		no	I don't understand the question.
Participant F		no	No commend
Participant G		Not completed	No commend
Participant H		no	We must ensure that all part of the investigation are done with the same quality but most of time the search and correlation is a decisive part
Participant I	yes		There are absolutes in most examinations, such pieces of data as the registry analysis of a Windows machine provides definitive answers as to owner, software etc.
Participant J		no	This depends on the questions you are trying to answer with the investigation.
Participant K	yes		Authentication for acquiring
Participant L			No commend
Participant M		no	No commend
Participant N	yes		Chain of custody and preservation

Question 5.

Please rate on a scale of 1-5 (1=not at all sufficient, 5= totally sufficient) whether you think current investigative processes in digital forensics are sufficient for addressing most aspects of the investigation.

Response: Yes: 8

No: 6

Summary: Data collected that most participants seven (7) are fairly satisfied with the existing investigation process. Two (2) participants are not satisfied and only three (3) are very satisfied with the existing investigation model.

From participant B, we note a good correlation between their answer for this question, which indicates no changes have to take place to the existing investigation process - and the question in Section A, question 6. In Section A, question 6:" Are there any aspects of current investigative processes that you would like to see changed?" participant B also answered NO.

A noticeable score of seven (7) participants' shows only 3 out 5, indicating they are either satisfied or borderline to unsatisfied with the investigation processes.

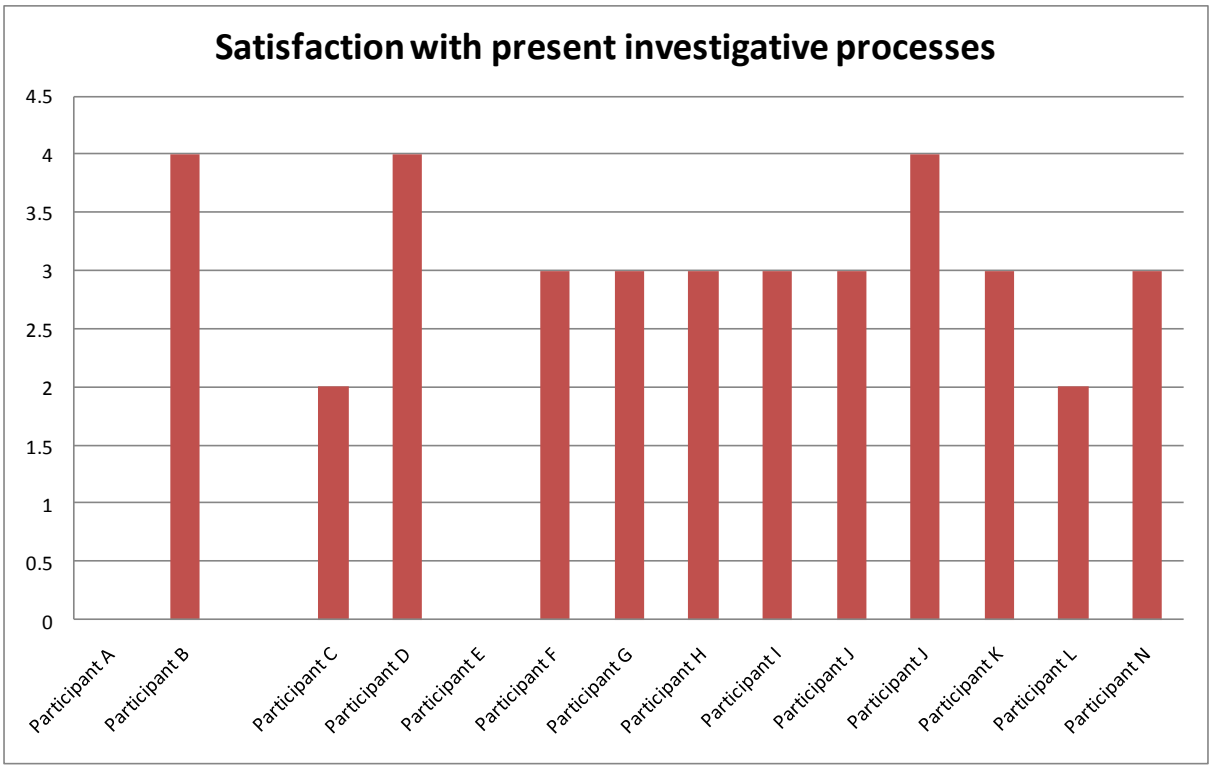
Scale: 1 2 3 4 5

Participant A		No c	20
Participant B		1 2 3 (4) 5	4
Participant C		1 (2) 3 4 5	2
Participant D		1 2 3 (4) 5	4
Participant E	Every job is different. The process used is dynamic and is not duplicated between jobs		20
Participant F		3	3
Participant G		3	3
Participant H		3	3
Participant I		3	3
Participant J		3	3
Participant J		4	4
Participant K		3	3
Participant L		2	2

Participant N		3	3

Satisfaction with present investigative processes.

A higher score indicate satisfaction with the existing forensic framework.



A higher score indicate satisfaction with the existing Investigation Processes

Question 6.

Are there any aspects of current investigative processes that you would like to see changed?

Response: Yes: 5

No: 4

Not Sure: 2

Summary: From Participant H we note a re-curing indicator for investigation automation. If changes are to be made, this participant clearly indicates in Section B, question 2 and 4, that databank creation and the type of data collected might help in investigation processes. However Participant H also points to the importance that automation might lose information if the basis of the platform is not correctly aligned to the type of investigation to be conducted.

One particular drawback is to get information that can be used for databank creation, as stated earlier in this thesis. This is also voiced in Participants H's commends in section B, question 5.

	Yes/No		Commend
Participant A		no	You are usually bound by the law of evidence as to what is considered valuable as if it isn't strong evidence then it may be considered irrelevant and if it is irrelevant it cannot be used in evidence. Usually an investigation has one objective to prove someone as done something wrong civilly or criminally and an investigator has the job of finding evidence to prove the case in accordance with the law.
Participant B		no	
Participant C	yes		The current state of examiner expertise is far below what is believed in general. More training is needed in advanced forensic analysis as well as in the investigative domain of knowledge.
Participant D		no	As every investigation is a one off and they are all different, each investigative process is individual to that case.
Participant E		Did not answer	Every job is different. The process used is dynamic and is not duplicated between jobs. I change the process to meet the needs of the investigation. Often I am restricted by the budget of the client.
Participant F		No commend	No commend
Participant G		No commend	No commend
Participant H	yes		Any automation process will be a great help, because nowadays some investigation can take very long time because of the huge amount of data. The drawback of automation is that if it's not perfect you can lose some evidence.
Participant I		Not sure	One of the issues that I think is being overlooked here is that each agency that conducts forensic examinations does so under

			their own principles, best practices and guidelines. Although I have seen some areas of the forensic process that could be more efficient. For instance, certain labs take in excess of 18 months for a single exam. There are many factors that contribute to this time frame, and some of those factors can be adjusted.
Participant J	yes		I'm not sure what you mean by "current investigative processes". A framework, versus a methodology for investigating a specific type of crime? The issue with most "frameworks" that I've seen is that they try to combine the investigation (related to criminal activity or violation of corporate policy) with the forensic examination (looking at the evidence to find information to help reach a conclusion).
Participant K		no	We build our own framework with regards to the evidence out of collection and practice???
Participant L	yes		No commend
Participant M	yes		No commend
Participant N		Not sure	The costs involved of E-discovery and related investigations are prohibitively expensive. Review time of data in expensive and with automation it could should be reduced.

Question 7.

Do you think the type of software you use to conduct forensic IT investigations plays a decisive role in extracting evidence that is critical to your investigation?

Response: Yes: 12

No: 1

Summary: In this question we note indicators that software plays a major role when investigations are conducted. When we compare this to the section Software Used, question 1, we note that participants prefer three types of software packages above the rest. Participants B,C and G states using Encase – as we saw from previous data in the thesis, while this is a preferred software package, it does not allow for customized script creation.

This makes customized investigations difficult and closely math earlier reference to “button investigators”. Thus reiterating the inability of investigators to comprehend the full spectrum of the investigation and only make use of the software packet features instead of scripting their own commands in a customizable manner to gather specific data. Only Participant H and J state using customizable script to access data.

	Yes/No		Commend
Participant A	yes		It doesn't play a decisive role as what you find does as in the evidence does that. The forensic software these days just automates many of the processes and makes it easier to work on a case. You still have to be able to interpret the results and prove how or why something happened.
Participant B	yes		I use both FTK and EnCase. I like FTK's indexing feature. I use FTK mainly to do cases which contain a lot of e-mail or cases where I need to do a lot of keyword searches. FTK's graphic feature does not work well. I like EnCase for investigations such as child pornography cases or any type of case that may contain a large number of images or videos. EnCase's keyword searching is a bit cumbersome vs. FTK's.
Participant C	yes		I use Encase 6.13 by Guidance Software. In the hands of a trained and experience investigator, pretty much anything that can be found, can be found.
Participant D	yes		If the software cannot be relied on to do the job it is designed to do, then your investigation will be flawed from the start. Individual testing of the software is imperative to the success of an investigation and knowing that the results are correct. Redundancy checks need to be also completed using comparative software.
Participant E	yes		No commend
Participant F	yes		No commend
Participant G		no	Encase – It provides a way to extract bit-to-bit image. Knoppix – Faster and less intrusive way to image hard drive. DtSearch – Scabale searching tool but limited reporting methodology.
Participant H	yes		We choose to develop our own solutions as an open source framework because we needed some features that other software can't provide. Also most of forensic software is old we choose to develop it with new technologies and special design which can permit to develop script very easily to face different cases.
Participant I	yes		I cannot answer this question the way in which it is phrased. An examination is a scientific process, cyclical in nature. There is no straight forward list to follow. However, I have provided a few of the main tools used.
Participant J	yes		I use a combination of open source, free, commercial and private/custom software. The choice of tool is not based on if the tool provides "decisive" evidence, but if the tool provides the answer, and how likely the tool is to provide the correct answer.
Participant K	yes		Impossible in that??
Participant L			No commend
Participant M	yes		No commend

Participant N	yes		No commend

Software Used

Order of investigation Software

Summary: Three software packages emerged as being the most often used in the investigation process. Encase, FTK Manager followed by EncaseFTK was the top choice.

Surprisingly, Helix and other open Linux based software, for instance Unbuntu was not high in demand.

Question 1.

Participant A

Order of investigation	Software
None specified	1. Encase or FTK imager or similar 2. Encase or FTK to conduct initial examination 3. May then need email examiner 4. Registry Viewer 5. Net Analysis 6. Password cracker 7. Scripts that then run within software like encase that recover data or parse it so it can be interpreted

Participant B

Order of investigation	Software
1. Request for Forensic Analysis	
2. Log In Evidence	
3. Create E0 files (Acquire Drive)	FTK Imager
4. Examine Image	FTK or EnCase

Participant C

Order of investigation	Software
Copy of disk	FTK Imager, Helix, Linen, Encase
Investigation	Encase

Participant D

Order of investigation	Software
1. Evidential Image	EnCase
2.	FTK Imager
3. Case Examination	Encase
4.	FTK
5.	DataLifter
6.	NetAnalysis
7.	CacheBack Access Data Registry Viewer Secret Explorer CD/DVD Inspector Beyond Compare Thumbs Plus

Participant E

Order of investigation	Software
1. Acquisition of data	Tableau TD1
2. Examination of image files	EnCase, FTK, Net Analysis, VM's etc
3. Reporting	MS Word

Participant F

Order of investigation	Software
1. preserve	EnCase, FTK, DD etc, obtain forensic copy of data
2. extract	EnCase, FTK, Intella, numerous other tools dependant on data type
3. examine	As above
4. review	" also DT Search, Summation etc
5. report	Dependant on data type but would include office & applications as well as Forensic tool output.

Participant G

Order of investigation	Software
1. Easy Recovery	Recover deleted Messages
2. Paraben	Convert to a standardized mail format
3. DtSearch	Searching and exporting resultset

Participant H

Did not answer

Participant I

Order of investigation	Software
	1. Helix
	2. EnCase
	3. UTK (Access Data)
	4. Custom applications
	5. X-Ways
	6. FTK Imager

Participant J

Order of investigation	Software
-	Helix, Ubuntu
-	Sleuthkit / Autopsy
-	LibForensics
-	LibForensics
-	Various PERL scripts
-	IDA Pro

Participant K

Order of investigation	Software
	Encase
	Ftk
	Prtk

Participant L

Did not answer

Participant M

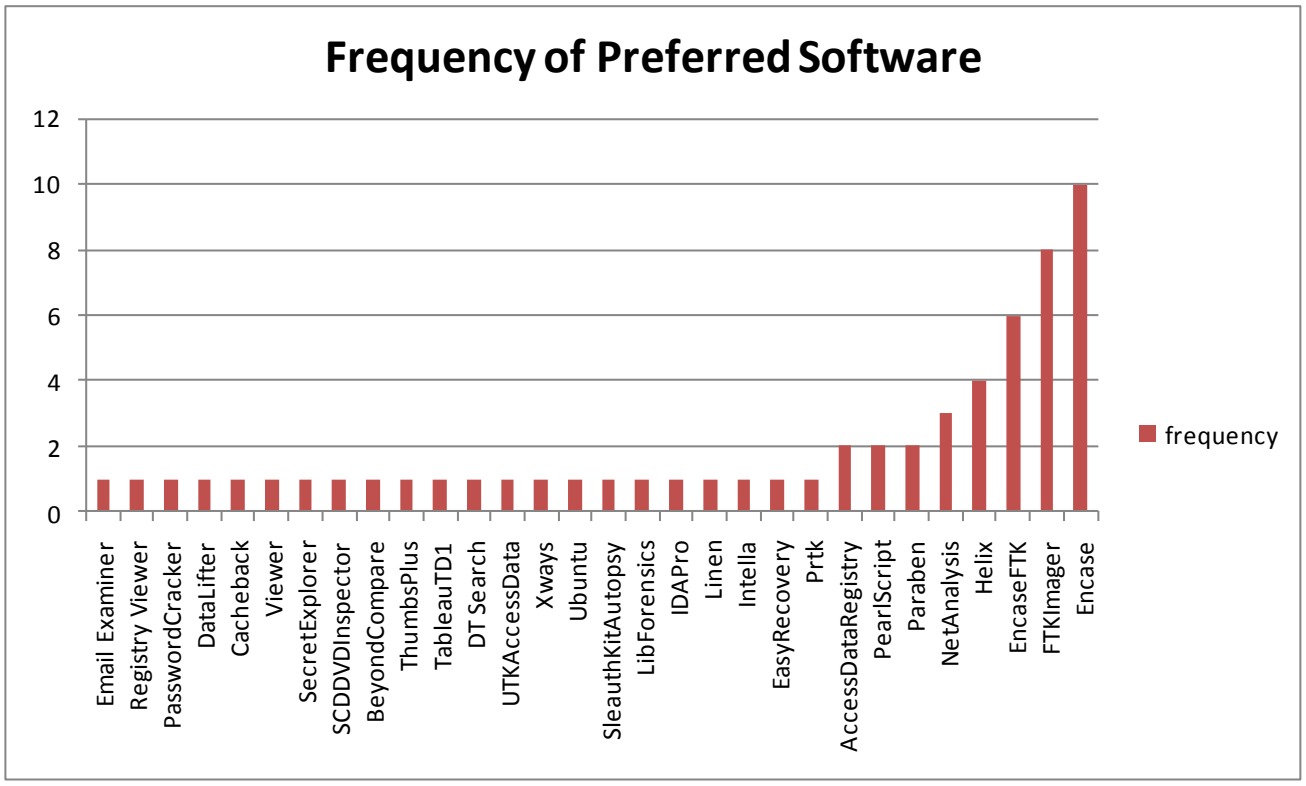
Did not answer

Participant N

Order of investigation	Software
1. Create DD image	AccessData Ftk Imager
2. Case Created	AccessData Ftk Imager
3. If further investigation is required	Guidance Software – Encase
4. Acquisition of portable devices	Paraben Software – Acquisition Toolkit
5. For Linux and Macintosh platform acquisition	Helix

Frequency of Preferred Software

A higher score indicate the frequency of a particular software packet used.



Data from Section B

From **Section B** we expected participation that related to the existences of a data bank and the participants'' expected contribution to such a database.

Question 1.

Do you think that having a database of previous cases (corpora) can help in digital forensic investigation and case analysis?

Response: Yes: 9

No: 4

Summary: Participant C points to a valid concern in data base creation. If all the reference cold be correctly build in the initial data based design, we could present a relational data structure that would allow for the searches to be stepped into more layers. Privacy issues and secure storage seems to be keeping most investigators form contributing to such a database.

	Yes/No	Commend	
Participant A		No	Every country is different, the laws are different the language, the date formats. You would have to have one that is country specific and outside of law enforcement no one is ever going to share information about cases as it is bad commercially and also you never know when a fellow forensic investigator is going to be on the other side of a case giving evidence against yours.
Participant B	yes		I currently keep my own database (using Microsoft Access). It is helpful at times to be able to find archived cases and refer back to them in assistance with current cases.
Participant C	yes		Only if such information creates a searchable knowledge base of specific techniques used to solve investigative challenges. i.e. locating some type of data and the technique used.
Participant D		no	The reason I state this, is because each case is and individual case. No two cases are the same and no two cases are looking for the same thing. Each case has individual requirements as requested by the contractor.
Participant E		no	Every job is different
Participant F	yes		Always – your ability to operate as an effective analyst is based to some part on training but primarily on experience. Applying knowledge gained from previous work is how process and methodology are developed and refined to become more efficient
Participant G	Yes no		Haven't worked on corpora
Participant H	yes		It can help to develop new tools or verify some existing tools but I thought it will be hard to have real cases data as they must be protected or destroyed most of the time

Participant I		unsure	A database would be useful in certain instances; however, there are many factors that would have to be addressed first. These factors include staffing to maintain the DB, whether or not this information would be accessible or shared with other entities or remain in house.
Participant J	yes		This could allow the examiner to learn from others. Similar situations and/or experiences can provide insight into the current examination.
Participant K		no	privacy
Participant L	yes		
Participant M	yes		A lot of cases are similar??
Participant N	yes		

Question 2.

What kind of data would you regard as important to have in the data bank? Please indicate in the right column or add additional attributes or clusters of associated types of data as you would require. (I.e. would knowledge of a criminal's behaviour show a potential pattern when you consider their educational background, capacity in storage, type of operating system they used?)

Summary: participants are unsure about the content of such a database – and concerns about the rapidly changing environment and case data makes this even more difficult.

Question 2 commends:

	Yes/No		Commend
Participant A			Only a very small portion of cases are criminal, most now days are civil as in corporate, employment, family and many other matters. Hence there are so many people in private enterprise doing computer forensics in the private sector it is all about chasing the dollar. The equipment is expensive the software is expensive and so is the training and staff costs. Hence you have to maximize the cases you can do to cover the ever increasing overheads. Most criminal work is very narrowly focused and if anything experts are doing work for the defence in a case and not for the Police.
Participant B			
Participant C			What's missing is examination specific type data. i.e. Location of chat artifacts, search strings for a particular type of chat, i.e. yahoo where the data must be recovered from unallocated space or another location.
Participant D			I have not completed this section as I do not consider that a database will help with an investigation. Education of the latest methods used by offenders and the latest aspects of the Operating Systems is more important, as if you don't know how the OS behaves then you are up against it from the start.

			Previous criminal behaviour/history is not relevant as their systems are always changing.
Participant E	yes		I preserve data, but do not conduct analysis for trends etc.
Participant F		no	Limited to data volume, date and time, and type of offending.
Participant G	yes		Hard drives, exchange, network shares, PDAs, thumb drives.
Participant H			In fact a databases is very important when hardware are very different like it's for cell phone today as we developed modules for cell phone analysis we understood very rapidly that we need such sort of database. Also for most common hardware it's less important. For HD it could be a database describing constructor zone because it changes from different manufacturer and model but at the contrary to know about mother board or network card difference will not permit to have a huge gain of information.
Participant I			
Participant J			

Question 3.

Do you collect data of previous cases currently?

Response: Yes: 6

No: 6

Summary: Six (6) participants collect data, but only Participant C use the data to test examination techniques.

	Yes/No		Commend
Participant A			
Participant B	yes		I do not keep them itemized, per se, but I do keep my reports and post my stats on a quarterly basis.
Participant C	yes		We collect information about examination techniques in an internal wiki.
Participant D		no	The only form of collection is my working files and final report. The evidential images are not kept, due to storage limitations. I do not refer back to any previous report, except to maybe use one as a template.
Participant E	yes		I preserve data, but do not conduct analysis for trends etc.
Participant F		no	Limited to data volume, date and time, and type of offending.
Participant G	yes		Hard drives, exchange, network shares, PDAs, thumb drives.
Participant H		no	We can't really collect data of previous cases as most of them must be 'destroyed' after the cases is finish
Participant I		No	I am not in a position to do this

Participant J	yes		As reference material.
Participant K			Not sure what this means – all work is controlled and accounted to the rules of the court. There is a prescription legal framework that defines what we may do and how??
Participant L		no	
Participant M	yes		
Participant N		no	

Question 4.

Do you think an automated digital forensic investigative process is feasible?

Response: Yes: 6

No: 8

Summary: From this result we note a strong response for and against automated procedures. Most concerns are that no investigation is ever the same as before and automated procedures might miss unique or complex associations which only an investigator can interpret. This is arguable; since human investigators might not always for associations with similar cases in the databank where core values of relations exists.

However some participants presented a clear notion towards automation, indicating the importance of having a basis of initial investigation processes. This notion is similar to the author's perspective that a tiered approach from initial level investigation leads to rationalization of redundant data, thereby allowing more detailed assessment of the factors at hand unique to the particular investigation scenario.

	Yes/No	Commend	
Participant A		No	Never. It is the evidence that is located that is the important aspect and how it is interpreted. You can have tools automatically extract data for you. But you have to interpret it and then prove it in court which is the MOST important part of the entire process. That the whole meaning of the word FORENSIC.
Participant B		no	
Participant C		no	Forensic automation is already becoming a problem by giving untrained examiners a false sense of security when in reality; they are not conducting an examination at all. When used properly some automation is good. However, it is not to the point where any time in the near future, an automated tool can conduct a thorough enough examination to be trustworthy.
Participant D	Yes/no		Both yes and no. Mainly NO from the point of view that each case is different. Some people might like a system to tell them what to do and when to do it, but that tells me that the person most probably has no formal Computer Forensic Training and is what we call a Point and Click consultant. Someone who has no Training in Forensics and or Police Investigative processes. You still have to have the tenacity to sit there for hours analyzing the recovered data.

Participant E	yes		This may be something for people with no investigative mindset. Good detectives and investigators do not operate by following a checklist. Real life does not work this way. The investigator needs to be able to “think outside the square”.
Participant F		no	It may be applicable for certain processes or activities (portions of cases), but I feel an automated complete examination is unfeasible. By automating the process you will remove the ability to change the process to address case specific exceptions.
Participant G	yes		It would ensure that the investigator is performing the tasks in a forensic manner and also ensure that none of the steps are missing in between. Also if a knowledge bank gets created along the way that will be helpful troubleshooting previous encountered problems.
Participant H	yes		We really think it's feasible for some part of the investigation but maybe all the analysis can't be automated. When we developed DFF we thought about automate digital forensic but we didn't include any automation yet, simply because lacks of time. But some building bricks is already in places to help put automation in places. For that you must though about very generic interfaces and not specialized your tools too fast. I thought you can use carving techniques for discovering file type and if your software architecture can permit it, launch script for analysis the different file of different type you discovers. For a simple cases think you just have a dump of a drive you don't know anything about, your automated process must detect the type of the disk and apply a file system drivers on it. After that your process must analyze each file detect their type and extract most of the metadata of the file type once you have reach certain level of `recursively` you maybe could correlate the data in an automated way to answer the questions of the user. Its means if the user search for images taken between date xx and yy, you could use the metadata to produce a report which will include automatically all the data found as that date. But as I say before the major drawback is that you can rely entirely on automated process because maybe some date where modified some images you are searching will be not include on the report and you can have lose some precious information’s. I thought too that automated process could be a help not to detect data we want but rather to exclude data we're sure we doesn't want in the same ways that NSRL database is used nowadays.
Participant I		no	An automated process would be at a disadvantage in court. It would set the stage for individuals to simply launch scripts instead of examining the evidence. There would be too many questions of the individual who launched the automated process, they may or may not understand the underlying scripts or processes and would open the door to the defense council. This could potentially lead to a weakening of the profession by having a push button examiner.
Participant J		no	I'm answering this as NO, because no two investigations are alike. There are common activities that can be automated (e.g. extracting all strings, finding pictures, etc. etc. etc.) just as in the traditional forensics world, there are "standard tests" run against evidence (e.g. gunshot residue, DNA comparison, etc.) The activities which are particular to a case would be difficult (theoretically impossible) to automate completely. The non-case-specific activities (e.g. string extraction) can be automated to a degree.

Participant K	yes		Requires human supervision for accuracy and completeness??
Participant L		no	
Participant M		no	Not at the moment – technology is developing too fast
Participant N	yes		

Question 5.

If we were to propose a global database of forensic case studies that can assist in pattern analysis and forensic crime prediction, would you contribute to it with your own cases?

Response: Yes: 5

No: 8

Summary: In this question, a difference came to light between investigators and red tape issues. This point to the difference between academic scientists, anticipating to find answers to pattern analysis of complex data and the button investigators that are kept back by red tape issues regarding privacy and non-disclosure agreements.

The feedback reflects the unresponsive trend of the digital forensic market in regard to participation that might hamper their own investigation niches or forensic interpretation – either reflecting to the software they are using or to the success of the investigation.

If the red tape investigators are not bound to these rules and get a standard global non-disclosure policy in place, then corpora creation might succeed.

	Yes/No		Commend
Participant A	Yes		That is slightly outside of computer forensics as that is criminal intelligence analysis. There are people doing this worldwide now and have been for over 20 years. They monitor crime trends and do it for each type of crime and then details within it. The NZ police have being doing it at least since the early 90s for all types of crimes. The Australian federal government has an entire Government Department doing just that.
Participant B	yes		I would be very interested in pattern analysis. I would have to get my supervisor's approval to release case data, but it should not be a problem.
Participant C		no	I fail to see the point of such a database.
Participant D		no	That would be a virtually impossible thing to do as there is a lot of very private and sensitive data stored on the computers. You would need their authorization to release the data and it is becoming very hard these days to even get the data to analyze, even if they are contracting you to conduct an examination. Each case I conduct today, on whatever level, I have to sign a

			confidentiality document before the examination is conducted. As an ex Police Analyst, I would find it very hard to believe that any pattern analysis would be produced and or crime predictions. This has already been tested in the US by the Criminal Justice system and the FBI at Quantico and has been discontinued.
Participant E		no	You are dreaming if you think that agency's could share personal information from their investigations on a global (or even national) basis. There is way too much red tape for this idea to ever be achievable.
Participant F		Possibly???	There are many privacy, confidentiality and access issues that would need to be addressed before I would commit to this.
Participant G		no	The cases that I work for are company's proprietary and confidential and I won't be able to share from it.
Participant H		no	We really can't provide any data of our client, but if you think about creating cases especially for the database in this case we can answer yes. We use existing cases we already use some cases we can find on internet like DFRWS challenges or honeynet-project challenges for training our employees, we also developed specially crafted dump to test our tools and other parts tools. I thought it's very important to know what the limitation of software is and to train forensic specialist.
Participant I		no	I am not able to, which is an issue you might run into, agencies tend not to share. Since there is such a closed community, the likelihood that agencies would actually contribute, weed through the red tape, deal with budgetary constraints and freely assist other agencies that are competing for funds is unlikely.
Participant J	yes		As long as the client consents.
Participant K	Yes/no		Requires disclosure regulations??
Participant L	yes		
Participant M		no	
Participant N		no	

Data from Section C

In this section we expect responses that show recommendations to improvement the existing investigation processes.

Question 1.

A 'traditional' framework for forensic IT investigation

Summary: We observed that participants maintain a good combination of traditional framework processes when they conduct their investigations. For instance, Reporting followed by Data Gathering and Analysis are the forerunners.

When we look at the lowest grading of Data Comparison, this slot well into the results form Section B, question 5, indicating that if participants do not want to contribute to a database creation, thus having data to analyze, then they have nothing to compare their results against. Therefore, we might argue that the "Data Comparison" is not as important to some investigators.

Participant A

	Not at all (1)	**Sometimes (2)**	**Most of the time (3)**	**Almost always (4)**	**All the time (5)**
Preparation/awareness					5
Analysis					5
Data gathering					
Method of capturing and duplication					5
Data comparison				4	
Reporting					5
Storage					5
Preservation					5
Post investigative analysis				4	

Participant B

	Not at all (1)	Sometimes (2)	Most of the time (3)	Almost always (4)	All the time (5)
Preparation/awareness					5
Analysis					5
Data gathering				4	
Method of capturing and duplication					5
Data comparison			3		
Reporting					5
Storage					5
Preservation					5
Post investigative analysis				4	

Participant C

	Not at all (1)	Sometimes (2)	Most of the time (3)	Almost always (4)	All the time (5)
Preparation/awareness					5
Analysis					5
Data gathering					5
Method of capturing and duplication					5
Data comparison					5
Reporting					5
Storage					5
Preservation					5
Post investigative analysis					5

Participant D

	Not at all (1)	Sometimes (2)	Most of the time (3)	Almost always (4)	All the time (5)
Preparation/awareness				4	
Analysis					5
Data gathering					5
Method of capturing and duplication				4	
Data comparison				4	
Reporting					5
Storage			3		
Preservation				4	
Post investigative analysis				4	

Participant E

Did not participate

Participant F

Participants' response: "Nothing in report???"

Participant G

	Not at all (1)	Sometimes (2)	Most of the time (3)	Almost always (4)	All the time (5)
Preparation/awareness			3		
Analysis					5
Data gathering					5
Method of capturing and duplication				4	
Data comparison			3		
Reporting				4	

Storage					5
Preservation					5
Post investigative analysis					5

Participant H

Did not participate

Participant I

Not completed

Participant J

	Not at all (1)	Sometimes (2)	Most of the time (3)	Almost always (4)	All the time (5)
Preparation/awareness				4	
Analysis					5
Data gathering			3		
Method of capturing and duplication			3		
Data comparison				4	
Reporting					5
Storage			3		
Preservation			3		
Post investigative analysis				4	

Participant K

	Not at all (1)	Sometimes (2)	Most of the time (3)	Almost always (4)	All the time (5)
Preparation/awareness					5
Analysis				4	
Data gathering				4	
Method of capturing and duplication				4	
Data comparison			3		
Reporting					5
Storage					5
Preservation					5
Post investigative analysis				4	

Participant L

	Not at all (1)	Sometimes (2)	Most of the time (3)	Almost always (4)	All the time (5)
Preparation/awareness				4	
Analysis				4	
Data gathering				4	
Method of capturing and duplication			3		
Data comparison			3		
Reporting					5
Storage				4	
Preservation			3		
Post investigative analysis				4	

Participant M

	Not at all (1)	Sometimes (2)	Most of the time (3)	Almost always (4)	All the time (5)
Preparation/awareness				4	
Analysis					5
Data gathering					5
Method of capturing and duplication					5
Data comparison			3		
Reporting					5
Storage		2			
Preservation					5
Post investigative analysis			3		

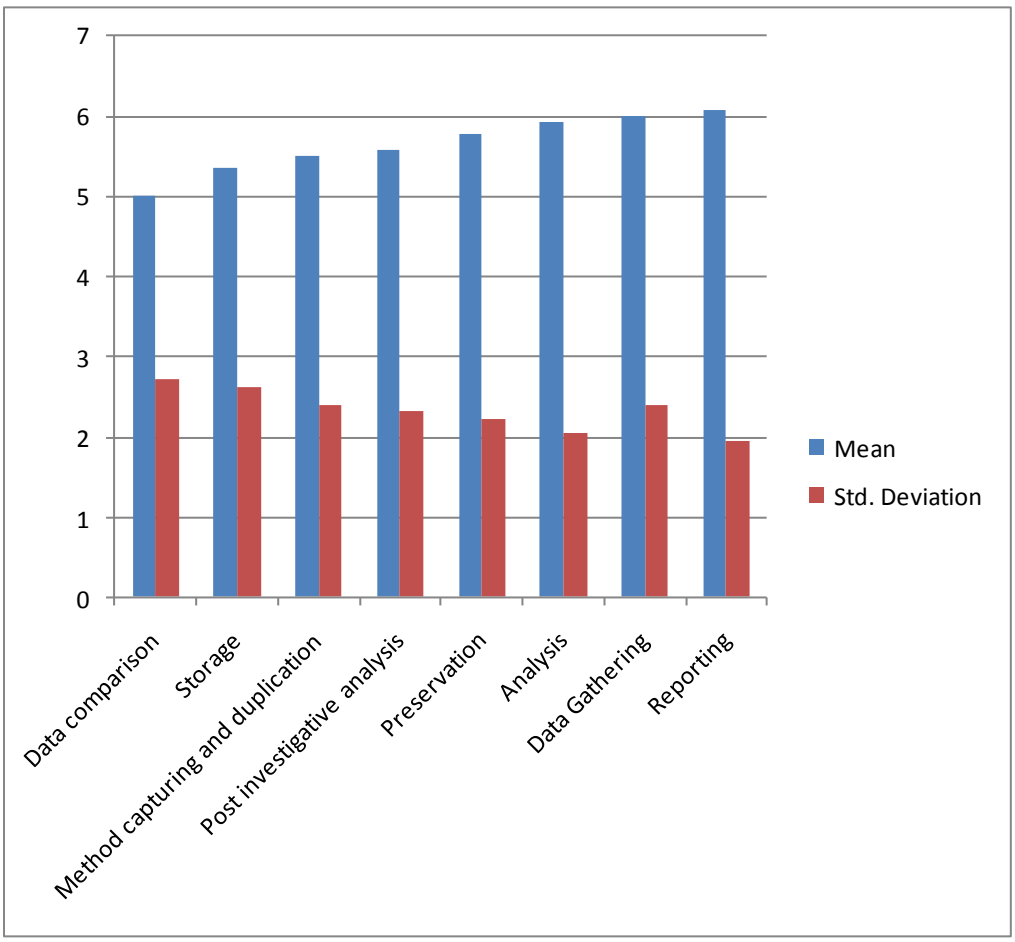
Participant N

	Not at all (1)	Sometimes (2)	Most of the time (3)	Almost always (4)	All the time (5)
Preparation/awareness					5
Analysis				4	
Data gathering				4	
Method of capturing and duplication			3		
Data comparison		2			
Reporting					5
Storage		2			
Preservation					5
Post investigative analysis					5

Using of a "traditional forensic framework"

A higher score indicate the frequency of use.

A higher score indicates higher importance.



The sequence of the investigation processes

Question 2.

Sequence of Investigation Processes

Summary: We used data in the "Present Sequence" and moved the same values to the "Your preferred order" for analysis.

Investigations started out very firm with data Collection and Knowledge of the Investigation. On the next level Identification, Storage and Preservation is set. A strange occurrence is to have Duplication only after Storage and Preservation – this might lead to disputed data corruption and place the chain of custody in question.

Surprisingly Prediction and Awareness is equally important in the first steps of investigation. This might be clear from the results sets obtained from participants, indicating that no predictive process exists at present. Although we had a few indicators for predictive analysis, this does not reflect in the existing investigation process.

Participant A

	Present Sequence	**Your preferred order**
Identification and Investigation of the incident	5	5
Analysis of the gathered data	7	7
Knowledge of the investigation	6	6
Collection, Gathering of the data	1	1
Examination, Cross referencing existing forensic incidents within similar context and occurrences	No	
Source identification – the originator of the digital event	No	
Storage of the data	4	4
Duplication of the data	3	3
Presentation of the data	10	10
Hypothesis	No	
Decision, Final finding of the case	9	9
Preservation	2	2
Examination	8	8
Analysis of the final case summary	11	11
Prediction of case as “most likely to happen again”	no	
Awareness of a potential digital event	no	

Participant B

	Present Sequence	**Your preferred order**
Identification and Investigation of the incident	2	2
Analysis of the gathered data	3	3

Knowledge of the investigation	4	4
Collection, Gathering of the data	1	1
Examination, Cross referencing existing forensic incidents within similar context and occurrences	9	9
Source identification – the originator of the digital event	10	10
Storage of the data	6	6
Duplication of the data	7	7
Presentation of the data	13	13
Hypothesis	11	11
Decision, Final finding of the case	14	14
Preservation	5	5
Examination	8	8
Analysis of the final case summary	12	12
Prediction of case as “most likely to happen again”		15
Awareness of a potential digital event		16

Participant C

Participant D

Participant E

Participant F

Participant G

Participant H

	Present Sequence	**Your preferred order**
Identification and Investigation of the incident	3	3
Analysis of the gathered data	10	10
Knowledge of the investigation	1	1
Collection, Gathering of the data	5	5
Examination, Cross referencing existing forensic incidents within	11	11

similar context and occurrences		
Source identification – the originator of the digital event	4	4
Storage of the data	8	8
Duplication of the data	7	7
Presentation of the data	14	14
Hypothesis	12	12
Decision, Final finding of the case	15	15
Preservation	6	6
Examination	9	9
Analysis of the final case summary	13	13
Prediction of case as “most likely to happen again”	16	16
Awareness of a potential digital event	2	2

Participant I

Can’t answer this question. The steps that you have listed here encompass many different people and many different disciplines.

Participant J

Note: I'm not sure how to answer this question appropriately, since I have dealt with scenarios where each could be considered a "1".

Participant K

We are at the service of the court and the legality/completeness??/compliance?? Is top priority. The list here is of a technical nature that the lab is accountable?? To. Order will vary between cases and at discontum?? Of investigation?? Context etc.

Participant L

	Present Sequence	**Your preferred order**
Identification and Investigation of the incident	15	15
Analysis of the gathered data	16	16
Knowledge of the investigation	1	1
Collection, Gathering of the data	2	2

Examination, Cross referencing existing forensic incidents within similar context and occurrences	3	3
Source identification – the originator of the digital event	4	4
Storage of the data	5	5
Duplication of the data	6	6
Presentation of the data	8	8
Hypothesis	7	7
Decision, Final finding of the case	9	9
Preservation	10	10
Examination	11	11
Analysis of the final case summary	12	12
Prediction of case as “most likely to happen again”	13	13
Awareness of a potential digital event	14	14

Participant M

Not answered

Participant N

	Present Sequence	**Your preferred order**
Identification and Investigation of the incident	1	5
Analysis of the gathered data	9	13
Knowledge of the investigation	2	6
Collection, Gathering of the data	3	7
Examination, Cross referencing existing forensic incidents within similar context and occurrences	000	4
Source identification – the originator of the digital event	4	8
Storage of the data	6	10
Duplication of the data	7	11
Presentation of the data	10	14

Hypothesis	n/a	2
Decision, Final finding of the case	11	15
Preservation	5	9
Examination	8	12
Analysis of the final case summary	12	16
Prediction of case as “most likely to happen again”	n/a	3
Awareness of a potential digital event	n/a	1

Sequence of Investigation Processes

A lower score reflects on the first activity – the higher score indicate a later activity in investigation procedures.