A Hierarchical, Objectives-Based Framework for the Digital Investigations Process

Nicole Lang Beebe
The University of Texas at San Antonio
Department of Information Systems and
Technology Management
6900 North Loop 1604 West
San Antonio, Texas 78249
nicole.beebe@utsa.edu

Jan Guynes Clark
The University of Texas at San Antonio
Department of Information Systems and
Technology Management
6900 North Loop 1604 West
San Antonio, Texas 78249
jan.clark@utsa.edu

ABSTRACT

Digital investigations, whether forensic in nature or not, require scientific rigor and are facilitated through the use of standard processes. Such processes can be complex in nature. A more comprehensive, generally accepted digital investigation process framework is therefore sought to enhance scientific rigor and facilitate education, application, and research. Previously proposed frameworks are predominantly single-tier, higher order process models that focus on the abstract, rather than the more concrete principles of the investigation. We contend that these frameworks, although useful in explaining overarching concepts, fail to support the inclusion of additional layers of detail needed by various framework users. We therefore propose a multi-tier, hierarchical framework to guide digital investigations. Our framework includes objectives-based phases and sub-phases that are applicable to various layers of abstraction, and to which additional layers of detail can easily be added as needed. Our framework also includes principles that are applicable in varied ways to all phases. The data analysis function intended to identify and recover digital evidence is used as an example of how the framework might be further populated and used. The framework is then applied using two different case scenarios. At its highest level, the proposed framework provides a simplified view and conceptual understanding of the overall process. At lower levels, the proposed framework provides the granularity needed to achieve practicality and specificity goals set by practitioners and researchers alike.

Keywords

Digital investigative process, digital forensics, computer forensics, analysis, framework

I. INTRODUCTION

The traditional physical forensic science discipline developed along side its underlying physical and biological sciences over the course of several decades (Palmer 2002). In contrast, the digital forensic science discipline is relatively nascent and significantly lags behind its better developed underlying computer science (Palmer 2001; Palmer 2002; Carrier and Spafford 2003). Because of this, digital forensics researchers, practitioners, and consumers are actively seeking a more comprehensive, generally accepted digital investigations process framework (Palmer 2001). Such a framework will provide a common starting place from which established theory (e.g. computer science theory and forensic science theory) can be scientifically applied to the digital forensic science discipline. The framework will also enable new theory development and the identification of research and development requirements. The resultant scientific rigor that will be applied using the framework as its foundation will transform current digital forensics practices into digital forensic *science*, as defined by academics and practitioners at the first Digital Forensic Research Workshop (DFRWS) in 2001:

> Digital Forensic Science – "The use of scientifically derived and proven methods toward the preservation, collection, validation, identification, analysis, interpretation, documentation, and presentation of digital evidence[1] derived from digital sources for the purpose of facilitation or furthering the reconstruction of events found to be criminal, or helping to anticipate unauthorized actions shown to be disruptive to planned operations." (Palmer, 2001: 16)

Although still relatively new, the digital forensic discipline impacts a diverse user community. This community ranges from digital forensic specialists in law enforcement, industry, and the military who conduct the digital forensic operations to educators and researchers who teach and conduct research in a variety of areas related to digital forensic science. An effective framework must be applicable to the entire community, offering the ability to improve both the theory and practice of digital forensic science.

A review of the prevailing digital investigation process models presented to date (Palmer 2001; DoJ 2001; Reith et al. 2002; Carrier and Spafford 2003; Mandia et al. 2003; Mohay et al. 2003; Nelson et al. 2004; Ó Ciardhuáin 2004; Casey and Palmer 2004) discloses a predominant focus on single-tier, higher order process models. While this is a natural starting point, the complexities of the digital investigation process simply cannot be represented at that level. A simple analogy is useful here. Arguably, flying an airplane can be a function of a higher order framework consisting of three phases: take-off, fly, and land. Few pilots could accomplish this task, however, without obtaining more detail regarding each of the phases. Likewise, greater detail pertaining to each phase of a digital investigation process framework is needed in order to improve usability for practitioners and researchers alike. We therefore propose a more comprehensive digital investigation framework that focuses on both theory and practice and includes lower order objectives-based sub-phases for each higher order phase.

This paper proceeds as follows: In Section II, we introduce the proposed framework and define its primary phase structure. The sub-phase structure for the Data Analysis Phase presented in Section III demonstrates how the framework can be further developed and applied. This is followed by example implementations of the proposed framework, utilizing commercial and law enforcement case scenarios in Section IV. Section V presents benefits and limitations of the framework; conclusions are presented in Section VI.

II. PROPOSED DIGITAL INVESTIGATIVE PROCESS FRAMEWORK

We sought to develop a framework that simplifies the complex, yet provides a mechanism for including the layers of detail needed by its users. Our primary goal was to ensure the framework's expansion capability while integrating previous frameworks and models to the extent prudent. The rationale in doing so was two-fold. First, we sought to leverage the philosophies and benefits of previously proposed frameworks and models. Second, in any community effort, it is important to create synergy between different perspectives. Any framework institutionalized through subsequent intellectual discourse and practical use must take into consideration differing perspectives, approaches, and vernacular.

The robustness of a framework is a function of its usability and acceptability. To achieve usability and acceptability, we incorporated phases, sub-phases, principles, and objectives. Phases and sub-phases are distinct, discrete steps in the process that are usually a function of time and suggest a necessarily sequential and sometimes iterative approach. Principles, on the other hand, are overarching procedures, guidelines, and/or methodological approaches that overlap some or all of the phases and sub-phases. Unlike phases, principles are not distinct, discrete steps in the process; instead, they represent goals and objectives sought throughout the process. Proper documentation is an example of a principle.

[1] The authors use the term "evidence" to apply to information of value derived from data, independent of whether the investigation is forensic or non-forensic in nature (the distinction being whether judicial actions are sought).

Because the framework presented is inherently a process model, the output of each phase serves as input to succeeding phases. This natural process model flow, along with key investigative principles, such as Information Flow (Ó Ciardhuáin 2004) and Case Management (Casey and Palmer 2004), serve to tie the phases of the framework together. The Information Flow principle unifies all phases of the proposed framework, whereas the Case Management principle unifies the phases applicable and conducted during the course of each unique investigation. Also unifying the phases of the framework are the investigative objectives upon which the investigation is based. A generic model of our proposed framework is shown in Figure 1.

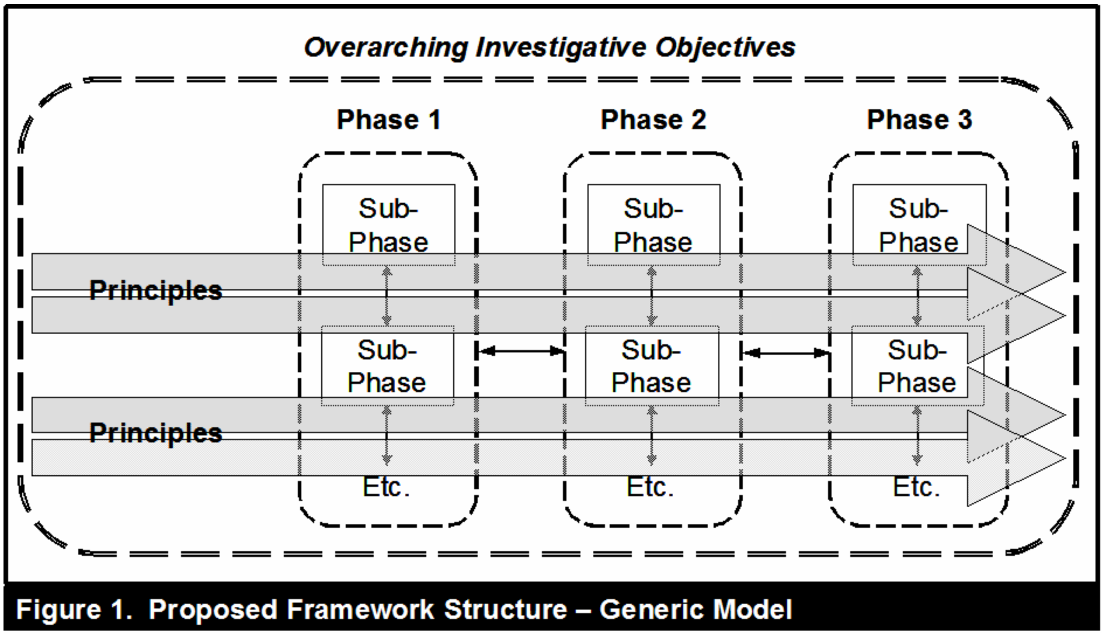


Figure 1. Proposed Framework Structure – Generic Model

In our experience, a digital investigation framework *must* be based on objectives, rather than tasks. This is because the uniqueness of each situation and “digital crime scene” (Carrier and Spafford, 2003) necessitates a non-checklist approach. A different subset of steps is likely taken in each situation. Therefore, the decision of which steps to take in any given situation can be more easily made when the steps are outlined in an objectives-based fashion, as opposed to a task-based fashion. As an example, it is easier for a practitioner to decide whether a step described as “Determine whether unauthorized software has been installed” is more relevant to the investigation than the task described as “Examine the Registry.” Furthermore, many tasks apply to more than one objective and can easily be matrixed to objectives, which allows practitioners to select relevant objectives-based tasks and then accomplish recommended sub-tasks accordingly. There are other benefits of this approach related to the furtherance of research and development activities, which will be discussed later in this paper.

Our proposed framework parsimoniously encapsulates all phases and activities outlined in prevailing models presented to date (Palmer 2001; DoJ 2001; Reith et al. 2002; Carrier and Spafford 2003; Mandia et al. 2003; Mohay et al. 2003; Nelson et al. 2004; Ó Ciardhuáin 2004; Casey and Palmer 2004). In this Section, we describe the first tier phases and process principles. Although all phases should consist of sub-phases, for the purpose of this paper, we focused on the Data Analysis sub-phases. We chose to focus on this phase because although Data Analysis is very important and complex, few researchers or practitioners have focused on it when developing prior frameworks. Additionally, it provides the best opportunity to demonstrate the necessity for a hierarchical, objectives-based digital investigations process. We expect future efforts within the community will develop the other sub-phase structures. Following is a discussion of each phase, including phase definitions and example activities that characterize each phase.

First-Tier Phases

First-tier phases are distinct and clearly defined. They are discrete in that clear delineations exist between phases. That is to say that each phase has a clear event that initiates it and clear output at its conclusion. The phases are sequentially ordered and are a function of time. First-tier phases are largely non-iterative within the scope of a single incident, but not prohibitively so (see Figure 2). We define "incident" as any event which is considered suspect or abnormal (e.g. policy violation, security breach, suspected criminal activity, etc.). There may be situations where within-investigation iteration is needed. However, we purport the non-iterative within-investigation view will prevail in most situations, especially as an organization's digital forensic investigation capability and maturity improves.

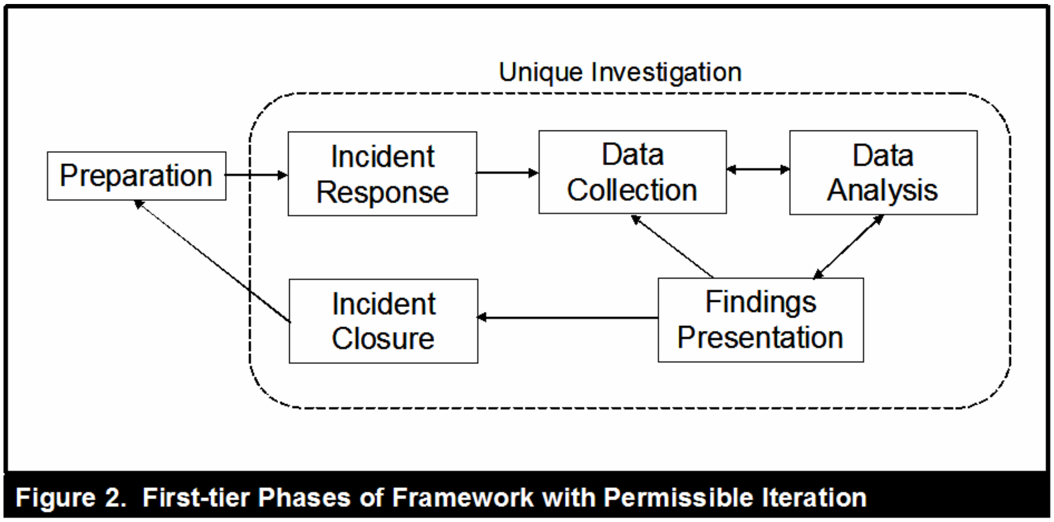


Figure 2. First-tier Phases of Framework with Permissible Iteration

Preparation Phase

Simply put, a digital investigation requires digital evidence which is not entirely existent by default and is frequently damaged or destroyed during standard containment, eradication, and recovery activities. Thoughtful preparation can improve the quality and availability of digital evidence collected, while minimizing organizational cost and burden. This equates to an organization's "forensic readiness" posture (Rowlingson 2004). The Preparation Phase includes those steps taken by companies to maximize digital evidence availability in support of deterrence, detection, response, investigation, and prosecution related to computer security incidents. Preparation activities include, but are not limited to:

- Assess risk considering vulnerabilities, threats, loss/exposure, etc.;
- Develop an information retention plan (both pre/post-incident);
- Develop an Incident Response Plan, including policies, procedures, personnel assignments, and technical requirements definition;
- Develop technical capabilities (e.g. response toolkits);
- Train personnel;
- Prepare host and network devices;
- Develop evidence preservation and handling procedures; and
- Develop legal activities coordination plan (both pre/post-incident).

Note that the preparation activities are focused on the victim organization, and not the investigator, who is presumably separate from the victim (either functionally, or organizationally). Investigators focus their preparation activities on Incident Response Planning, technical capability development, training, and evidence preservation handling procedure development.

Incident Response Phase

The Incident Response Phase consists of the detection and initial, pre-investigation response to a suspected computer crime related incident, such as a breach of computer security, the use of a computer to view contraband material (e.g. child pornography), etc. The purpose of this phase is to detect, validate, assess, and determine a response strategy for the suspected security incident. Incident response activities include, but are not limited to:

- Detect or suspect unauthorized activity;
- Report detected or suspected unauthorized activity to proper individual(s)/authority;
- Validate the incident;
- Assess damage/impact via interviews of technical/business personnel, review pertinent logs, review network topology, etc.;
- Develop a strategy regarding containment, eradication, recovery, and investigation, considering business, technical, political, and legal factors/goals;
- Coordinate, as applicable, managerial, human, legal, and law enforcement resources; and
- Formulate the Investigation Plan for data collection and analysis.

The final, seventh step—formulate an Investigation Plan—will vary in form depending on the investigative objective of the victim organization. If the victim organization seeks law enforcement involvement, the scope of the victim organization's Investigation Plan will be limited, and the law enforcement agent(s) will develop the primary Investigation Plan. In that case, the incident response activities conducted by law enforcement will primarily consist of: (1) information collection activities that support the development of the Investigation Plan and facilitate a proper crime scene response posture and data collection effort, and (2) acquisition of proper legal authority (e.g. preparation of affidavits and receipt of search warrants, obtaining proper legal consent, etc.)

Data Collection Phase

Data and information required to validate an incident and determine its impact will be initially collected during the Incident Response Phase. Once a decision has been made to investigate the incident, regardless of its scope or anticipated legal or administrative actions, the formal Data Collection Phase ensues. The purpose of the Data Collection Phase, therefore, is to collect digital evidence in support of the response strategy and investigative plan. Data collection activities include, but are not limited to:

- Complete "live response" data collection, which likely began during the Incident Response Phase;
- Obtain network-based evidence from applicable sources (e.g. intrusion detection systems, routers, firewalls, log servers, etc.);
- Obtain host-based evidence from applicable sources (e.g. volatile data, system date/time information, hard drives or forensic duplicates thereof, etc.);
- Obtain removable media evidence from applicable sources (e.g. backup tapes, floppy disks, CD-ROMs, flash memory devices);
- Install activity monitoring capability (e.g. network monitors, system monitors, surveillance cameras);
- Ensure integrity and authenticity of the digital evidence (e.g. write protection, hashes, etc.); and
- Package, transport, and store the digital evidence.

Data Analysis Phase

The Data Analysis Phase is arguably the most complex and time consuming phase in the digital investigations process. The purpose of the Data Analysis Phase is confirmatory analysis (to confirm or refute allegations of suspicious activity) and/or event reconstruction (answer "who, what, where, when, why, and how" type questions). Data collected during the Data Collection Phase is surveyed, extracted, and reconstructed during the Data Analysis Phase. Data analysis activities include, but are not limited to:

- Transform the voluminous amount of data collected during the Data Collection Phase into a more manageable size and form for analysis;
- Conduct an initial data survey to recognize obvious pieces of digital evidence and assess the skill level of the suspect(s);
- Employ data extraction techniques (e.g. keyword searches, extraction of unallocated space and file slack, file timeline/mapping, hidden data discovery/extraction, etc.); and
- Examine, analyze, and event reconstruct the data to answer critical investigative questions.

Presentation of Findings Phase

The purpose of the Presentation of Findings Phase is to *communicate* relevant findings to a variety of audiences, including management, technical personnel, legal personnel, and law enforcement. In naming this phase, we selected "presentation of findings" over "reporting" because it suggests careful consideration about how to best communicate information to various audiences. A technical report, which is the natural inclination of digital forensic analysts, tends to document relevant information, but does so in a manner that is not necessarily helpful for those who then act upon the information provided. The presentation of findings may be written, oral, or presented in both formats. The presentation(s) are intended to provide both succinct and detailed confirmatory and event reconstruction information regarding the data examined in the Data Analysis Phase.

Incident Closure Phase

The Incident Closure Phase, as the name implies, focuses on closure of the investigation. However, it is important to not only close out this investigation and act upon decisions related to it, but also to attempt to preserve knowledge gained to enhance subsequent investigations. The steps include:

- Conduct a critical review of the entire process and investigation to identify and apply lessons learned;
- Make and act upon decision(s) that result from the findings presentation phase;
- Dispose of the evidence (e.g. return to owner, destroy, cleanse and re-use (forfeiture)—all as applicable and legally permissible); and
- Collect and preserve all information related to the incident.

Comparison to Previous Frameworks and Models

As mentioned previously, synergy with previously proposed frameworks and models was a primary goal of the proposed framework. Table 1 demonstrates the comprehensiveness of the proposed framework, as compared to previously proposed models. Table 2 then provides a mapping of the specific phases and steps from previously presented models to the proposed model's first tier phases.

Table 1. Mapping of Previous Models/Frameworks to the Proposed Framework

(Beebe/Clark Framework Phases)

	Prep.	Incident Response	Data Collect.	Data Analysis	Findings Present.	Incident Closure
Palmer, 2001 (DFRWS model)		✓	✓	✓	✓	✓
Department of Justice, 2001			✓	✓	✓	
Reith et al, 2002 (Abstract Model)		✓	✓	✓	✓	✓
Mandia et al, 2003	✓	✓	✓	✓	✓	
Carrier and Spafford, 2003	✓	✓	✓	✓	✓	✓
Nelson et al, 2004		✓	✓	✓	✓	✓
Ó Ciardhuáin, 2004		✓	✓	✓	✓	✓
Casey and Palmer, 2004		✓	✓	✓	✓	

Table 2. Mapping of Previous Models/Frameworks to the Proposed Framework (Beebe/Clark Framework)

	Prep.	Incident Response	Data Collect.	Data Analysis	Findings Present.	Incident Closure
Palmer, 2001 (DFRWS model)						
Identification		✓				
Preservation			✓			
Collection			✓			
Examination				✓		
Analysis				✓		
Presentation					✓	
Decision						✓
Department of Justice, 2001						
Collection			✓			
Examination				✓		
Analysis				✓		
Reporting					✓	
Reith et al, 2002 (Abstract Model)						
Identification		✓				
Preparation [for the current investigation]		✓				
Approach strategy		✓				
Preservation			✓			
Collection			✓			
Examination				✓		
Analysis				✓		
Presentation					✓	
Returning evidence						✓
CONTINUED ON NEXT PAGE						

Mandia et al, 2003						
Pre-incident preparation	✓					
Detection of incidents		✓				
Initial response		✓				
Formulate response strategy		✓				
Data collection			✓			
Data analysis				✓		
Reporting					✓	
Carrier and Spafford, 2003						
Readiness	✓					
Deployment		✓				
Digital crime scene investigation			✓	✓	✓	
Preservation			✓			
Survey				✓		
Documentation	✓	✓	✓	✓	✓	✓
Search and collection			✓	✓		
Reconstruction				✓		
Presentation					✓	
Review						✓
Nelson et al, 2004						
Initial assessment		✓				
Approach strategy (“design”)		✓				
Resource determination		✓				
Copy evidence			✓			
Risk identification & mitigation		✓				
Test approach strategy (“design”)			✓	✓		
Data analysis and recovery				✓		
Data investigation				✓		
Report					✓	
Critique						✓
Ó Ciardhuáin, 2004						
Awareness		✓				
Authorization		✓				
Planning		✓				
Notification		✓				
Search/Identify			✓			
Collection			✓			
Transport			✓			
Storage			✓			
Examination				✓		
Hypothesis				✓		
Presentation					✓	
Proof/Defense					✓	
Dissemination						✓
Casey and Palmer, 2004						
Incident alerts or accusation		✓				
Assessment of worth		✓				
Incident/crime scene protocols		✓	✓			
Identification or seizure			✓			
Preservation			✓			
Recovery				✓		
Harvesting				✓		
Reduction				✓		
Organization and search				✓		
Analysis				✓		
Reporting					✓	
Persuasion and testimony					✓	

As shown, each of the previous frameworks (with the exception of the DoJ Model) address most or all of the First Tier Phases of our proposed framework. At first glance, one would assume that our model contributes little more than what has already been proposed. However, none of the other models provides sufficient detail at each of sub-phases to enable all members of the digital forensics community to utilize and understand the nuances of the framework. Following is a more detailed discussion of our underlying principles and objectives-based sub-phases.

Digital Investigation Principles

Certain principles apply to all phases of the digital investigations process, and should therefore not be cordoned off as distinct, discrete phases or steps. Doing so diminishes their impact on the overall process. Such principles represent overarching goals and may necessitate different actions during each phase of the digital investigations process in order to achieve those goals. Principles can naturally translate to constraints, or be considered constraints in and of themselves. Examples of digital investigation principles include, but are not limited to:

- Evidence preservation,
- Documentation,
- Proper investigative authority,
- Sensitivity and/or classification,
- Investigative priority,
- Information flow and controls (Ó Ciardhuáin 2004),
- Case management (Casey and Palmer 2004), and
- Process improvement feedback.

To illustrate the importance and application of digital investigation principles, *evidence preservation* and *documentation* are discussed in detail.

Evidence Preservation Principle

The primary goals of the evidence preservation principle are (1) to maximize evidence availability and quality, and (2) maintain the integrity of the evidence during the digital investigation process. The application of the evidence preservation principle varies within each phase as follows:

- **Preparation Phase:** Ensure the availability and quality of digital evidence when needed.
- **Incident Response Phase:** Ensure evidence preservation during initial validation and assessment ("live response") activities.
- **Data Collection Phase:** Ensure data is collected in a forensically sound manner. Some example activities include creating forensic duplicates, employing write-protection technology, calculating checksums and hashes, and applying environmental protection activities.
- **Data Analysis Phase:** Forensic working copies are created as needed. Additionally, the analyst must be cognizant of which steps and processes modify working copies (e.g. file access times) and perform steps methodically from least invasive to most invasive and/or continually return to use of clean copies.
- **Presentation of Findings Phase:** Communicate findings in a manner that facilitates future corroboration.
- **Incident Closure Phase:** Properly dispose of evidence and retain information related to the entire process.

Documentation Principle

The goal of the documentation principle is to permanently (or semi-permanently as applicable) record all information relevant to and/or generated during the digital investigative process to support decision making and the legal, administrative, etc. processing of those decisions. Information that should be documented during each phase includes, but is not limited to the following (Palmer 2001; DoJ 2001; Reith et al. 2002; Carrier and Spafford 2003; Mandia et al. 2003; Nelson et al. 2004):

- **Preparation Phase:** Risk assessment/management information and decisions; policies; procedures; "known good" system information and hashes; training program requirements and progress; and legal coordination.
- **Incident Response Phase:** All information related to detection of the incident and any information pertaining to "who, what, where, when, why, and how" type questions; witness statements; and damage information, including direct costs and personnel time.
- **Data Collection Phase:** Information pertaining to the "state" of systems when data is collected (physical connections, processes running, network interface mode, date/time, open ports, etc.); evidence identification mechanisms (marking); and chain of custody information.
- **Data Analysis Phase:** Digital forensics tools used; processes, actions, and approaches taken during the analysis; and all findings, including things later deemed irrelevant.
- **Findings presentation phase:** Communication of findings from both technical and non-technical view points, and a formal record of relevant assumptions, observations, and conclusions. The findings documentation should be timely and professional, follow the "ABC's of writing" (accuracy, brevity, and clarity), and be restricted to only what is known, not supposed.
- **Incident Closure Phase:** Permanently (or semi-permanently) retain all related documentation created during the digital investigations process.

Second-Tier Phases (Sub-Phases)

The complexity level associated with the digital investigative process necessitates a multi-tier, hierarchical framework with objectives-based sub-phases and task hierarchies (objective-task matrices). The second tier sub-phases should be inclusive of all possible types of crime and digital evidence and consist of task hierarchies subordinate to specific objectives of interest. While the objectives-based sub-phases (OBSP) will remain largely consistent from situation to situation, the specific objectives-based tasks selected in each situation will vary according to the unique needs of each investigation. Because some tasks and sub-tasks may be applicable to more than one objective, the proposed framework lends itself to the development of useful matrices. Tasks can be matrixed to the set of digital forensic objectives, enabling the digital forensic examiner to quickly determine which objectives and in turn which specific tasks are applicable to the incident and approach strategy at hand. Experience conducting digital forensic investigations and educating others in doing so punctuates the importance of this approach. The number of possible digital forensic tasks that can be accomplished on any piece of digital evidence is staggering. Investigators require an efficient mechanism to identify which tasks are needed for the investigation at hand. Focusing cognitive effort on the digital investigation objectives (e.g. determine if unauthorized system modifications have occurred, determine which accounts have been compromised) is much more manageable and effective than attempting to identify which exact tasks from a seemingly endless range of possible tasks are applicable. The proposed framework facilitates the development of task-objective matrices to reduce this cognitive burden.

It is also important that the digital investigative process framework be robust enough to apply to various layers of abstraction (Carrier 2003). Abstraction layers are used to analyze and translate data into more manageable formats, either via translation from a lower level representation to a higher, more human readable level representation (e.g. translation from binary to ASCII), or via data reduction mechanisms to lessen the amount of data to be analyzed by humans (e.g. intrusion detection system). Example layers of abstraction include physical media, media management, file system application, and network. Separate frameworks for each layer of abstraction or each type (i.e. media analysis vs. network analysis) would be too cumbersome. In the proposed framework the abstraction layer corresponds to the translation rule that will operate on the input. This is of particular importance with regard to the Data Analysis Phase, as shown in Section III.

A generic representation of the sub-phase structure is depicted in Figure 3. Note that each objective-based sub-phase (OBSP) may be related to other OBSPs within the given phase. The next section of this paper is dedicated to applying the framework to the Data Analysis Phase and infusing it with layers of detail to achieve practicality and specificity goals.

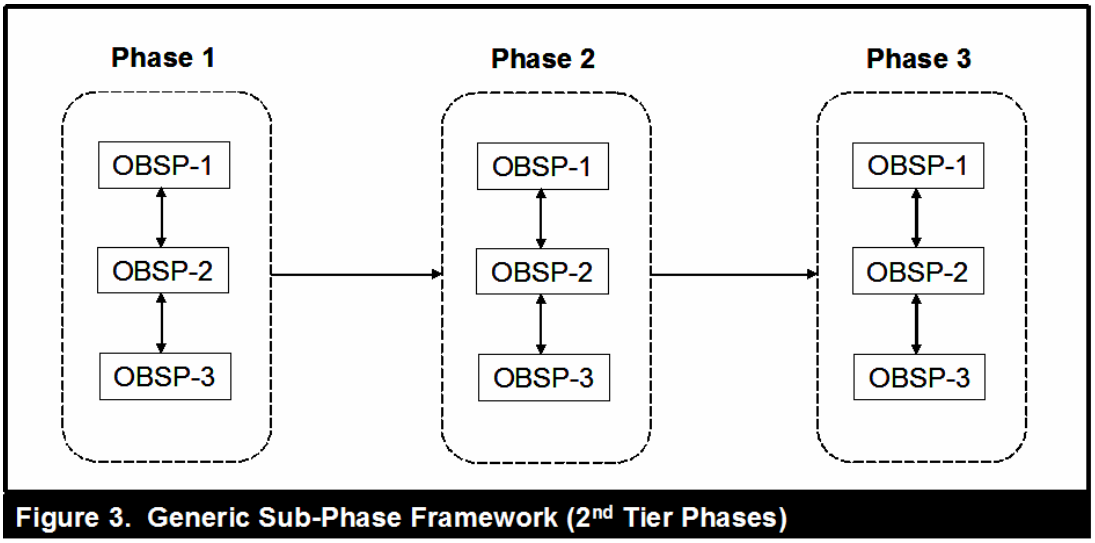


Figure 3. Generic Sub-Phase Framework (2nd Tier Phases)

III. FRAMEWORK APPLICATION TO THE DATA ANALYSIS PHASE

Review of Previously Proposed Analytical Phases

As previously presented, the purpose of the Data Analysis Phase is confirmatory analysis (to confirm or refute allegations of suspicious activity) and/or event reconstruction (answer “who, what, where, when, why, and how” type questions). Data collected during the Data Collection Phase is surveyed, extracted, and reconstructed during the Data Analysis Phase.

Digital investigation processes presented by prior researchers have differed regarding the exact definition and, more importantly, the boundaries of the Data Analysis Phase. Several frameworks and models dedicate separate phases for ‘examination’ and ‘analysis’ (Palmer 2001; DoJ 2001; Reith et al. 2002). In doing so, the examination phase is primarily characterized by search and extraction activities, whereas the analysis phase is primarily characterized by subsequent activities that generate useful information from the extracted data. Nelson et al. (2004) present a similar two phased approach to examination and analysis, except they apply different labels to their phases.

Mandia et al. (2003) describe a single data analysis related phase, but subdivide it into “Data Preparation” and “Data Analysis” sub-phases. In this case, Data Preparation includes: file list creation, deleted data recovery, unallocated space recovery, statistical data collection, partition table and file system identification, file signature analysis, and known system file identification. The Data Analysis sub-phase includes: email/attachment extraction, installed application review, string searches, software analysis, logical file review, Internet browser history review, live system data collection review, network based evidence review, identify and decrypt encrypted files, and specialized analyses.

Carrier and Spafford (2003) describe three first-tier phases that relate to the overall data analysis function. These include the ‘Survey Phase,’ the ‘Search and Collection Phase,’ and the ‘Reconstruction Phase.’ The Survey Phase “…finds obvious pieces of digital evidence for the given class of crime … [and] …

show[s] the investigator the skill level of the suspect and what analysis techniques the investigation will require" (Carrier and Spafford, 2003: 11). The Search and Collection Phase consists of data extraction and processing (e.g. keyword searches, unallocated space analysis, file activity timeline analysis, code reverse engineering, encryption analysis, and log reviews). The Reconstruction Phase "…put[s] the pieces of the digital puzzle together" (Carrier and Spafford, 2003: 12) and answers investigative "who, what, where, when, why, and how" questions.

Mohay et al. (2003) describe two first-tier phases that relate to the overall data analysis function. These include: (1) 'the live system processing and data collection,' and (2) 'the analysis of secured data.' Their process is written more from a network forensics perspective, thus 'live system processing and data collection' includes volatile data acquisition, copying system files, logical volume imaging, and obtaining system date/time information. Likewise, 'analysis of secured data' includes logical analysis of the media structure, collecting operating system configuration information, file system mapping information collection and analysis, file signature analysis, identifying file content and type anomalies, evaluating program functionality, text string and key word searching, evaluating virtual memory, and evaluating ambient data. Overarching analysis techniques include system usage analysis, Internet usage analysis, time-line analysis, link analysis, and password recovery and cryptanalysis.

Data Analysis Sub-Phase Structure

We suggest the analytical phases be bounded by a single Data Analysis phase accompanied by an iterative set of sub-phases. The typical digital forensic analysis process is characterized by decisions to search for certain data artifacts, subsequent data extraction and analysis, and then decisions to search for new, different, and/or additional data artifacts. These subsequent decisions are due to a variety of reasons, including, for example, approach strategy refinement based on information obtained and discovery of unanticipated data artifacts or evidence. This is the reason Data Analysis sub-phases must be iterative in nature.

The proposed Data Analysis sub-phases include: Survey, Extract, and Examine (referred to as the "SEE Data Analytical Approach"). As is the case with general land surveying, the primary purpose of the Survey Sub-Phase is mapping. In a land survey, surveyors are tasked with providing a detailed, precise description of a land area including its topography, elevation, boundaries, geographical coordinates, conformity with standards, and object/spatial relationships. Analogously, in a digital data analysis survey, the analyst is tasked with providing detailed, precise descriptions of various aspects of a digital object's "landscape." Some examples include mappings of file systems, logical disk partitioning, disk geometry, "landmark" locations, and irregularities. This mapping data provides and enables the following: familiarity with the digital object under analysis, indications of suspect skill level, and location of obvious and potential evidence.

The Survey Sub-Phase mapping data then facilitates data extraction activities. The purpose of the Extract Sub-Phase is to extract data from the digital object according to stated objectives, using the mapping data from the Survey Sub-Phase. Techniques such as keyword searches, deconstruction of proprietary data formats, mining for hidden data, filtering, pattern matching, and file signature analysis are examples of Extract Sub-Phase activities.

The purpose of the third and final sub-phase, the Examine Sub-Phase, is to examine the extracted data to achieve confirmatory and/or event reconstruction goals. During this phase, conclusions are made regarding the presence or absence of digital evidence (confirmatory analysis) and/or the answers to "who, what, where, when, why, and how" questions (event reconstruction analysis). Analytical techniques such as log reviews, image and text viewing, chronological and correlation assessment, and decryption are examples of Examine Sub-Phase activities.

The SEE Data Analytical Approach can be applied iteratively any number of times, with regard to any data analysis objectives, and targeting any type of evidence at any layer of abstraction. This approach is depicted in Figure 4.

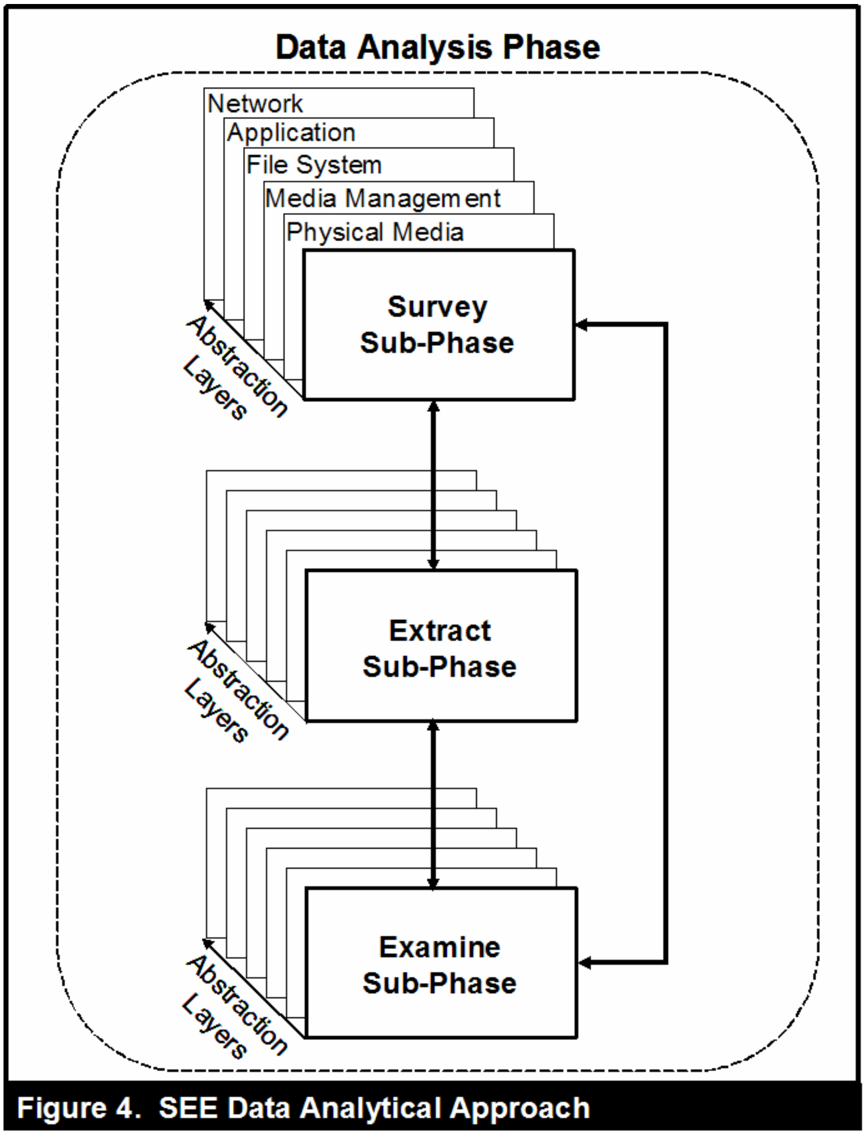


Figure 4. SEE Data Analytical Approach

Data Analysis Objective-Task Structure

The SEE Data Analytical Approach provides a phased approach to digital forensic analysis activities. Specific data recovery and analysis objectives and subordinate task hierarchies are overlaid on top of the Survey—Extract—Examine sub-phases. Arguably, the list of data recovery and analysis objectives is considerably shorter than specific tasks undertaken to achieve the objectives, especially when considering the number of tasks that are applicable to multiple objectives. As such, construction of objective-task matrices (the number of which is determined by the degree of "drill-down" desired) can greatly help examiners determine an analysis strategy. An examiner simply selects the data analysis objectives relevant to the current investigation and the objective-task matrices generate appropriate suggested task lists for consideration. An important caveat is that this approach is intended only as a decision support tool to ease the cognitive load amidst a complex process. It will not generate "checklists", and the output should not be misconstrued as such.

Table 3 provides a sample objective-task matrix (DoJ 2001; Vacca 2002; Kruse and Heiser 2002; Mandia et al. 2003; Casey 2003a, b; Mohay et al. 2003; Nelson et al. 2004). The sample provided is by no means all inclusive and is only partially indicative of one level of detail. As previously stated, additional matrices and/or matrix detail (task – sub-task – activity) can be developed to provide additional guidance based on the needs and desires of examiners and/or organizations. Exact objective-task and task-sub-task hierarchies for the data analysis phase are beyond the scope of this paper and must be developed collaboratively via intellectual discourse. The sample objective-matrix task is provided to illustrate how the framework can be infused with detail to provide the level of specificity and practicality sought by practitioners, researchers, and educators.

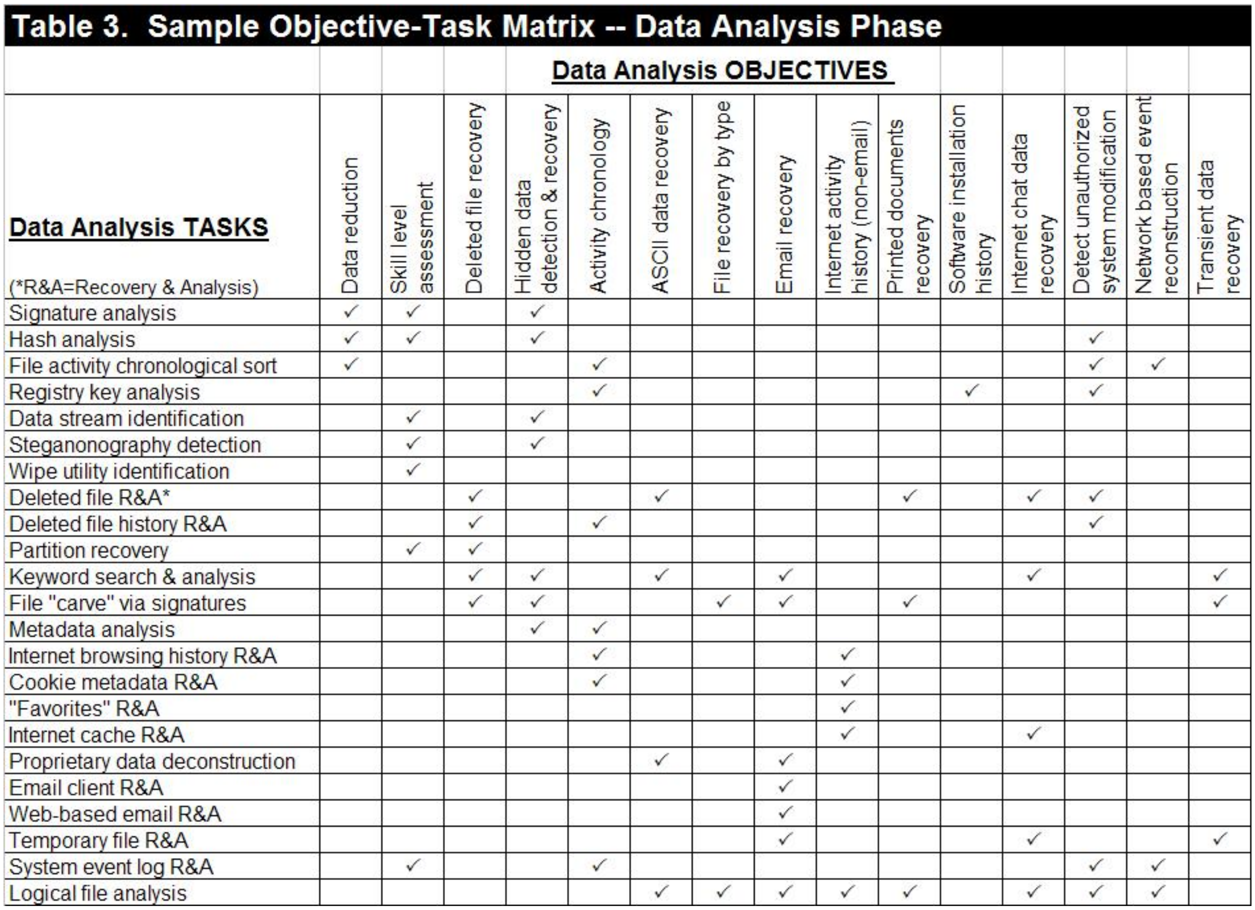


Table 3. Sample Objective-Task Matrix -- Data Analysis Phase

	Data Analysis OBJECTIVES														
Data Analysis TASKS (*R&A=Recovery & Analysis)	Data reduction	Skill level assessment	Deleted file recovery	Hidden data detection & recovery	Activity chronology	ASCII data recovery	File recovery by type	Email recovery	Internet activity history (non-email)	Printed documents recovery	Software installation history	Internet chat data recovery	Detect unauthorized system modification	Network based event reconstruction	Transient data recovery
Signature analysis	✓	✓		✓											
Hash analysis	✓	✓		✓									✓		
File activity chronological sort	✓				✓								✓	✓	
Registry key analysis					✓						✓		✓		
Data stream identification		✓		✓											
Steganonography detection		✓		✓											
Wipe utility identification		✓													
Deleted file R&A*			✓			✓				✓		✓	✓		
Deleted file history R&A			✓		✓								✓		
Partition recovery		✓	✓												
Keyword search & analysis			✓	✓		✓		✓				✓			✓
File "carve" via signatures			✓	✓			✓	✓		✓					✓
Metadata analysis				✓	✓										
Internet browsing history R&A					✓				✓						
Cookie metadata R&A					✓				✓						
"Favorites" R&A									✓						
Internet cache R&A									✓			✓			
Proprietary data deconstruction						✓		✓							
Email client R&A								✓							
Web-based email R&A								✓							
Temporary file R&A								✓				✓			✓
System event log R&A		✓			✓								✓	✓	
Logical file analysis						✓	✓	✓	✓	✓		✓	✓	✓	

As more layers of detail are added to the matrices, database technology becomes more and more helpful. Databases by their very nature represent matrices of data. The more complex the matrices, the more helpful database technology can be. The proposed framework lends itself to users' varied inclinations to make such matrices as simplistic or comprehensive as they desire or prefer. For example, some investigators may wish to limit the level of detail and specificity of their matrices to those shown in Table 3 (basic objective-task matrix). Others, on the other hand, may wish to develop and utilize more comprehensive matrices that map objectives, to tasks, to sub-tasks, to further sub-tasks, and even differentiate task applicability based on digital investigative principles (e.g. investigative objective being forensic or non-forensic), device under analysis, target operating system, etc. Users of such databases can then query the database based on the unique parameters of their case (objectives, device under analysis, etc.) and be subsequently guided (not directed) through a series of recommended tasks, sub-tasks, and actions at various stages of the investigation.

IV. CASE SCENARIO IMPLEMENTATION OF THE PROPOSED FRAMEWORK

In this section, two case studies are presented to demonstrate implementation of the proposed framework. One case study is a server intrusion at a private company (non-governmental organization), and the other pertains to a law enforcement investigation for possession of contraband material (child pornography). These two scenarios were selected to highlight the applicability of the proposed framework to both forensics and non-forensics based investigations (the distinction being whether or not presentation of findings in a judicial setting is anticipated). These scenarios were also selected to mirror those selected in Carrier and Spafford's (2003) article discussing their proposed Digital Investigations Process, thereby facilitating framework comparison by the reader. The reader is reminded that these scenarios, although common occurrences, are fictional in nature, and the discussion will remain high-level in nature, due to obvious space constraints.

Commercial Server Intrusion Case Study

As with Carrier and Spafford's (2003) proposed server intrusion scenario, the victim organization is a medium-sized manufacturing company. The company's name is Primo Manufacturing, Inc. The alleged victim system is Primo's primary public DNS server. They became aware that the server was hacked, due to a phone call from another company alleging that Primo's DNS server was scanning other systems on the Internet for systems vulnerable to a secure shell (SSH) vulnerability.

Preparation Phase

Prior to the incident, Primo accomplished the following during the Preparation Phase:

- Conducted a risk assessment and outlined/implemented a risk management plan;
- Determined which logs would be saved for which devices, where they would be stored, and how backups would be accomplished;
- Developed an Incident Response Plan, delineating user, manager, system administrator, etc. roles, responsibilities, coordination, etc.;
- Developed incident response toolkits, including a CD-ROM of trusted tools for the Solaris DNS system;
- Hired a digital forensics consulting company on retainer to handle rigorous forensic data analysis;
- Trained all employees on the Incident Response Plan, assuring all personnel would know how to proceed when incidents are reported or suspected;
- Trained one system administrator (referred to as the Incident Response System Administrator) on incident response procedures, including basic triage and evidence preservation techniques;
- Prepared host and network devices according to the risk management and information retention plan. As a result, logging was enabled on the DNS server with logs sent to a central log server and server time was synchronized via network time protocol (NTP); and
- Developed evidence preservation and handling procedures that (1) maximized evidence preservation while awaiting a management decision concerning the forensic or non-forensic objectives of the investigation, and (2) relaxed evidence handling procedures when a non-forensic investigation is pursued.

In accordance with the proposed framework, Primo considered and applied framework principles while accomplishing these Preparation Phase activities. Information documented during this phase included the risk assessment itself, as well as the risk management decisions made and rationale thereof; configuration management information pertaining to incident response toolkits; the digital forensics consultant contract;

and training data/information. Most importantly, information documented by Primo during this phase included the forensic system data (i.e. system and network activity logs) and the Incident Response Plan.

Primo Manufacturing, Inc. ensured applicable logging was enabled on their firewalls, network-based intrusion detection system (IDS), DNS server, other servers, and desktops. To ensure this data remained available with assured integrity, it was backed up to a well-protected central back-up server. This activity was directed as a result of Primo's Incident Response Plan, which was developed earlier during the Preparation Phase. Primo's Incident Response Plan outlines roles, responsibilities, directed actions, decision criteria, and standards pertaining to investigative principles in various situations (forensic vs. non-forensic, low priority vs. high priority, low sensitivity vs. high sensitivity, etc.). For example, Primo's Incident Response Plan calls for maximum evidence preservation during the Incident Response Phase until a management decision is made regarding the nature of investigation pursued (forensic vs. non-forensic). Then, personnel are directed toward different evidence preservation standards accordingly. In sum, Primo produced an Incident Response Plan capable of providing personnel adequate direction in nearly all conceivable situations. They then disseminated, trained, and tested the plan across their organization.

Incident Response Phase

The Incident Response Phase of this investigation began upon receipt of a call from an external source complaining that a computer with one of Primo's IP addresses was scanning external computer assets for a secure shell (SSH) vulnerability. Primo's goals during this phase were to validate whether or not the allegation was founded, assess the impact of the compromise, engage in necessary containment activities, and determine a response strategy regarding investigation and/or recovery, as determined by management objectives concerning the incident. These goals largely equate to the Incident Response sub-phase structure (Validate → Assess → Contain → Strategize), although a specific sub-phase structure was not proposed in this paper.

The Incident Response Plan became the basic roadmap for all Primo employees to follow during this Phase. The system administrator who received the allegation documented the information accordingly and contacted her manager, as well as Primo's designated Incident Response System Administrator (the one specifically trained in incident response procedures). The Incident Response System Administrator confirmed the veracity of the allegation via network traffic analysis and victim system analysis using his incident response toolkit, doing so in a manner which minimized data alteration on the victim system. He remained mindful of physical investigation considerations, as well as digital investigation considerations, as pointed out in Carrier and Spafford's (2003) server intrusion case study. The Incident Response System Administrator then conducted basic triage in a forensically sound manner to determine the scope of the intrusion (e.g. no sensitive company information was compromised, and the primary DNS server appeared to be the only system compromised). He determined that that the probability was high that the suspect was external to the organization. He also contained the compromise by taking the system off-line (without powering it off) and replacing it with the unaffected secondary DNS server.

The Incident Response System Administrator provided all known information to management. Primo management decided that because no sensitive company information was compromised and the monetary impact of the matter was low, attribution and retribution would not be sought and that a non-forensic investigation was appropriate for the purpose of eradication and recovery. This management decision represented one of the two key pieces of output for the Incident Response Phase. It also affected the production of the second piece of output—the Investigation Plan. The Investigation Plan was informally developed by the Incident Response System Administrator, with input from others as needed, and outlined basic non-forensic investigative objectives, such as:

- Determine the vulnerability exploited (method of compromise);
- Determine how to protect against vulnerability exploitation; and

- Determine if any other systems are vulnerable.

Data Collection Phase

The Data Collection Phase began upon receipt of Primo's management decision regarding the investigative course of action and development of the Investigation Plan. Primo employees collected data with the goals of Investigation Plan in mind and in accordance with standards set forth in the Incident Response Plan. They then determined what data might be needed to support the data analysis goals and identified the data collection targets . Finally, following the non-forensic goals outlined in the Investigation Plan, they collected that data in a manner consistent with the investigation principle standards established in the Incident Response Plan.

In this scenario, data collection targets included: DNS server logs (/var/adm/utmp, /var/adm/wtmp, /var/adm/lastlog, /var/adm/sulog, /var/adm/messages, /var/log/syslog, /var/cron/log, etc.), a list of running processes and installed services, a physical memory dump to capture rogue processes running on the compromised system, file system mapping information to capture file date/time and permission information, and hashes of standard Solaris binaries. All of this data was captured without regard to the forensic integrity of the victim system, because a non-forensic investigation decision was made. In other words, data was copied off of the DNS server in a manner/form convenient to the employee, and a forensic image of the DNS server was not obtained. Data copied off the system was hashed in accordance with Primo's Incident Response Plan, because such actions are non-labor/cost intensive. Although actions were documented, they were not done so as rigorously as if this were deemed a forensic investigation (e.g. providing support for potential future judicial actions).

Data Analysis Phase

During this phase, the examiner conducted the analysis on the compromised DNS server, because again, a non-forensic investigation decision was made. The logs, memory dump, etc. collected during the Data Collection Phase served more as "insurance" in case they were lost or corrupted in some manner while conducting the analysis. The examiner started by surveying the victim system for obvious "landmarks," such as hidden files, rogue processes running, and deleted logs. All of the log files were present and appeared to be intact, however, the examiner discovered one hidden file that contained source code and evidence of a rogue process running on an unauthorized port. The server had not been rebooted in at least a month, and the date/time of the rogue process had been initiated one week prior to the incident. Based on this date/time information, the examiner extracted file system information, focusing on file activity surrounding the time the rogue process began. He also extracted log files. He examined the extracted file system data and logs, focusing on the time frame when the rogue process began, and was able to identify a basic chronology of nefarious activity. It became clear that the hacker escalated privileges on a valid user account using a known buffer overflow vulnerability that has no patch. The examiner cracked the password of the compromised user account and noted it was the same as the user ID, presuming then that the hacker guessed the user account password. Upon extracting hashes of the Solaris system binaries and comparing them with "known goods," the examiner concluded a rootkit had not been installed on the DNS server.

Based on the information obtained during the Data Analysis Phase, Primo's Incident Response System Administrator concluded that the server could be rebuilt using the same processes and procedures used in the past, but that all users and administrators needed to be reminded about safe account creation/handling procedures. Output of the Data Analysis Phase included the conclusions of how the server was compromised and details regarding protective actions needed to better protect Primo resources in the future.

Presentation of Findings Phase

During the Presentation of Findings Phase, Primo's Incident Response System Administrator presented relevant findings in a variety of ways to relevant audiences. He provided the Chief Information Officer with an oral report of the findings and recommendations for action. He followed that up with a written memorandum for record, which included information about the compromised account, who created it, and who was responsible for maintaining and/or deleting it. In both reports, he clearly outlined the (limited) damage to Primo Manufacturing, Inc. and recommended immediate and long-term protective actions. He then prepared a more technical report for the information technology team, which emphasized lessons learned and remedial actions, but also provided technical detail on the exploit, rogue process, etc. for the team's educational purpose.

Incident Closure Phase

Based on the information provided during the Presentation of Findings Phase, Primo management decided to have the IT staff rebuild the DNS server and place it back on line, ensuring all accounts were valid and properly protected. Additionally, Primo management decided to direct a forced password change on all user accounts organization-wide and mandate remedial information/computer security training for all employees. These decisions were implemented during the Incident Closure Phase. Additionally, the Incident Response System Administrator coordinated a critical review of the entire process and investigation to identify and incorporate lessons learned. Finally, Primo collected all information pertaining to the incident and investigation and stored it in accordance with their Incident Response Plan.

Law Enforcement Contraband (Child Pornography) Case Study

The commercial server intrusion case study discussion above demonstrates how the phases and principles of our proposed framework apply toward a non-forensic investigation. Conversely, the law enforcement contraband (child pornography) case study will naturally assume a forensic investigation is pursued.

The law enforcement contraband case study scenario is modeled after Carrier and Spafford's (2003) law enforcement contraband scenario and is synopsized as follows: Upon investigating a web server that contained child pornographic images, law enforcement officials identified several potential suspects, including Mr. Smith, who paid fees to download contraband images from the web server. During the physical investigation, law enforcement officials correlated the member information to financial records, Internet Service Provider subscriber records, etc. to identify Mr. Smith as a suspect and his place of residence as the site from which he downloaded the contraband images.

Preparation Phase

During the Preparation Phase, prior to this particular investigation, the law enforcement agency conducted agent training, response planning, technical capability development, and evidence preservation handling procedure development. Similar to the commercial server intrusion case study, all investigation principles were considered and incorporated into applicable agency plans and policy documents, which serve as the primary output of this phase for law enforcement agencies.

Incident Response Phase

Again, the Incident Response Phase began with notification. In this scenario, the law enforcement agency identified an individual, Mr. Smith, suspected of possessing and/or distributing contraband (child pornography), and they subsequently notified their computer crime investigators. The agents engaged in information gathering activities to facilitate coordination of legal authority and response. For example, the agents obtained all known information about Mr. Smith (e.g. identification, criminal history, firearm registration, residence location and description, etc.), including any known information about his computing resources and skill (e.g. Internet connectivity mechanism, number of computers, operating systems, etc.). The agents then coordinated and obtained legal authority (e.g. search warrant) in concert with "physical crime scene investigators."

Upon receipt of the proper legal authority, the Incident Response Phase concluded with a proper search of Mr. Smith's residence. The law enforcement agents contained the digital crime scene by disconnecting

Mr. Smith's cable modem from the PC in order to prevent potential data alteration. They properly documented the scene from both a physical crime scene and a digital crime scene perspective. They used various forms of camera/video photography, sketches, and note taking to document all relevant crime scene information. Specific information documented during the search pertaining to the "digital crime scene" included equipment information and cable connections.

Based on the information known from the physical investigation and media found during execution of the search warrant, the agents developed an Investigation Plan. Because this type of contraband investigation is handled frequently by the agency, the Investigation Plan remained informal (non-written) and was based on previous investigative experience with similar cases. This Investigation Plan, together with the legal authority (i.e. search warrant), comprised the output for this phase.

Data Collection Phase

During the Data Collection Phase, agents seized all potentially relevant digital media in a forensically sound manner. (Unlike the commercial server intrusion, non-forensic investigation case study, no analysis on the original media would be conducted.) Digital media seized in this case included one computer with one hard drive installed, several CD-ROMs, and several Post-It™ notes (with logins, passwords, and website addresses). Because the computer was already powered off when the agents arrived, there was no need to photograph the computer monitor and pull the plug from Mr. Smith's computer (since it was a Windows™-based, non-server system). Finally, because the search warrant permitted seizure outright without obtaining additional evidence, the computer and other digital media were seized and brought back to the agency computer crime laboratory for proper forensic imaging, hashing, and analysis (triage and imaging were not conducted in the field). The Data Collection Phase concluded in the agency's laboratory with the forensic imaging and hashing of the media seized.

Data Analysis Phase

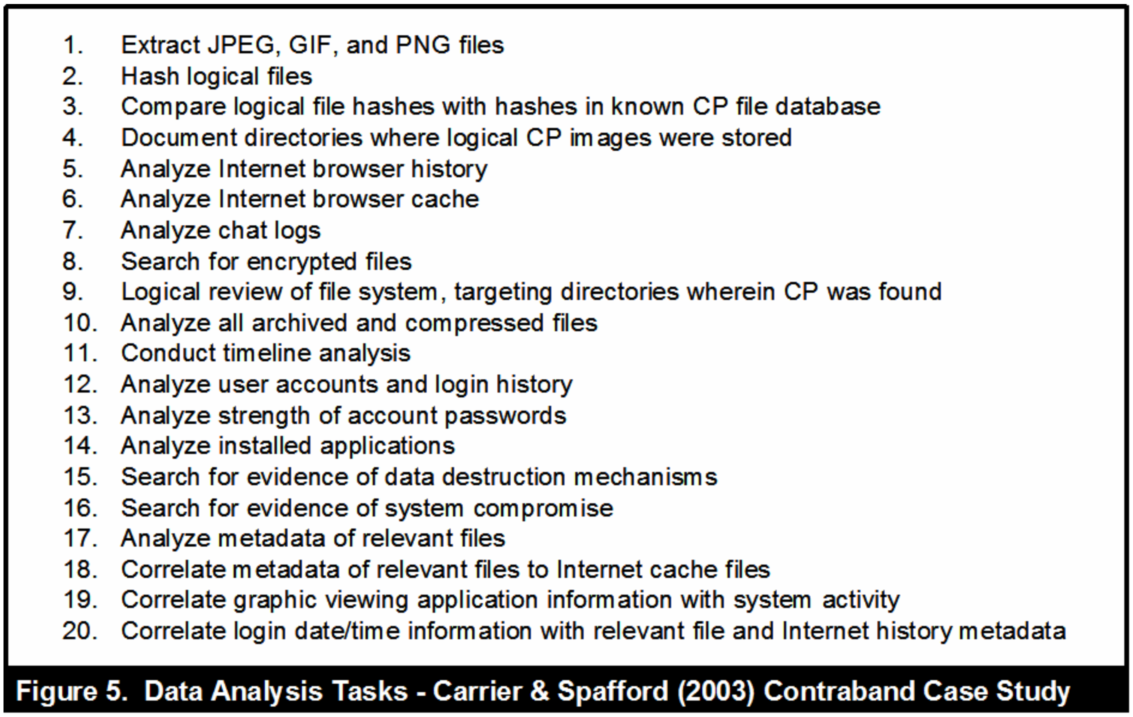
During the Data Analysis Phase, the law enforcement agents surveyed the "digital landscape" of the digital objects seized, extracted potentially relevant data (e.g. graphic images, file hashes, Internet cache and history, etc.), and examined the extracted data from both confirmatory and event reconstruction perspectives.

In this scenario, the investigator's cognitive burden was alleviated by having the opportunity to select investigation objectives from a reasonably sized list of potential data recovery objectives. Such investigation objectives included:

1. Recover contraband images
2. Attribute possession to individual(s)
3. Demonstrate knowledge of possession and distribution
4. Reconstruct events regarding possession (time, method, etc.)
5. Confirm or refute possible defenses

Upon selecting these five data recovery objectives, subsequent data survey, extraction, and examination tasks were illuminated—presumably the same twenty as in Carrier and Spafford's (2003) law enforcement contraband scenario (shown in Figure 5). The difference in this scenario is the objectives-based approach and its decreased cognitive burden on the part of the investigator, as well as a potentially decreased error rate associated with forgetting an important data analysis task.

The Survey, Extract, and Examine Sub-Phases were accomplished iteratively. As with all phases, investigation principles were applied accordingly. For example, copious notes were taken pertaining to actions taken, software used, data extracted, and information uncovered in support of the documentation principle. The entire examination was conducted on forensically validated working copies and/or images to support the evidence preservation principle in a forensic investigation scenario.



1. Extract JPEG, GIF, and PNG files
2. Hash logical files
3. Compare logical file hashes with hashes in known CP file database
4. Document directories where logical CP images were stored
5. Analyze Internet browser history
6. Analyze Internet browser cache
7. Analyze chat logs
8. Search for encrypted files
9. Logical review of file system, targeting directories wherein CP was found
10. Analyze all archived and compressed files
11. Conduct timeline analysis
12. Analyze user accounts and login history
13. Analyze strength of account passwords
14. Analyze installed applications
15. Search for evidence of data destruction mechanisms
16. Search for evidence of system compromise
17. Analyze metadata of relevant files
18. Correlate metadata of relevant files to Internet cache files
19. Correlate graphic viewing application information with system activity
20. Correlate login date/time information with relevant file and Internet history metadata

Figure 5. Data Analysis Tasks - Carrier & Spafford (2003) Contraband Case Study

In this scenario, the agents recovered known and suspected child pornographic images from logical files of Mr. Smith's user account, as well as Internet cache and unallocated space. They discovered print spool images of known and suspected child pornography, matching the filenames and images found in Mr. Smith's user account. They correlated the downloaded images with connectivity to the web server originally investigated. They did not find any evidence to support the distribution allegation.

Presentation of Findings Phase

During the Presentation of Findings Phase, the computer crime investigators orally briefed the primary (physical crime scene) investigators about the data analysis findings. The agents then produced a report of findings pertaining to the digital investigation. The report included a synopsis of the items analyzed and support requested, a summary of findings, a detailed discussion of findings, and a technical glossary. The primary investigators then incorporated the digital crime investigation report into the overall case report, as applicable, and presented the case to the appropriate legal authorities (e.g. Assistant U.S. Attorney).

Incident Closure Phase

During the Incident Closure Phase, legal actions were taken. Mr. Smith was arrested, indicted, and convicted of possession of child pornography. The agents disposed of the evidence based on direction from legal authorities and permanently archived the case file and all supporting material. Finally, the agents discussed and incorporated lessons learned to improve their policies and procedures.

Framework Application to Forensic and Non-Forensic Investigations

The two case studies presented show how the proposed framework and its digital investigation principles remain intact, while its implementation flexes, depending on the Investigation Plan and its investigative objectives. Table 4 below demonstrates how the proposed framework flexes to support both forensic and non-forensic investigations within the same scenario (the commercial server intrusion case study).

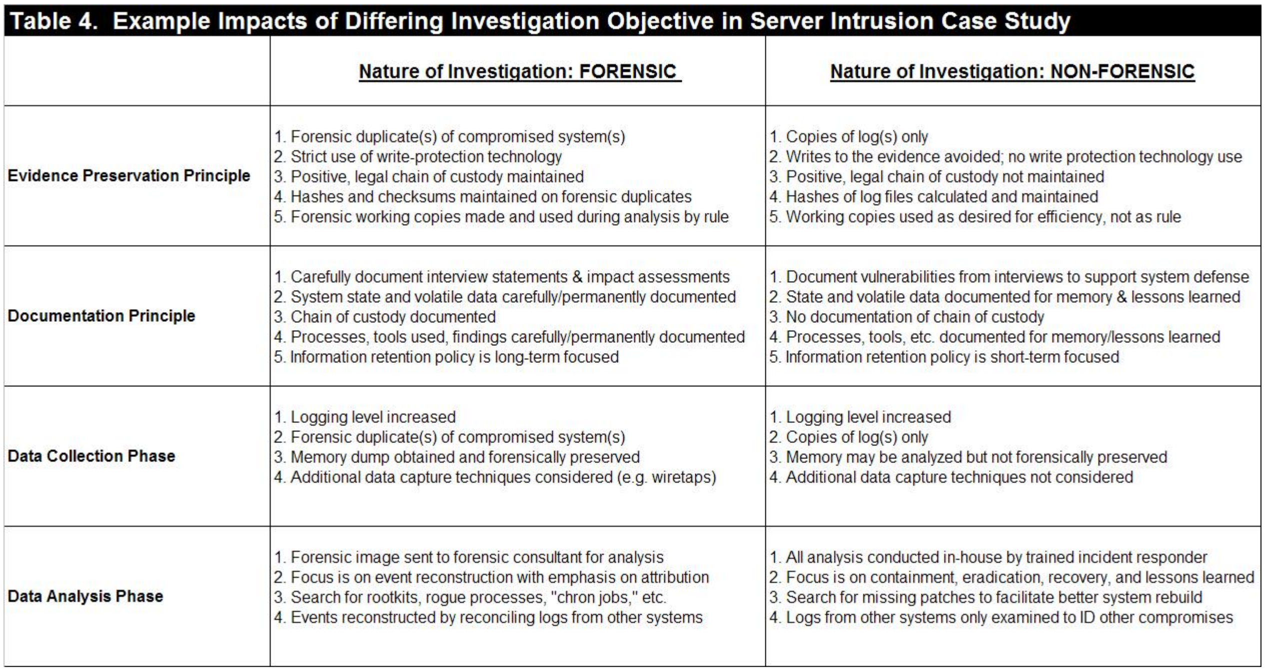


Table 4. Example Impacts of Differing Investigation Objective in Server Intrusion Case Study

	Nature of Investigation: FORENSIC	Nature of Investigation: NON-FORENSIC
Evidence Preservation Principle	1. Forensic duplicate(s) of compromised system(s) 2. Strict use of write-protection technology 3. Positive, legal chain of custody maintained 4. Hashes and checksums maintained on forensic duplicates 5. Forensic working copies made and used during analysis by rule	1. Copies of log(s) only 2. Writes to the evidence avoided; no write protection technology use 3. Positive, legal chain of custody not maintained 4. Hashes of log files calculated and maintained 5. Working copies used as desired for efficiency, not as rule
Documentation Principle	1. Carefully document interview statements & impact assessments 2. System state and volatile data carefully/permanently documented 3. Chain of custody documented 4. Processes, tools used, findings carefully/permanently documented 5. Information retention policy is long-term focused	1. Document vulnerabilities from interviews to support system defense 2. State and volatile data documented for memory & lessons learned 3. No documentation of chain of custody 4. Processes, tools, etc. documented for memory/lessons learned 5. Information retention policy is short-term focused
Data Collection Phase	1. Logging level increased 2. Forensic duplicate(s) of compromised system(s) 3. Memory dump obtained and forensically preserved 4. Additional data capture techniques considered (e.g. wiretaps)	1. Logging level increased 2. Copies of log(s) only 3. Memory may be analyzed but not forensically preserved 4. Additional data capture techniques not considered
Data Analysis Phase	1. Forensic image sent to forensic consultant for analysis 2. Focus is on event reconstruction with emphasis on attribution 3. Search for rootkits, rogue processes, "chron jobs," etc. 4. Events reconstructed by reconciling logs from other systems	1. All analysis conducted in-house by trained incident responder 2. Focus is on containment, eradication, recovery, and lessons learned 3. Search for missing patches to facilitate better system rebuild 4. Logs from other systems only examined to ID other compromises

V. FRAMEWORK DISCUSSION

Benefits of the Proposed Framework

The primary goals of any framework should be to:

- Achieve scientific rigor and relevance;
- Simplify complex processes to facilitate understanding of the underlying structure;
- Retain enough granularity, or the flexibility to incorporate granularity needed to exploit the framework in unique situations; and
- Delineate standard assumptions, concepts, values, and practices.

The proposed hierarchical, objectives-based framework achieves these goals. Rigor is achieved through the use of a phased approach and the inclusion of important principles. Relevance is achieved through the use of objectives-based sub-phases and the ability to infuse the framework with objective-task matrices. The hierarchical nature of the proposed framework allows complex process simplification by allowing its users to conceptually focus in on higher ordered tiers. At the same time, granularity is facilitated by the framework's ability to expansively include multiple layers of detail. Finally, the proposed framework delineates standard assumptions, concepts, values, and practices through the use of constraints, definitions, principles, objectives, and task hierarchies.

Carrier and Spafford (2003) presented a five-point requirement set by which proposed digital investigative process models and frameworks may be further judged. These requirements are summarized as follows:

- Basis in existing physical crime scene investigation theory;
- Practicality—matching steps taken in actual investigations;
- Technology neutrality to ensure the process isn't constrained by current products and procedures;

- Specificity to facilitate technology requirement development; and
- Applicability to all possible user communities.

The first-tier framework, which consists of Preparation, Incident Response, Data Collection, Data Analysis, and Incident Closure Phases, as well as the SEE Data Analytical Approach, which consists of Survey, Extract, and Examine Sub-Phases, both leverage lessons learned over time from the physical crime scene investigation process. Principles such as data preservation and documentation permeate all phases of the investigative process. Confirmatory analysis goals are differentiated from event reconstruction analysis goals. Finally, Investigative approach strategy development is a function of investigative objective selection.

The proposed digital investigative framework offers unique benefits over previously proposed frameworks in the areas of practicality and specificity. Previously proposed frameworks lack detail, and/or an apparent path to incorporate the level of detail needed to achieve practicality for investigators. This inadequate level of detail also hinders requirements development and gap analysis activities by researchers and tool developers. While the present effort is admittedly incomplete in that it only presents an initial sub-phase structure for the Data Analysis Phase, it proposes a multi-tier, hierarchical, objectives-based framework that facilitates the addition of multiple layers of detail to achieve the needed levels of practicality and specificity. A framework that can handle the infusion of layers of detail will ultimately guide the practitioner regarding “how to” and “where to” find digital evidence. It will also facilitate the mapping of tool capabilities to evidence recovery objectives (again, via matrix development), thereby facilitating gap analysis and directing subsequent research and development efforts.

The proposed framework also fulfils the remaining two requirements—technology neutrality and wide user community applicability. Technology neutrality is relatively easy to achieve. The lack of neutrality in some models introduced to date is predominantly a function of the lack of concerted effort to apply scientific principles and develop a digital investigations process for the field. The latter requirement, wide user community applicability, is evident in the proposed first-tier framework. The challenge is in developing sub-phases that meet the needs of all potential user communities. We argue that the SEE Data Analytical Approach and objectives-task matrices are applicable to all user communities, including law enforcement, industry, military, and non-military government communities. Each community’s needs are served by having the ability to determine its own set of objectives for both data recovery and principle attainment, potentially unique to each situation. The framework is able to flex based on each unique set of objectives.

The framework’s amenability to the development of helpful matrices provides further credence of its flexibility. Matrices that correlate tasks to objectives, tools to tasks, and capabilities to tools can be easily developed. This will provide practitioners, researchers, developers, and legal community members with an efficient means to:

- Clearly delineate a data analysis approach strategy and keep analytical objectives at the forefront;
- Identify analytical steps to take to achieve data analysis objectives;
- Establish analytical process standards and guidelines;
- Determine which tools can be used for various data analysis tasks;
- Identify and track which tools have been scientifically tested;
- Track margin of error associated with tools; and
- Determine where research and development is needed.

Limitations of the Proposed Framework

As stated previously, the proposed set of objectives is arguably incomplete and is proposed to stimulate additional discussion and development within the digital forensic science and practitioner community. This initial effort was predominantly focused on traditional, main-stream computer and network forensics in which hard drives are the primary digital device analyzed. Additional work is needed to ensure the proposed model is applicable across various layers of abstraction (Carrier 2003) and to other digital devices, such as personal data assistants (PDAs), digital cameras, telephones, and removable data storage devices. Collaborative inputs from both academic and practitioner communities are needed to increase robustness and rigor of the proposed framework.

It is also conceivable that platform (i.e. operating system) specific renditions of the task hierarchy may be needed. In the case of analyzing hard drives, a significant area of emphasis for digital forensics, tools, techniques, and procedures vary widely between file systems and operating systems. This again supports the argument for creating an objectives-based digital investigations process framework. Still, once objectives are identified, task hierarchies might be enhanced via additional levels of specificity unique to file systems, operating systems, etc.

Finally, the proposed framework would benefit from additional case studies—hypothetical and/or actual—with a methodical approach to identifying objectives and task hierarchies. Doing so would optimize objective and task development efforts. Additionally, it would facilitate scenario development to help users, researchers, and tool developers understand how to leverage and apply the proposed framework.

VI. CONCLUSION

Digital investigations, whether forensic in nature or not, require scientific rigor and are facilitated through the use of standard processes. Because such processes are inherently complex when developed with wide applicability in mind (independence of technology, user, or objective), the proposed framework was developed. The first tier structure represents a simple, easy to grasp framework with wide applicability. While framework simplicity represents one side of a proverbial two-sided coin, granularity and specificity represent the other side. The framework must facilitate the infusion of detail required by its users, be they practitioners or researchers. The proposed framework's emphasis on sub-phase and objective-task hierarchical structures represents the ability to infuse such detail. This makes the framework both usable and flexible. Previously proposed frameworks are predominantly single-tier, higher order process models that focus on the abstract, rather than the more concrete principles of the investigation. We contend that these frameworks, although useful in explaining overarching concepts, fail to support the inclusion of additional layers of detail needed by various framework users.

Throughout the course of this paper, we introduced the reader to our proposed framework, including its phases and principles. We presented a typical sub-phase structure and example objective-task matrix and demonstrated its utility by focusing on their application to the Data Analysis Phase. We introduced two case studies that demonstrate the synergy of the proposed framework with those proposed previously, and that show how it would realistically be applied in both forensic and non-forensic investigations. Finally, benefits of the proposed framework and digital investigation process philosophy were outlined, while acknowledging the framework's limitations.

ACKNOWLEDGEMENTS

The authors wish to thank attendees of the Digital Forensic Research Workshop 2004, who provided invaluable feedback on an earlier version of this paper. Additionally, the authors wish to thank members of the commercial industry, law enforcement, and the U.S. military for their input, including Mr. Kevin Mandia, Mr. Bob Renko, Captain Adam Fraser, Special Agent Paul Alvarez, and Special Agent John Wood. Finally, the authors wish to thank the anonymous reviewers of this manuscript.

REFERENCES

Carrier, Brian "Defining Digital Forensic Examination and Analysis Tools Using Abstraction Layers.," *International Journal of Digital Evidence* (1:4), Winter 2003, pp 1-12.

Carrier, Brian, and Spafford, Eugene H. "Getting Physical with the Digital Investigation Process," *International Journal of Digital Evidence* (2:2), Fall 2003, pp 1-20.

Casey, Eoghan *Digital Evidence and Computer Crime - Forensic Science, Computers and the Internet* Academic Press, Cambridge, 2003a, p 265.

Casey, Eoghan (ed.) *Handbook of Computer Crime Investigation*. Academic Press, London, 2003b.

Casey, Eoghan, and Palmer, Gary L. "The Investigative Process," in: *Digital Evidence and Computer crime,* E. Casey (ed.), Elsevier Ltd., 2004.

DoJ "Electronic Crime Scene Investigation - A Guide for First Responders," U.S. Department of Justice, pp. 1-82.

Kruse, Warren G., and Heiser, Jay G. *Computer Forensics - Incident Response Essentials* Lucent Technologies, Indianapolis, 2002, p 398.

Mandia, Kevin, Prosise, Chris, and Pepe, Matt *Incident Response & Computer Forensics*, (Second ed.) McGraw-Hill/Osborne, Emeryville, 2003, p 507.

Mohay, George, Anderson, Alison, Collie, Byron, Vel, Olivier de, and McKemmish, Rodney *Computer and Intrusion Forensics* Artech House, Boston, 2003, p 395.

Nelson, Bill, Phillips, Amelia, Enfinger, Frank, and Steuart, Chris *Guide to Computer Forensics and Investigations* Thomson Learning Inc. - Course Technology, Canada, 2004, p 689.

Ó Ciardhuáin, Séamus. "An Extended Model of Cybercrime Investigations," *International Journal of Digital Evidence* (3:1), Summer 2004, pp 1-22.

Palmer, Gary L. "A Road Map for Digital Forensics Research - Report from the First Digital Forensics Research Workshop (DFRWS) (Technical Report DTR-T001-01 Final)," Air Force Research Laboratory, Rome Research Site, Utica, pp. 1-48.

Palmer, Gary L. "Forensic Analysis in the Digital World," *International Journal of Digital Evidence* (1:1), Spring 2002, pp 1-6.

Reith, Mark, Carr, Clint, and Gunsch, Gregg "An Examination of Digital Forensic Models," *International Journal of Digital Evidence* (1:3), Fall 2002, pp 1-12.

Rowlingson, Robert "A Ten Step Process for Forensic Readiness," *International Journal of Digital Evidence* (2:3), Winter 2004, pp 1-28.

Vacca, John R. *Computer Forensics - Computer Crime Scene Investigation* Charles River Media, Inc., Hingham, 2002, p 731.

ABOUT THE AUTHORS

Nicole Lang Beebe (nbeebe@utsa.edu) is a Research Assistant at the University of Texas at San Antonio, where she is working on her PhD in Information Systems. Previously, she was a Senior Network Security Engineer with the Science Applications International Corporation (SAIC), where she conducted commercial digital forensics investigations and information/network security vulnerability assessments for government and commercial customers. She has been a federally credentialed computer crime investigator for the Air Force Office of Special Investigations (AFOSI) since 1998 (Reservist since 2001). She is a Certified Information Systems Security Professional (CISSP), an EnCase Certified Examiner (EnCE), and holds degrees in electrical engineering and criminal justice.

Jan Guynes Clark (jgclark@utsa.edu) is a Professor at the University of Texas at San Antonio, which is a National Security Agency (NSA) designated Center of Academic Excellence. Dr. Clark is a Certified Information Systems Security Professional (CISSP), has a Ph.D. in Information Systems, and numerous publications on a variety of information systems topics.