January 17, 2012

From: James Darnell, Chair SWGDE
To: John Paul Jones II
Executive Secretary
RDT&E IWG

Subject: RDT&E IWG Letter to SWGDE

Dear John Paul Jones:

This letter is in response to the list of questions received on October 26th, 2011 regarding published literature in regard to digital evidence analysis. SWGDE viewed many of the questions asked by the SoFS as subcategories of broader categories and were combined as such. To assist in mapping the original questions with our answers, all questions are listed before our answers. Additionally, when document abstracts were used for descriptions, the text is displayed in *italics*.

1. **What literature exists that describes the underlying assumptions and propositions in computer forensics and how relevant scientific fields are used to develop forensically correct extraction and examination procedures?**

2. **What literature exists that describes the extraction and analysis of evidence regarding criminal attacks on computer and network systems?**

5. **What literature exists that describes how computer forensics extraction software and hardware are selected for particular tasks, and how examiner judgment plays a role in this selection?**

13. **What literature exists that describes technical specifications for various hardware and software used in computer forensics?**

15. **What literature exists that describes the potential for corruption or cross-contamination of data in a digital environment, and how this is affected by choice of computer forensic tool?**

17. **What literature exists that describes detection of alterations/tampering of storage media, files or file metadata using computer forensics tools?**

SWGDE believes assumptions and propositions include the following:

- We view a computer as a device to store information as ones and zeros and the role of the
  examiner is to extract information represented thereby
- Ones and zeros are repeatedly and reliably discerned regardless of how they are accessed
- Hashing is a reliable method of ensuring files are exact copies and the file content has not changed due to examiner interaction
    - Hashing is consistent across all types of media
- Hardware, software, and firmware does not materially affect the integrity of the ones and zeros
    - Even in the presence of hardware (storage device) failure, ones and zeros can be reliably extracted
- File headers and footers are reliable characteristics of specific file types that allow them to be carved

In forensic science, evolving technologies dictate constant updates to examiner methods and tools. However, digital forensic science differs in that changes in technology alter the entire subject landscape (not to mention the tools and techniques required to address the new technology) and constantly introduces new variables that result an intolerably large set of permutations that is feasible to scientifically address.

Assumptions in digital forensics have not been subject to the rigor of a scientific vetting process. While reference literature does exist for discrete tools within different sub-disciplines, i.e. imaging, carving, etc., there has yet to be exhaustive treatment of these underlying assumptions. The underlying science does not dictate how the application of that science materially affects the outcome of that application. SWGDE suggests that the scientific method employed by examiners allows for reliable and repeatable results.

SWGDE would welcome any recommendations for scientific research projects that could provide the level of scientific rigor required to enhance the foundations of the digital forensic discipline.

**3. What literature exists that describes databases and other reference material fundamental to computer forensics practice, and how are these maintained?**

SWGDE is unaware of any exhaustive treatment of databases and other fundamental reference material. There are several databases, such as the NSRL database cited below, which provide examiners with information critical to timely and accurate retrieval of evidence from digital media.

**NIST (created in 2000, updated quarterly). National Software Reference Library. National Software Reference Library (NSRL) Project. Available at http://www.nsrl.nist.gov/index.html**

The National Software Reference Library (NSRL) provides a repository of known software, file profiles, and file signatures for use by law enforcement organizations in computer forensics investigations.

In addition, there are numerous laboratory and vendor created and maintained databases that signify detectable behavior pattern such as kernel access, credit card skimming, etc. Depending on the source, many of the databases may not be shared across the industry.

**4. What literature exists that describes standards for interoperability of forensic tools, allowing different vendor products to interoperate?**

There are no existing standards. There is constant discussion in the community as to the need for standardization.

**Carrier, Brian (2003). Open Source Digital Forensic Tools: The Legal Argument. Available at http://www.digital-evidence.org/papers/opensrc_legal.pdf**

A related document that supports the use of open source tools that lends them to interoperability.

**6. What literature exists that describes how and/or when computer forensics examiners interpret recovered digital media, such as identifying objects or identifying specific persons in images, video, or audio? What literature exists that describes additional examiner qualifications needed for this purpose? What literature exists regarding statistical models used to estimate the certainty of conclusions?**

**7. What literature exists that describes the validity of interpretations that result from computer forensics analysis?**

Automated tools exist in this area; however, the validity of such tools is tested independently without formal peer-reviewed processes and publications.

Regarding published literature, SWGDE knows of no scientific documents that specifically address human and automated interpretation or the training and/or additional qualifications necessary to make such interpretations.

Automated identification tools lend themselves to statistical validation and SWGDE would welcome any research projects to lend more scientific rigor to this analysis.

**8. What literature exists that describes sources of error in computer forensics attributable to human error, including perceptual and cognitive factors such as expectation bias and fatigue?**

**9. What literature exists that describes the error rates of computer forensics examiners, such as in proficiency testing or blind examinations? This includes how factors such as experience or training influence performance.**

**14. What literature exists that describes the types of errors that can be made by various computer forensic tools and the accuracy of such tools?**

There are no known studies on human error rates in the digital forensics discipline.

As for tools, several organizations publish tool testing reports on a regular basis. These reports include error characterization.

Tool testing portals include the following:

**NIJ Computer Forensics Tool Testing (CFTT) Project Web Site located at** http://www.cftt.nist.gov/.

The goal of the Computer Forensic Tool Testing (CFTT) project at the National Institute of Standards and Technology (NIST) is to establish a methodology for testing computer forensic software tools by development of general tool specifications, test procedures, test criteria, test sets, and test hardware.

**National Repository for Digital Forensic Information (NRDFI) located at** http://www.nrdfi.net.

NRDFI hosts digital forensic tools and validations of those tools that allow for their use in law enforcement.

**10. What literature exists that describes the training and education requirements for computer forensics practitioners, and whether these requirements are different for particular tasks in different areas of computer forensics, including the investigatory vs. laboratory functions of computer forensics?**

**Brinson, A., Robinson, A., & Rogers, M. (2006). A cyber forensics ontology: Creating a new approach to studying cyber forensics. *Digital Investigation*, 3, S37-S43.**

*The field of cyber forensics, still in its infancy, possesses a strong need for direction and definition. Areas of specialty within a professional environment, certifications, and/or curriculum development are still questioned. With the continued need to standardize parts of the field, methodologies need to be created that will allow for uniformity and direction.*

*This paper focuses on creating an ontological for the purpose of finding the correct layers for specialization, certification, and education within the cyber forensics domain. There is very little information available on this topic and what is present, seems to be somewhat varied. This underscores the importance of creating a method for defining the correct levels of education, certification and specialization. This ontology can also be used to develop curriculum and educational materials. This paper is meant to spark discussion and further research into the topic.*

**ASTM E2678 (2009). Standard Guide for Education and training in Computer Forensics. Available on the ASTM website www.astm.org.**

Developed from a NIJ publication, this guide describes qualifications for a career in computer forensics; career paths for practitioners; structures of associate, baccalaureate, and graduate degree programs; a structure for academic certificate programs; and the training and continuing education of practitioners.

**Irons, A.D., Stephens, P., & Ferguson, R.I. (2009). Digital Investigation as a distinct discipline: A pedagogic perspective. *Digital Investigation*, 6 (1-2), 82-90.**

*Is Digital Investigation sufficiently different in character from existing academic disciplines such as Computer or Forensic Science to be called a distinct discipline? Is it a profession in its own right? The authors outline why the debate is a significant one in terms of its consequences for professional standards, quality control, academic and personal accreditation. The paper emphasizes the differences in the way we teach digital investigations in comparison to computer science covering theory, practice, the education versus training debate, the interdisciplinary nature of the subject, a problem solving and problem based approach, and the need to emphasis professionalism and ethics. The arguments for four alternative positions are proposed: Digital Investigation as a branch of Computer Science, Digital Investigation as a branch of Forensic Science, Digital Investigation as an inter-disciplinary science and Digital Investigation as a distinct discipline. The experience gained in the development and delivery of three typical academic programs in the area is used to support one position, namely that Digital Investigation is a distinct discipline that merits professional status.*

**SWGDE/SWGIT (2010). Guidelines & Recommendations for Training in Digital & Multimedia Evidence. Available on the SWGDE website at www.swgde.org/current-documents.**

These guidelines assist in the design of an agency level training program. It defines different categories of training (awareness, skills, knowledge), and different categories of job function (manager, examiner, technician, first responder). For each job function category, it elaborates on the knowledge and skill set required. For examiners, discipline-specific guidance is provided.

**SWGDE (2011). Audio Core Concepts. Available on the SWGDE website at www.swgde.org/current-documents.**

This document provides an outline of the technician level knowledge and abilities all practitioners of forensic audio should possess. The provided elements are intended as a basis for training and testing programs (certification, competency, and proficiency). The skills set intentionally mirrors the tasks defined in SWGDE's *Best Practices for Forensic Audio*. The document does not describe the additional skill sets specific to more advanced analysis, such as enhancement, complex media repairs, or signal analysis.

**Ballou, S. & Gilliland, R. (2011). Emerging paper standards in computer forensics. *Digital Investigation*, 8 (2), 96-97.**

The authors, members of ASTM Committee E30 on Forensic Science, describe different types of standards related to computer forensics being developed to help scientists perform their work effectively and describes the work of the American Society for Testing and Materials (ASTM) in this area.

**Allegra E., Di Pietro, R., La Noce, M., Ruocco, V., & Vincenzo Verde, N. (2011). Cross-border co-operation and education in digital investigations: A European perspective. *Digital Investigation*, 8 (2), 106-113.**

*Co-operation and education are fundamental issues when dealing with national and international organizations involved in digital forensic investigations. Although these two aspects are often separately handled, they are strictly connected. On the one hand, different agencies can leverage on co-operation for the training of their investigators while, on the other hand, co-operation is possible only if an adequate level of education on digital forensic matters is reached. In this paper, the concrete outcome of a complete training program that involved several European antitrust agencies is reported (named EAT_FIT, European Antitrust Training in Forensic IT). We sum up the activities and the techniques that are generally used in antitrust investigations, and we outline the rationales used to set up such a training course. Assessment data collected both during and after the training highlight the needs and the difficulties faced by the digital forensic practitioners working in the field.*

11. **What literature exists that describes what constitutes a validated tool in computer forensics, and how such tools are tested?**
12. **What literature exists that describes the risks of utilizing unvalidated tools in computer forensics?**

**NIST (2001). General Test Methodology for Computer Forensic Tools, Version 1.9. Computer Forensics Tool Testing (CFTT) Project. Available from www.cftt.nist.gov.**

NIST's stated goal in this paper is to provide a measure of assurance for the software tools used by law enforcement in computer forensics investigations. This document provides a description of the general approach taken to develop the test methodology for computer forensic tools and the rationale behind this approach. NIST cites as a complicating factor the lack of standards and specification describing what forensic tools should do and the need for these tools to survive the scrutiny of a judicial process. The CFTT effort is supported by SWGDE and its individual members. One of SWGDE's goals is to see this Project provide a means to reduce the burden on individual labs to perform their own testing, by defining validation requirements suitable for the entire computer forensics community.

**Lyle, J. (2002). NIST CFTT: Testing Disk Imaging Tools. *Proceedings of the Digital Forensics Research Workshop 2002*. Available from www.dfrws.org/2002/program.shtml.**

The author, a senior member of NIST's CFTT Project, describes the validation methodology NIST uses for testing disk imaging tools.

**Beckett, J. & Slay, J. (2007). Digital Forensics: Validation and Verification in a Dynamic Work Environment. *Proceedings of the 40th Hawaii International Conference on System Sciences*, 6, S12-S22.**

*Many forensic computing practitioners work in a high workload and low resource environment. With the move by the discipline to seek ISO 17025 laboratory accreditation, practitioners are finding it difficult to meet the demands of validation and verification of their tools and still meet the demands of the accreditation framework. Many agencies are ill-equipped to reproduce tests conducted by organizations such as NIST since they cannot verify the results with their equipment and in many cases rely solely on an independent validation study of other peoples' equipment. This creates the issue of tools in reality never being tested. Studies have shown that independent validation and verification of complex forensic tools is expensive and time consuming, and many practitioners also use tools that were not originally designed for forensic purposes. This paper will explore the issues of validation and verification in the accreditation environment and propose a paradigm that will reduce the time and expense required to validate and verify forensic software tools.*

**Carrier, B. (2002). Defining Digital Forensic Examination and Analysis Tools. *Digital Forensics Research Workshop*, Syracuse, NY.**

*This paper documents the use of abstraction layers in digital forensics.*

**SWGDE (2009). Recommended Guidelines for Validation Testing, Version 1.1. Available on the SWGDE website at www.swgde.org/current-documents.**

This document describes the critical need for validation testing of tools before use on casework. It outlines why to validate, when to validate, and a process for doing so. It provides a test plan template, a test scenario report template, and examples of each to help implementers.

**Garfinkel, S., Farrell, P., Roussev, V., & Dinolt, G. (2009). Bringing science to digital forensics with standardized forensic corpora. *Digital Investigation*, 6, S2-S11.**

The authors describe the need for standardized forensic corpora to establish ground truth for methods under test, to allow direct objective comparison of different methods, and to establish reproducibility both for validation and for advancement of methods. They also go into detail about the modalities, sensitivity, issues around IRB use restrictions, and metadata of the proposed corpora. They then review some existing corpora to illustrate the pros and cons of each. The authors end with a plea for scientific rigor in digital forensics and describe the goals for education, testing, and research that standardized corpora will bring.

**Guo, Y., Slay, J., & Beckett, J. (2009). Validation and verification of computer forensic software tools—Searching Function. *Digital Investigation*, 6, S12-S22.**

The authors point out the differences in various definitions of "validation" and "verification" (VV) from different sources related to software development and computer forensics (or "electronic evidence" as they call it). They also acknowledge that the dynamic evolution of technology leads to a dynamic evolution of evidence from these devices, which in turn leads to a dynamic evolution of the tools used to collect that evidence. They warn that tools are not always made for a forensic purpose and make the distinction between validating the results of a tool (black box testing) and validating the tool itself. The authors mention the challenging VV work being done at NIST, Carrier, SWGDE, and by tool vendors. They also underscore a key point that tests by vendors are not publicly documented (presumably to protect proprietary code or trade secrets) and generally cite repeatability from other tools as proof of validation without acknowledging that two tools may both be wrong. The authors declared motivation is to move the focus of VV from specific evidence ("Does this tool work on this device?") to a higher level theory of digital forensic analysis from which VV requirements manifest. The authors propose a new high-level VV framework and apply their framework to the task of data searching, developing requirements and a corresponding reference test set test can be applied to any tool that performs searches.

**Lyle, J. (2010). If *error rate* is such a simple concept, why don't I have one for my forensic tool yet? *Digital Investigation*, 7, S135-S139.**

*The author identifies critical issues that distinguish errors in the digital domain from those in the physical world. In particular, errors in physical processes, including forensic tests, are commonly associated with random variables and characterized statistically by precision and accuracy. In computer forensics, many errors tend to be systematic in nature repeated testing will return the same result (high precision), rightly or wrongly. The author describes other possibilities where statistical analysis may play a role, but any error rates developed are likely to be a function of many dynamic variables, such as the operating system, which will change over time thereby nullifying the applicability of the computed error rate. The author identifies three broad sources of error: the algorithm intended for the process, the software implementation of the algorithm, and the performance of the process by a person. The author's conclusions include that a general error rate may not be meaningful in computer forensics and to consider the source of error the algorithm or the implementation.*

**Beckett, J. & Slay, J. (2011). Scientific underpinnings and background to standards and accreditation in digital forensics. *Digital Investigation*, 8, 114-121.**

*With its use highlighted in many high profile court cases around the world, Digital forensics over the last decade has become an integral part of the modern legal system and corporate investigations. As the discipline grows and it use becomes widely accepted, there is a need to align it with traditional forensic sciences and move towards strengthening an accreditation regime for the discipline. This paper examines the origins of science and scientific method to form the core premises for establishing criteria to assess digital forensics as a science and hence justifying the basis for standards and accreditation.*

**Garfinkel, S. (2010). Digital forensics research: The next 10 years. *Digital Investigation*, 7, S64-S73.**

*Today's Golden Age of computer forensics is quickly coming to an end. Without a clear strategy for enabling research efforts that build upon one another, forensic research will fall behind the market, tools will become increasingly obsolete, and law enforcement, military and other users of computer forensics products will be unable to rely on the results of forensic analysis. This article summarizes current forensic research directions and argues that to move forward the community needs to adopt standardized, modular approaches for data representation and forensic processing.*

**Casey, E. (2011). The increasing need for automation and validation in digital forensics. *Digital Investigation*, 7 (3-4), 103-104.**

Although automation is a necessary part of today's digital forensic paradigm, automated digital forensic tools must provide sufficient transparency to enable investigators to catch errors and omissions in automated processes.

**Carrier, Brian (2010), *The Digital Forensic Tool Testing Project, http://dftt.sourceforge.net***

*This website takes the approach of fabricating datasets with known features that are thoroughly documented to facilitate testing and validation.*

**Erin Kenneally, Gatekeeping Out Of The Box: Open Source Software As A Mechanism To Assess Reliability For Digital Evidence, http://www.vjolt.net/vol6/issue3/v6i3-a13-Kenneally.html**

*This article examines digital evidence reliability by first identifying and differentiating the two competing categories of software from which this evidence is derived: proprietary and Open Source. The next section explores the standards for software reliability in both the industrial marketplace and the legal arena. Specifically, the current standards are addressed in light of their value to industry and the law, as well as their respective historical origins This sets the stage for a reconciliation of standards for reliability as between industry and the courtroom.*

**Brian Carrier, Open Source Digital Forensics Tools: The Legal Argument, http://www.digital-evidence.org/papers/opensrc_legal.pdf**

*This paper addresses open source digital forensic analysis tools and their use in a legal setting. To enter scientific evidence into a United States court, it must be reliable and relevant. The reliability is tested by applying Daubert guidelines. This paper examines the guidelines and shows that open source tools may more clearly and comprehensively meet the guidelines than closed source tools would.*

**16. What literature exists that describes the effectiveness of computer forensics tools and procedures on degraded or corrupted digital media?**

**Jim Lyle, Issues with imaging drives containing faulty sectors, http://www.dfrws.org/2007/proceedings/p13-lyle.pdf**

This paper documents several experiments using non-commercial imaging tools and their behavior when encountering faulty sectors on a hard drive.

**18. What new technologies and areas of research should be pursued with regard to computer forensics? (Note- this question does not require a list of references, it is for informational purposes only.)**

New technologies that represent challenges in digital forensics with little published scientific research include the following:
Cloud forensics,
Volatile memory forensics,
Live network triage,
Insider threat, and
Gaming systems

SWGDE thanks you and the RDT&E IWG Co-Chairs for this opportunity to relay information concerning our field. Please do not hesitate to contact us for further information.

Sincerely,

James Darnell
Chair SWGDE