Public Comments Received

on NISTIR 8354-DRAFT
Digital Investigation Techniques: *A NIST Scientific Foundation Review*

Published July 18, 2022

NISTIR 8354-DRAFT: Digital Investigation Techniques: *A NIST Scientific Foundation Review* was released for public comment on May 9, 2022. That draft document is available at https://nvlpubs.nist.gov/nistpubs/ir/2022/NIST.IR.8354-draft.pdf.

A public comment period was held from May 9, 2022, to July 11, 2022. This document lists all 15 public comments in the chronological order in which they were received. Submitter email addresses and phone numbers have been redacted.

NIST hosted a webinar on June 1, 2022, to review the content of the draft report and address questions. A recording of the webinar can be found at https://www.nist.gov/news-events/events/2022/06/webinar-digital-investigation-techniques-nist-scientific-foundation. The 24 questions/comments received during the Q&A portion of the webinar are included in the public comments as PC7.

NIST NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
U.S. DEPARTMENT OF COMMERCE

Public Comment #	Commenter Name & Affiliation
1	Yuri Gubanov (Belkasoft)
2	Dan Mares (Retired)
3	Lou Giannelli (Skyborg Cyber Security SME)
4	Nina Sunde (Norwegian Police University College)
5	Maxim Suhanov
6	Matt Bergin (Infosec)
7	Webinar Q&A (24 comments received)
8	Maxim Suhanov
9	Dawn M. Ryan (Anne Arundel Community College)
10	Graeme Horsman (Cranfield University)
11	Constantinos Patsakis (University of Piraeus) Nikolaos Mantas (independent DFIR researcher)
12	Institute of Electrical and Electronics Engineers (IEEE)
13	National Association of Criminal Defense Attorneys, Fourth Amendment Center
14	The Innocence Project
15	Daniel Kahn Gillmor (ACLU)

PC1

From: Yuri Gubanov
Date: Thu, May 12, 2022 6:00 AM -0400
To: ScientificFoundationReviews <ScientificFoundationReviews@nist.gov>
Subject: Digital Investigation Techniques: A Scientific Foundation Review - Draft report comments

Hello,

A couple of comments on the report:

1. You mentioned software write blockers, though did not mention it is not the best way to write blocking. In our article https://belkasoft.com/5-bloopers-of-digital-forensic-investigator we mention an article on what may go wrong with software way
2. There is nothing about SSD specifics when write blocking. SSDs cannot be fully protected neither with software not even hardware blockers
3. Sad not to see Belkasoft in the list of software tools tested. There are tools which do not exist anymore, like Lantern, but no Belkasoft, which is top-3 DFIR software as per last 4 year Forensic 4:cast awards
4. For deleted data recovery from SQLite: it is not just "records which are marked as deleted". For example, there is unallocated space in SQLite, which can be carved for deleted data.
   See https://belkasoft.com/sqlite-forensics-with-belkasoft-x for more ways to recover SQLite deleted data. I believe, there is a pending report on Belkasoft SQLite recovery review made by NIST, which you can refer to
5. There are more than just 3 ways to recover deleted data. Recycle Bin analysis and analysis of file system snapshots are also widely used. For the latter, previous versions of snapshots may contain data which is deleted in later snapshots and in the current state of a drive

Hope this helps.

Regards,
Yuri
Belkasoft

PC2

From: Dan Mares
Subject: NISTIR 8354-DRAFT comment
Date: 13 May 2022 13:34
To: Lyle, James R. (Fed)

Hey, long time no see, hear, yell, etc.
How you doing? retired yet?

anyway, i got hold of this document:
NISTIR 8354-DRAFT
Digital Investigation Techniques:
A NIST Scientific Foundation Review

and decided to put my $.02 in.
attached is a response to this document with some of my comments.
unfortunately i made it a real patch work, so you will have to forgive that
i bounced all around with my comments. but all the comments target or center around
just a few items pertaining to processing possible forensic evidence, and a thought that
one or all of these 4 items would give a defense attorney a lot of ammunition to attack.

1. properly finding and identifying (cataloging, listing) the items of interest based on specific file structure(s) where you must understand that not all instances will allow a full bit-image to be made of the evidence

2. properly hashing the original information for good evidence identification and cataloging.
similar to what a physical investigator might do to catalog, list, and maintain integrity of the evidence
hashing is addressed in the article, but not targeted specifically at an initial capture of the state of the evidence.

3. properly copy the identified evidence without altering the original, and when laying it down at the analysis site, properly restoring its total tree, and MAC stuff.

4. properly preserving (zipping; a fancy way of copying,etc) the final product for posterity.

These 4 items should be basic for any test of software. Can it find, catalog, identify, hash, preserve.
No matter what type of case: graphic, theft, porn, text, phone, etc.
All the items mentioned above seems like they should be basic tests for test platforms and software.

That being said, i checked the sample data sets in the https://cfreds.nist.gov/all location, and could not find a simple full tree to test any or all of the above 4 in a simple test set designed for the major file systems: ie: EXT2, NTFS, FATxx etc.

Anyway, hope reading my item will not send you into a deep sleep.

Take care.

--
Dan Mares
Retired

Items from the document are where possible identified by italic text.

My comments are normal style.

My executive summary:

Too much stress is placed on processes down the list as opposed to some of the most basic first steps that should be accomplished. These first steps:
List, Catalog, Identify,
Hash
Copy

Should be placed within the tool testing minimal requirements.

Take a look at four references and evidence of importance to all of my comments:
https://www.dmares.com/maresware/articles/copy_that.htm
https://www.dmares.com/maresware/articles/hash_it_out.htm
https://www.dmares.com/maresware/articles/ZIP_IT.htm
https://www.dmares.com/maresware/articles/list_it.htm

==========
==========
Checked: https://cfreds.nist.gov/all for sample files which might test a simple cataloging and hashing of a tree file system. Couldn't find one specifically setup for that test.

==========

In your executive summary:

Item 2. Acquire digital data. This is accomplished by copying data to make an image file of the acquired digital data. Copying digital data accurately is based on established engineering techniques such as error detecting and correcting codes to ensure that data is copied accurately. This is discussed in Sec. 4.2.

Item 3. Ensure integrity of acquired data. Cryptographic hashing is used to ensure that if acquired digital data is changed inadvertently or deliberately, the change can be detected. This is discussed in Sec. 4.3

Comments: Especially within item 2, acquire digital data, even though it is not part of the acquire process, a pre-step should be the identification or (cataloging/listing) of all the data which you wish to acquire. The document in most instances pre-supposes that a complete bit-image will be done. However in the real corporate world the full bit-image of a multi-terabyte server may not always be possible, and thus a single tree directory structure may be the only source of the evidence. No matter what file system, this single tree environment was not considered in almost all of the sections.

Executive Section 5. Page 2 Navigate the acquired digital data.

This is accomplished by unraveling, i.e., parsing the layout of the acquired data. This is best performed using a software tool. There is the risk that an incorrect implementation will not correctly interpret the structure of a particular file system, e.g., not showing all acquired active files. This is discussed in Sec. 4.5.

Item 7: ***KEY TAKEAWAY #4.2:*** *Searching tools ….*
You mention searching tools. Although a lot of analysis is conducted using searching tools, you can't really search data unless you know how much and where it is. This searching takes place only after you have a good idea of the data to search, and have possibly copied and hashed its contents so that your hashing tool does not corrupt the process. Nowhere have I seen significant references to this initial data identification and preservation process where your searching tools are mentioned.

Item 11 ***Key Takeaway section 4.6****: page 4 It is not feasible to test all combinations of tools and digital evidence sources.*

However, no tests or mention of tests to conduct a very basic process of determining a basic listing and/or hash of ALL the files within the suspect system. This should be a first step in any test or analysis.

Other Statements:

Section 2 items:

2.6 File Systems (page 17)

In this section, it doesn't appear that any information or mention is made of the specific items which may be accessed or seen by the normal user. In particular, the items that may prove to be evidentiary in nature are the number of files, the locations of the files, the hashes, the MAC times. These items are probably the most basic of any file system, and should be at a minimum considered a basic test or item to accumulate and record. After all, part of collecting evidence, no matter what type of case, it so catalog and record all the evidence. This file systems section, if it mentions it at all, doesn't put much emphasis on the collection and recording of the state of the evidence (files) at the time of seizure. The bit image software packages that do a forensic bit image "image" retain that information, but consideration should be given to any file system/hardware setup where a full image cannot be done, and the catalog, hash, copy of ONLY those areas identified as possible evidence storage locations can be accessed.

The identification, cataloging, hashing and copying of this initial evidence should be a primary concern in any file system. It doesn't appear to be stressed as much as a defense attorney may challenge that process.

Might I suggest that in the https://cfreds.nist.gov/all portal where sample data is provided you could add a few simple items. Mainly sample data from the most popular file systems; IE: FAT32, NTFS, EXFat, EXT2. These sample could have specific files and structures unique to that file system so that when a person uses a forensic tool to test, they have a sample file system with unique items from that suspect system. For instance, EXT2 might have significant long filenames, along with NTFS. While with the NTFS section you add alternate data streams, and 0 byte files. Items which software programs that are designed to process those file systems may have troubles finding, identifying, listing and processing (hashing) those types of files specific to that file system.

If you check my test pages listed at the end of this document, you will see that many of the test software I used failed a significant amount of NTFS specific file system capability. This type of test might be set up for other commonly found file systems. After all, if a forensic tool can't find a long filename, it might be challenged easily in court.

And finding, listing, capturing MAC dates, hashing and forensically copying of the files should be among the FIRST processes a forensicator should perform. After all, aren't these items among the basic "evidence" identification and capture routines.

Key Takeaways section 4.2 etc: Does nothing about mentioning the reliability of copying evidence outside of the image, and/or reliably finding/listing all available files within the current file structure.

In *Table 1.2: Page 7*: Under the Digital World column:
the item: "Files stored on the computer hard drive, removable media"

Suggests that an operation should be able to find, list, and process those stored files. However, no test were seen or found within the document which indicated that finding and listing the files within the hard drive was one of the "basic" tests which could/should be conducted.

On page 7 of the report, the chart:

Acquiring (or gaining access to) the digital data.
• *Ensuring the integrity of the data.*
• *Reconstructing and recovering deleted artifacts.*
• ***Identifying relevant artifacts.***
• ***Extracting relevant artifacts***.
• *Classifying relevant artifacts.*
• *Assembling a narrative of what happened.*

However, the bullet item: Identifying relevant artifacts, and Extracting relevant artifacts
Should basically mean: cataloging/listing the files, and copying/extracting the appropriate files to a work environment. Neither of these two items seem to be fully covered in the test processes.

==========

All throughout the document pages 9, etc. the discussion talks about

*"**In the end, the job of the digital examiner is to use tools to find relevant information from digital evidence"***
However, again, nowhere is the most simple of requirements of the tools to find, list, catalog, copy any relevant data.

==========

Again, on page 11, p.2:
Early in the development of digital computers the need for reliability was recognized and the means to ensure reliable data transfer was developed. Transferring data within a computer has to be extremely reliable because "in a digital computer ... a single failure usually means the complete failure [that] . .

"reliable data transfer" , copy from source to destination, again, doesn't seem to be covered within the document or testing processes.

==========

Chapter 4 of the document:

Figure 4.1: The collection steps ensure the integrity of the acquired data to provide a stable source for the analysis of the data.

Ensure the integrity of the data requires accurate and "complete" hashing of ALL the data on the specific file system. Your tests stuites relating to hashing do not allow for coverage of all available data on some of the major (NTFS) file systems.

These two sections:

2. Acquire digital data. This is accomplished by copying data to make an image file of the acquired digital data. Copying digital data accurately is based on established engineering techniques such as error detecting and correcting codes to ensure that data is copied accurately. This is discussed in Sec. 4.2.
3. Ensure integrity of acquired data. Cryptographic hashing is used to ensure that if acquired digital data is changed inadvertently or deliberately, the change can be detected. This is discussed in Sec. 4.3

Both above are basic for a reliable investigation, and should be included in the tool testing CFTT environment. Section 4.2 talks about this requirement but is not part of any section of CFTT.

Section 4.3 **Integrity Verification**
validating via hash:

You have published hash standards, but have not provided hashing CFTT basic data sets.

"Searching tools" are profusely mentioned. But accurate searching on only accomplished after you have properly identified, hashed, and "copied" all relevant files/data to a reliable safe work environment.

Section 4.8
4.8 Verification of Techniques and Validation of Tools

Deals, mentioned profusely the hashing techniques, but again fails terribly at finding/cataloging and properly copying in a forensically sound manner to obtain the files necessary for further analysis. The selection of a valid algorithm is necessary, but if you don't know which/what files you are to hash, then the evidence cannot properly be validated.

Section 4.10.1 Error Rates
Error rates are fine, but there are some items which cannot and should not accept even a minimal error rate. Such as listing files within the tree. If an acceptable error rate is used, one may miss that one file which will prove the case. The same argument can be made for "copying" necessary evidence files for analysis. An acceptable error rate maybe should be 0 missed files.

The error rates acceptable for these two items should only accept those items which are ONLY technically feasible with extreme low level tampering of the file system. Normal file visibility and accessibility thru the basic file system :IE: windows, should require a 100% accurate response.

Key takeaway 4.4 addresses this subject minimally, but not enough emphasis is placed upon an acceptable error for straight forward file system processing of the files. There should be no "**random**" errors when say, a program fails to find and list, or hash a hidden file outside of a traditionally protected system file which normal use would not allow access to.

Key takeaway 4.6*: page 47: It is not feasible to test all combinations of tools and digital evidence sources.*

True it is not feasible. But some basic file system requirements should be minimal based on the file system being processed. IE: process hidden files, for NTFS, process long file names, for NTFS, process alternate data streams. For other cell phone, or linux OS's there may be other simple file system files that should always be found, processed, and available for analysis. This requirement should be made part of all tool tests. Meaning some minimal requirements that everyday users with a minimal technical knowledge should be able to see, look, and feel that file.

4.10.4 TOOL TESTING RESULTS

One of the items is Text String searching.
Even though this may be an important test, as mentioned before, I would suggest that finding/cataloging/listing, copying, ALL the files should be a BASIC requirement, and then the absolute capability of HASHING all those files should be a requirement well before Text Strings search.

Finding, cataloging, listing, copying those items of interest should be part of the list. And any, and all of these items are not listed at all.

You may say, disk imaging, is sufficient to satisfy the copying requirement. However, in a lot of large corporate places where multi terabyte servers are involved, the disk image is not possible, and alternate identifying and copying processes should be identified.

4.10.5: NIST test data sets for tool testing:

Again, the continuously mention of write blockers, and forensic image is used for the data sets, and string searches is used. But the simple basic example data sets for file lists, hash, copy is not mentioned enough or stressed its importance.

Optional features are mentioned, but those “optional features” are never highlighted.

PC3

From: GIANNELLI, LOUIS M CTR USAF AFMC AFRL/AFRL
Sent: Tuesday, May 17, 2022 8:58 AM
To: ScientificFoundationReviews <ScientificFoundationReviews@nist.gov>
Subject: NISTIR 8354-DRAFT DIGITAL INVESTIGATION TECHNIQUES

NIST Special Programs Office - Scientific Foundation Review

NIST draft states: This document is an assessment of the scientific foundations of digital forensics. [Abstract]

My comment:
What is the primary audience of this assessment? The language used throughout the NIST draft is non-technical, and contains basic topical information for a non-technical audience. How is this review an assessment on digital forensics, written in a non-technical language, and a content designed to explain fundamentals well-known to the digital forensic community? Therefore, I have to ask again: who is the audience of this review? The language and scope of this review is too basic for the digital forensic community.

Lou Giannelli, CISSP
Applied research Solutions
Skyborg Cyber Security SME
AFRL/RYAR

PC4

From: Nina Sunde
Date: Mon, May 23, 2022 8:39 AM -0400
To: ScientificFoundationReviews <ScientificFoundationReviews@nist.gov>
Subject: Comment to NISTIR 8354-DRAFT

To: National Institute of Standards and Technology
Date: 23 May 2022

Comment to:

NISTIR 8354-DRAFT
Digital Investigation Techniques: A NIST Scientific Foundation Review
(James R. Lyle, Barbara Guttman, John M. Butler, Kelly Sauerwein, Christina Reed, Corrine E. Lloyd)

Comment:
The report acknowledges the challenge of (cognitive) bias (p. 40) and concludes that is a need for better understanding of bias (p. 56). Although I agree with the conclusion that more research is needed, I suggest that the review points to the research that already exists on contextual bias in digital forensic investigations:

> Sunde, N., Dror, I. E. (2019). Cognitive and human factors in digital forensics: Problems, challenges, and the way forward. *Digital Investigation*, 29, 101-108. https://doi.org/10.1016/j.diin.2019.03.011.
>
> Sunde, N., Dror, I. E. (2021). A Hierarchy of Expert Performance (HEP) applied to Digital Forensics: Reliability and Biasability in Digital Forensics Decision Making. *Forensic Science International: Digital Investigation, 37,* 301175. https://doi.org/10.1016/j.fsidi.2021.301175.

The research provided insight into the effects of contextual information on DF practitioner decision-making, and indicated that not directly task-relevant contextual information influenced the amount of traces discovered by the DF practitioners (Sunde and Dror, 2021). In addition to "better understanding of bias" (p. 56) there is a need for research on measures that minimize cognitive bias while ensuring that the DF practitioner has access to the necessary information to perform a targeted and effective investigation. I therefore suggest the following adjustment to the last bullet point under section 5 (p. 56):

- Better understanding of bias and effective bias minimization measures. Because of the nature of most digital evidence case work, forensic examiners are exposed to knowledge about people involved in the case, such as seeing their photos and reading their text messages. In addition, the forensic examiner may need to interact with an investigator.

Sincerely,

Nina Sunde
Police Superintendent
Department for Post Graduate Studies
The Norwegian Police University College

PhD student
Department of Criminology and Sociology of Law
Faculty of Law
University of Oslo

PC5

From: Maxim Suhanov
Sent: Friday, May 27, 2022 6:11 AM
To: ScientificFoundationReviews <ScientificFoundationReviews@nist.gov>
Subject: Comments on NISTIR 8354-DRAFT

Hello.

Please find attached my comments on NISTIR 8354-DRAFT.

Comments on “Digital Investigation Techniques: A NIST Scientific Foundation Review” (NISTIR 8354-DRAFT)

Glossary and Acronyms

Page	Text	Comments
iv	**Advanced Format** Created to address technical issues with the 512-byte storage device sector size by changing storage device sector size from 512-bytes to a multiple of 512-bytes such as 4096-bytes, i.e., storage devices with a sector size larger than 512-bytes.	Not all storage devices with a sector size larger than 512 bytes are *Advanced Format* (e.g., DVDs are not *Advanced Format* despite their 2048-byte sector size).
v	**Disk Imaging** The process of acquiring the digital contents of a storage device (fixed disk, removable disk, flash drive, etc.). This acquires all the data on a device including files, metadata, and contents of unallocated areas of the device.	Usually, disk imaging involves copying data which is exposed to a host (but not all exposed data is often copied – e.g., S.M.A.R.T. values can be skipped). Some data stored on a device is not exposed to a host (this includes raw data from service areas and blocks used for overprovisioning), thus not copied by most disk imaging tools (it is impossible to read such data using standard ways of communicating to a storage device). The definition can be changed to reflect this.
vi	**File System** A method for organizing files on a storage device. Common file systems on Windows systems are NTFS, ExFAT and FAT. LINUX systems use ext4 and FAT. Apple Macs use HFS+, APFS, FAT and ExFAT.	Ext2 is still popular in the Linux world (it is often used for volumes that do not require journaling, like “*/boot*”).

vi	**Fixed media** A storage device that is physically installed in a computer.	1. Two definitions (of “Fixed media” and “Removable media” respectively) are not clear enough. According to them, it is possible for removable media to become fixed once it is attached to a computer port (it becomes *“physically installed in a computer”*). 2. Are all hot-pluggable drives removable? Are USB flash drives removable? Are USB HDD/SSD enclosures removable? Although many types of storage devices can be hot-plugged, not all of them are considered as removable (e.g., SATA HDDs). The current definitions (of “Fixed media”, “Removable media”, and “Storage Device”) do not cover this case. 3. Also, there is an issue with USB devices. USB flash drives are removable, because they can be attached and removed while the computer is running. But USB flash drives have non-removable media, because they do not contain swappable memory chips. Even most USB HDD/SSD enclosures have non-removable media, because their internal drives cannot be swapped on the fly (without resetting the enclosure, so they require an initiator-target nexus loss event). As a result, USB flash drives must report their removable medium bit as 0 (but most of them do not). But if this bit is correctly set to 0, some popular operating systems treat such a storage device as non-removable (see SanDisk Answer ID 12830: https://kb.sandisk.com/app/answers/detail/a_id/12830), because they see no difference between removable devices with non-removable media and devices with removable media. For further discussion, see “SPC-6: Removable Medium Bit Expectations” (https://www.t10.org/cgi-bin/ac.pl?t=d&f=20-082r1.pdf). So, the current definitions can be adjusted to define removable (and non-removable) storage devices, as well as removable (and non-removable) media.
ix	**Removable media** A storage device that is either (1) a data container that is inserted and removed from a data reader or (2) a storage device that can be connected or removed from a computer while the computer is running.	
ix	**Storage Device** *[…]* * Fixed media physically installed in a computer. The computer must be powered off to install or remove the storage device. * Removable media. Can be installed or removed while the computer is running. Small storage devices are called flash drives or thumb drives (they are about the size of a human thumb). These devices are usually connected via a USB interface. *[…]*	
viii	**Metadata** Metadata is a description of stored data. Categories of metadata	There are more metadata types than listed in the definition (e.g., RAID metadata – this type of metadata describes how to assemble the array of physical drives to get one or more virtual drives).

	include: (1) application metadata (in a document this could be author, organization, etc., in a database such as SQLite there is metadata to describe the layout of the stored data within the database), (2) file system metadata (placement of the file within the file system, owner, permissions, MAC times, etc.), (3) partition metadata that identifies the type of file system the partition contains and global file system parameters, and (4) device metadata describes the layout of partitions on a device.	So, the phrase *"Categories of metadata include"* can be changed to *"Categories of metadata include, for example"* (or to something similar).
viii	**NTFS** New Technology File System. Microsoft Windows file system introduced in 1993, revised several times over the years.	The NTFS file system has been revised many times, so the word "several" is misleading ("multiple" is more suitable). Modern revisions do not increment the format version number, though. For example, the following features have been implemented in the past years (all of them affect the on-disk format, but do not change the version number of the whole file system): - energy-efficient metadata logging (requiring less writes per logged operation); - per-directory case sensitivity; - reparse point security (trust levels); - storage reserve (reserved file system areas and file tags); - very large clusters (up to 2M bytes).
viii	**Operating System** The software that creates the digital environment for running software on a computer or other digital device. Most operating systems are variants of either MS Windows (95, 98, 2000, Vista, XP,	Although many operating systems have gone away, the phrase "[m]ost operating systems are variants of *[two operating system families]*" is still not suitable here. Embedded systems, feature phones, various hardware devices often use other operating systems (e.g., Azure RTOS ThreadX and MOCOR).

	10, etc.) or UNIX (BSD, Linux, Mac OS, iOS, etc.).	
viii	**Partition** A contiguous area of a storage device used to contain a formatted file system.	Partitions do not necessarily contain a formatted file system. Other types of data that can be stored in a partition: - a swap space (occupying a whole partition); - a software RAID volume (which provides a layer between a partition on a physical drive and a file system) or a volume created using similar technology (e.g., Linux LVM). Operating systems can reserve some disk space for their internal use by creating an unformatted partition. Also, operating systems can merge two or more partitions into one virtual partition (by providing a way to split data between multiple existing partitions), this is another layer between a partition on a physical drive and a file system.
viii	**Partition Table** A table describing the layout of a physical storage device that has been divided into partitions, each partition contains a separate file system.	
ix	**Storage Device** An electronic or optical device that can store data for later retrieval. *[…]*	This definition does not include magnetic tape cartridges (they are still used for backups).
ix, x	**UICC card** A Universal Integrated Circuit Card (also called a SIM card) contains phone number and account information for mobile devices. An integrated circuit card that securely stores the international mobile subscriber identity (IMSI) and the related cryptographic key used to identify and authenticate subscribers on mobile devices.	Not all SIM cards store a phone number. Also, “UUIC card” stands for “universal integrated circuit *card card*”.
x	**Write Blocking** Techniques designed to prevent any modification to digital	The write blocking techniques cannot prevent modifications performed by a storage device itself. So, write blocking does not prevent “any modification”.

	media during acquisition or browsing.	For example, write blocking does not stop file-system-aware SSDs from reclaiming unallocated space of supported file system types. Also, write blocking does not prevent modifications not exposed to a host (like wear leveling, which changes the layout of data on the physical layer, while data returned to a host remains the same).

Executive Summary

Page	Text	Comments
4	11. **KEY TAKEAWAY #4.6**: It is not feasible to test all combinations of tools and digital evidence sources.	It is practically impossible to test all combinations of tools and source data (different file system types and their states on storage devices, as well as different types of underlying media). But different types of storage devices can produce unexpected inconsistencies during acquisitions. So, various types of storage devices have to be carefully selected for a specific test (e.g., both HDDs and SSDs should be tested against an acquisition tool). *More information about this takeaway is provided below (see page 12 of this document).*
4	12. **KEY TAKEAWAY #4.7:** Extensive tool testing of over 250 widely used digital forensic tools showed that most tools can perform their intended functions with only minor anomalies.	It is unclear what constitutes a minor anomaly. Some clarification is needed in the executive summary. *More information about this takeaway is provided below (see page 19 of this document).*

Chapter 2

Page	Text	Comments
11	*[…]* and each partition is formatted with a selected file system including layout of file placement on the device	A structure describing which data blocks are used by a given file belongs to file system (volume) metadata; this is not partition metadata. Partitions should not be confused with file systems (volumes), a partition can contain a file system, but it is not equal to a file system (also, it is possible to create a file

	in partition metadata, and any stored data.	system on a storage device without using a partition table, such a file system starts at LBA of 0).
14, 15	**2.3 Time** Times and dates can often exhibit subtle nuances that are prone to misunderstanding. *[…]*	There are several remarkable issues worth mentioning. 1. Local timestamps can be ambiguous, because two different UTC timestamps can point to the same local date and time values due to time folding (when clocks go back). Here is an example: `$ TZ='Europe/Stockholm' date -d @1667092000 '+%F %T'` `2022-10-30 02:06:40` `$ TZ='Europe/Stockholm' date -d @1667088400 '+%F %T'` `2022-10-30 02:06:40` Printing the time zone explicitly can help mitigate the problem: `$ TZ='Europe/Stockholm' date -d @1667092000 '+%F %T %Z'` `2022-10-30 02:06:40 CET` `$ TZ='Europe/Stockholm' date -d @1667088400 '+%F %T %Z'` `2022-10-30 02:06:40 CEST` Still, there are cases when clocks go back without changing the time zone identifier. 2. Time zones are affected by political issues. People living in a specific area can use a different time zone than specified officially. Or two time zones can coexist. Also, applying time zone offsets to timestamps in the future is dangerous (nobody knows which time zone offset will be in effect by the time).
16	For example, a partition table is a configuration file that describes the layout of a storage device.	A partition table is not a configuration file. Popular partition table formats (e.g., MBR and GPT) store their metadata in blocks (sectors) outside of any file systems (and, thus, outside of files).

18	LINUX systems use ext4 and FAT.	Ext2 is still popular in the Linux world.
20	*[…]* TRIM for SATA devices *[…]*	The Trim command is not specific to SATA drives, it could be sent to USB drives too (if the SCSI ATA PASS-THROUGH command is supported by a USB drive or an SSD enclosure).
23, 24	The primary sources of knowledge about digital forensic techniques are the following: * Vendor-Independent Forensic Technique training classes * Tool Vendor offered classes * Forensic Tool Vendor white papers and other support documents * Forensic Professional Organizations * Standards Organizations * Online training videos * Blog Posts * Academic Peer Reviewed papers in Conferences and Journals * Academic course work * Reference books * Operating system and computer hardware vendors support documents * Reverse engineering of software: Operating system, file system or application	1. There are more primary sources of knowledge: - source code of operating systems, drivers, applications (including leaked source code), debugging symbols (the latter partially overlaps with reverse engineering, though); - podcasts; - talks at conferences. 2. On page 23, “[o]nline training videos” are listed. On page 27, this item is called “[o]nline [v]ideos”. Based on the context, “online videos” is a better description. *Similarly, there are other items called differently in the list (on pages 23-24) and in the description (on pages 24-29), they will not be explicitly mentioned here.*
27	**2.8.6 Online Videos**	
29	**2.8.11 Software Developer Documentation** Documentation about operating system internal organization is often available and provides a rich source of information about trace artifacts	It should be noted that such documentation can be misleading, wrong, or out of date. Software developers can forget to update the documentation once the implementation has changed. Or software developers can document the design (or even the design expectations), but not the actual implementation.

	that may be of forensic value. Documentation of individual applications and sometimes the source code of the applications might be available.	

Chapter 4

Page	Text	Comments
34	The acquired data is placed into a container file that represents the acquired data.	It is possible to clone a storage device. In this case, data from a source storage device is copied to a destination storage device directly (no file containers are used). Data on the source storage device located at LBA of 0 will be copied to LBA of 0 on the destination storage device, data on the source storage device located at LBA of 1 will be copied to LBA of 1 on the destination storage device, etc.
35	Whenever possible, acquisition should be done in conjunction with either a hardware write blocking device or a software write blocking tool to avoid modification of the original data.	There is an approach called “quasi software write blocking” – when an operating system is built not to write to attached storage devices. Examples of such operating system builds include Windows FE and Helix3 Pro. In this case, there is no component that inspects I/O requests to block unwanted writes (like a write blocking driver), but the operating system is designed to work in the “read-only” mode (most Linux distributions fail here, but the concept still exists). (More information can be found here: https://github.com/msuhanov/Linux-write-blocker/tree/master/research.)
35	**4.2.2 Mobile Device Acquisition** For mobile device forensics, there are many considerations and options for acquiring and analyzing data from a mobile device (SWGDE 2016a, 2016b, 2019b):	1. File system acquisitions can be full (all allocated files and directories are copied when a mobile device is rooted or jailbroken, either temporarily or permanently) or partial (many files and directories are not copied in this case because of security permissions). 2. Many mobile devices support memory cards. Memory cards can be acquired too (typically, they are

	* Logical acquisition: *[…]* * Selective acquisition: *[…]* * File system acquisition: *[…]* * Physical acquisition: *[…]* * Universal Integrated Circuit Card (UICC), also called a Subscriber Identity Module (SIM Card) acquisition: *[…]*	removed from a mobile device and imaged separately). This option can be added to the list.
36	After digital data has been acquired to an image file it needs to be verified that the acquired data has not been changed. Cryptographic hashing is used to detect inadvertent or deliberate changes.	Cryptographic hashing alone does not prevent deliberate changes that involve updating the hash value to reflect the changed data (i.e., it is possible to change the acquired image file, calculate the hash over altered data, and then specify it as the original hash value recorded during the acquisition, or it is possible to alter data on a source storage device and then acquire it again). This is why cryptographic hash values should be recorded in a way that protects them from alteration (e.g., by maintaining a chain of custody or by using digital signatures). Simply storing cryptographic hash values along with acquired data is not enough (this is why evidence containers like E01 are not “self-authenticating”, they only protect against inadvertent changed and data transfer errors).
37	If the deleted data has been overwritten or allocated to a new object, the deleted data cannot be recovered.	This is true, but such a deleted object can be found in another source (e.g., in a volume shadow copy).
43	The general validation and verification for a given version of a tool can be done once. It does not need to be performed by every lab.	1. Software tools depend on the environment (the operating system, shared libraries, and their configuration), so a tool running in more than one environment can produce different results. Some examples are: - the same UTC timestamp can be converted to the same local time zone (and vice versa) differently depending on the time zone library and its version; - the same Unicode string (like a file name) can be displayed (reported) differently depending on the Unicode support implemented in the operating system (a list of supported Unicode characters can vary);

		- a tool (and, in general, a technique) relying on file system and volume management drivers can produce different results depending on the versions of these drivers and the version of the operating system; - similarly, a tool using shared libraries to parse data can produce different results depending on the version of various third-party libraries and their data (especially, when on-disk data format updates are not backward-compatible). And more specific examples are: - Tools that process data from volume shadow copies by mounting the disk image in the Windows operating system and then using the devices exposed by the volume snapshot driver can see no volume shadow copies when running under Windows 11 (the volume snapshot driver does not expose volume shadow copies if the underlying storage device is read-only, this problem does not exist previous versions of Windows); - Tools that parse Windows event logs by using API functions provided by the Windows operating system can display the same event (in the human-readable form) differently depending on the operating system version used (because string templates used in different versions of the operating system can vary). This also affects event logs created by third-party applications. This why validation for a given version of a tool should be done more than once in order to cover more environments in which this tool is used. This requirement can be relaxed (within reasonable limits) for tools that do not parse data, for tools that are included into “stable” environments (e.g., live distributions and kiosks), and for specific functions of tools that are known to be “stable” across different environments (e.g., using a bundled third-party library). 2. Some software can be updated without changing its version number. In 2016, PassMark released an updated version of the OSFClone tool (this is a live distribution used for acquisitions), which included a configuration change to disable the swap space activation on attached storage devices. The version number of the tool did not change, though. Another example is software packaged by many Linux distributions. This software is often patched to

		introduce distribution-specific changes (fixing compiler warning, typos, vulnerabilities, and even some other bugs), while the version number remains the same (only the revision number is changed, which is a distribution-specific field, often not reported by a tool when displaying its version). A more specific example is the dcfldd tool. In one version of the Debian operating system, the following version of the tool is shipped: 1.3.4.1-10, but the tool reports its version as 1.3.4-1 (the revision number, 10, is omitted). This revision includes multiple patches, see the Debian changelog: https://metadata.ftp-master.debian.org/changelogs//main/d/dcfldd/dcfldd_1.3.4.1-10_changelog. Along those patches, there is one moving the storage device size probing function to a newer, 64-bit, call ("10_fix-probing-of-large-block-devices.patch"), thus *introducing* one bug that can be encountered when testing the patched revision of the tool (but both revisions, patched and unpatched, report the same version number). With the patch applied, the tool gives the following output: `# dcfldd if=/dev/sda of=/dev/null sizeprobe=if` `[0% of **249511424Mb**] 92160 blocks (2880Mb) written. 24:03:54 remaining.^C` `92204+0 records in` `92204+0 records out` Without that patch, the output is: `# dcfldd if=/dev/sda of=/dev/null sizeprobe=if` `[0% of **487327Mb**] 96000 blocks (3000Mb) written. 00:00:00 remaining.^C` `96082+0 records in` `96082+0 records out` The real size of the block device ("/dev/sda") is, in bytes: `# blockdev --getsize64 /dev/sda` `**510999396352**` Or 487327 MiB. The patched revision of the tool report an invalid size. The reason is that the patched revision takes the device size, in bytes, and multiplies it by the sector size, in bytes

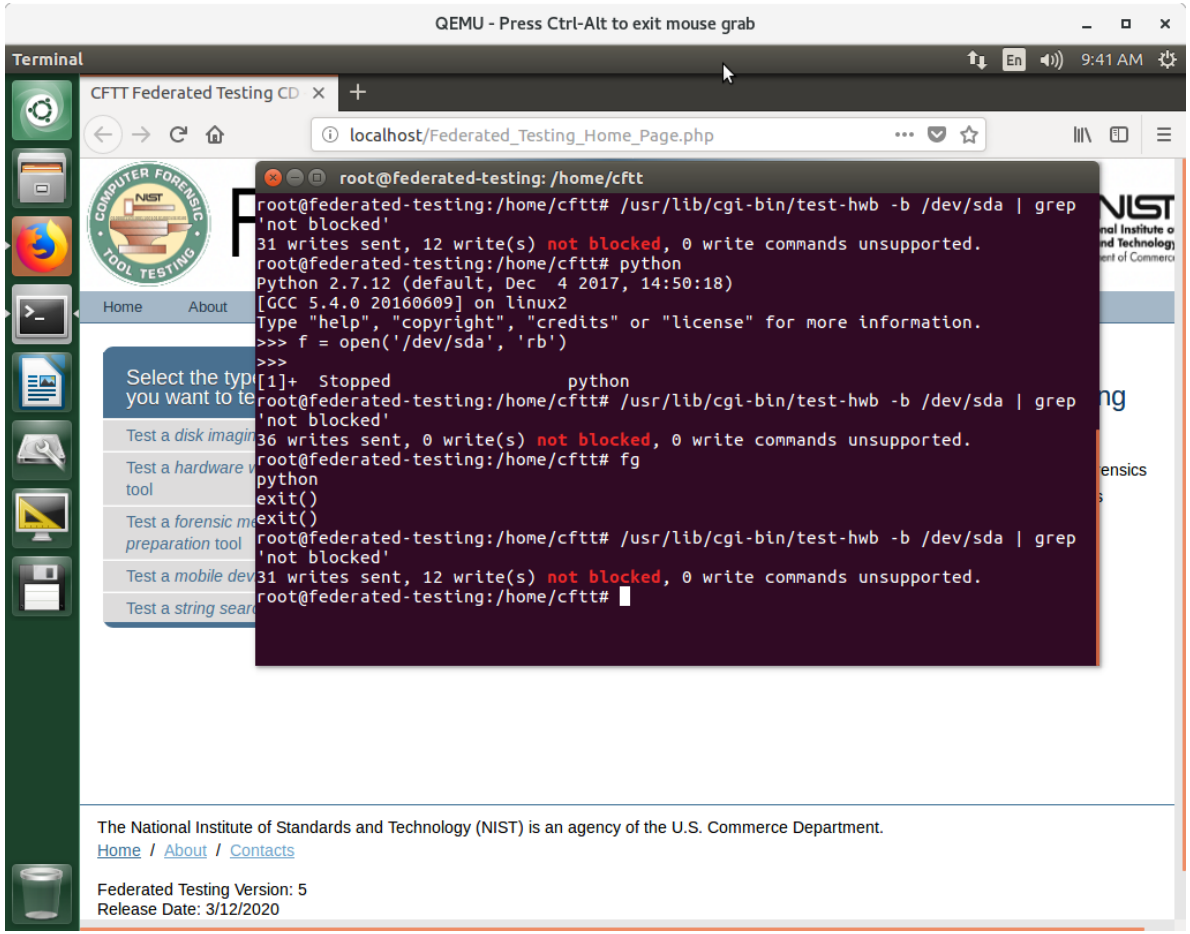
		(249511424÷487327=512). This is why validation must account possible code and configuration changes not affecting the version number.
47	**KEY TAKEAWAY #4.6:** It is not feasible to test all combinations of tools and digital evidence sources.	It is practically impossible to test all combinations of tools and source data. A single file system type has many thousands of possible states (this includes all flags stored in the file system header and in core metadata locations, all possible states of low-level structures, and all types of underlying media), there is no way to test every state against a selected tool, while testing only those file system states that are considered “common” or “real-world” can result in important issues being undetected. A volume formatted using the NTFS file system can be in the following states: a.1. unmounted properly (clean); a.2. not unmounted properly, without an ongoing I/O operation interrupted (when disconnecting the volume); a.3. not unmounted properly, with an ongoing I/O operation interrupted; b.1. not marked as corrupt; b.2. marked as corrupt; c.1. starting at the first block (sector) of a partition; c.2. starting at LBA of 0 (no partition table is used on a storage device, this is often found on USB flash drives); d.1. located on the SSD; d.2. located on the HDD; d.3. located on the USB flash drive; e.1. using the $LogFile version 1.1; e.2. using the $LogFile version 2.0; f.1. using 1024-byte file record segments; f.2. using 4096-byte file record segments. Although the list above is incomplete, it gives dozens of possible combinations to test. It is possible to

		test each state in the corresponding group separately, but this will never produce complete results, because some issues can be uncovered only when a specific combination of states is encountered. Here are two examples: - It is important to run tests using, at least, HDDs and SSDs. BitCurator is a distribution used for archival and digital forensics acquisitions, it can be installed on a physical machine. If an SSD is attached as a secondary drive and its file system is mounted in the read-only mode for data preview, the unallocated space of this file system can be trimmed on Sundays or Mondays. See the corresponding issue: https://github.com/BitCurator/bitcurator-distro-main/issues/102. - It is important to account complex file system states when testing data acquisition tools and hardware write blockers. Tableau TD3 is a forensic imager and a network-based write blocker, it has several ports marked as write-blocked (including SATA and USB). When a storage device is attached to a “write-blocked” port, a write command can be sent to it. All of the following conditions must be met to trigger the issue: a. a source drive has at least one Ext4 volume supported for data preview (e.g., this is an HDD with a partition table and its only partition is formatted using the Ext4 file system); b. an error (e.g., an I/O error) is recorded in the file system journal; c. no error is recorded in the file system superblock. (These conditions can be encountered in the real world, but they are unlikely to be selected during a usual validation test.) Under these conditions, the Ext4 driver of the Tableau TD3 device (its firmware is Linux-based) will transfer the error code from the journal to the superblock, thus issuing a write command through a “write-blocked” port. The problem arises from the fact that the device implements no write blocking and it

		automatically mounts a file system for data preview (later, a user can browse this file system using a web browser; the file system is mounted automatically, even when no network-based functions are requested by a user), so the read-only mode is not enforced. See the corresponding paper: https://github.com/msuhanov/Linux-write-blocker/tree/master/research. And a test image (to be written to a partition of an HDD to be tested): https://github.com/msuhanov/articles/blob/master/misc/ext4.raw. This problem was discovered in 2016, but the most recent firmware (version 2.1.1) is still affected. The problem does not affect Tableau TX1 (a more recent product). This is why validation requires prior reverse engineering to locate weak spots (instead of blindly testing thousands of file system states, which is practically impossible).
50	Tools often omit readable sectors surrounding a bad sector, usually related to how the file system blocks disk sectors for the interface (USB, SATA, Firewire, etc.) used to access the hard drive.	1. Disk imaging tools skip readable sectors surrounding an unreadable one because of the error granularity. This has nothing to do with file systems (although file systems can track unreadable sectors using larger blocks like clusters, this has no direct relation to disk imaging). For example, if a disk imaging tool reads 8 sectors at a time and one of them is unreadable, the whole read request fails (as a result, all of 8 sectors are reported as unreadable). In this situation, a disk imaging tool can read the failed range again, requesting 1 sector at a time. If the tool does not do that, an anomaly is observed (not all readable data is acquired). 2. Hardware write blockers can affect the error granularity. For example, Tableau T356789iu is a forensic bridge that can report 128 sectors as unreadable when only one unreadable sector is encountered (and there is no way to read those 127 sectors from the host side). This problem does not affect the latest firmware (the only firmware version affected is 1.3.0). The problem arises from the fact that the bridge provides a read-ahead cache, which is kept in its random-access memory. All read requests coming from the host are served through this cache. And the cache is populated with a poor error granularity.

50	The most serious failure ever observed was reported in 2003 for SafeBack Version 2.0 (National Institute of Justice and National Institute of Standards and Technology 2003). *[…]*	There was another failure of equal value observed. In the test results for the dcfldd tool (version 1.3.4-1), there is a case where data on a destination storage device becomes misaligned after an unreadable sector is encountered on a source storage device. According to the report (https://www.dhs.gov/sites/default/files/publications/DCFLDD%201%203%204-1%20Test%20Report_updated.pdf): When a drive with faulty sectors was imaged (test case DA-09) the data cloned to the target drive became misaligned after faulty sectors were encountered on the source drive. For example, sector 6,160,448 on the target drive contained the contents of sector 6,160,392 from the source, sector 6,160,449 on the target contained the contents of source sector 6,160,393, and so on. The size of the offset or misalignment between the data on the source and target drives grew as more faulty sectors were encountered on the source. This means that the dcfldd tool produces inaccurate results for drives containing unreadable sectors. Since unreadable sectors can occasionally be found on HDDs and SSDs, as well as on USB flash drives and memory cards, the tool is unreliable and should not be used for acquisitions. Otherwise, misaligned data would be extremely hard to account (e.g., a tool parsing a file system does not expect a data block to be stored at a different sector number). More information about the issue can be found here: https://github.com/adulau/dcfldd/issues/1.
51	Except for one model device, hardware write-block devices always blocked write commands. The firmware for the one blocker that allowed write commands was quickly fixed by the vendor.	1. There was another failure observed. Coolgear SS-127ASD is a SATA/PATA-to-USB adapter with the write blocking switch. A similar piece of hardware is also sold under other brands (e.g., AgeStar). According to three reports, this device does not block some write commands:

50	Most write blockers on the market were able to block commands that would have changed a drive. The few exceptions were for uncommon commands or, in one case, where a vendor was unaware of a change to a chipset (that was quickly fixed).	https://www.dhs.gov/sites/default/files/publications/testresults_for_cool_gear_ss-127asd_sata-ide_adapter.pdf, https://www.dhs.gov/sites/default/files/publications/testresults_for_cool_gear_sata-ide_adapter_windows.pdf, and https://www.dhs.gov/sites/default/files/publications/testresults_for_cool_gear_sata-ide_adapter_linux.pdf. 1.1. On page 50, one more issue with hardware write blockers is mentioned. 2. It should be noted that such tests utilize the “read-write-read-compare” approach (a sector is read to get its original data, then an attempt is made to write to that sector, and the sector is read again to see if anything has changed in its data), or a similar approach. In theory, a hardware write blocker can fail to block a write command, but an attempt to read a successfully overwritten sector would produce old (intact) data for that sector (while data on the storage device has changed), because all read requests are served through a read-ahead cache, which is not invalidated on that write operation. To account such cases, a hardware write blocker should be rebooted before checking the sectors for possible data modifications (to reset its random-access memory and, therefore, the cache). A similar issue affecting the “test-hwb” tool was reported to NIST in 2018 (this issue has a similar nature, but it occurs on the host side). The tool fails to detect modifications on a storage device if a corresponding block device is opened in another program (like a partition table viewer). The issue is still present in the Federated Testing suite, version 5:

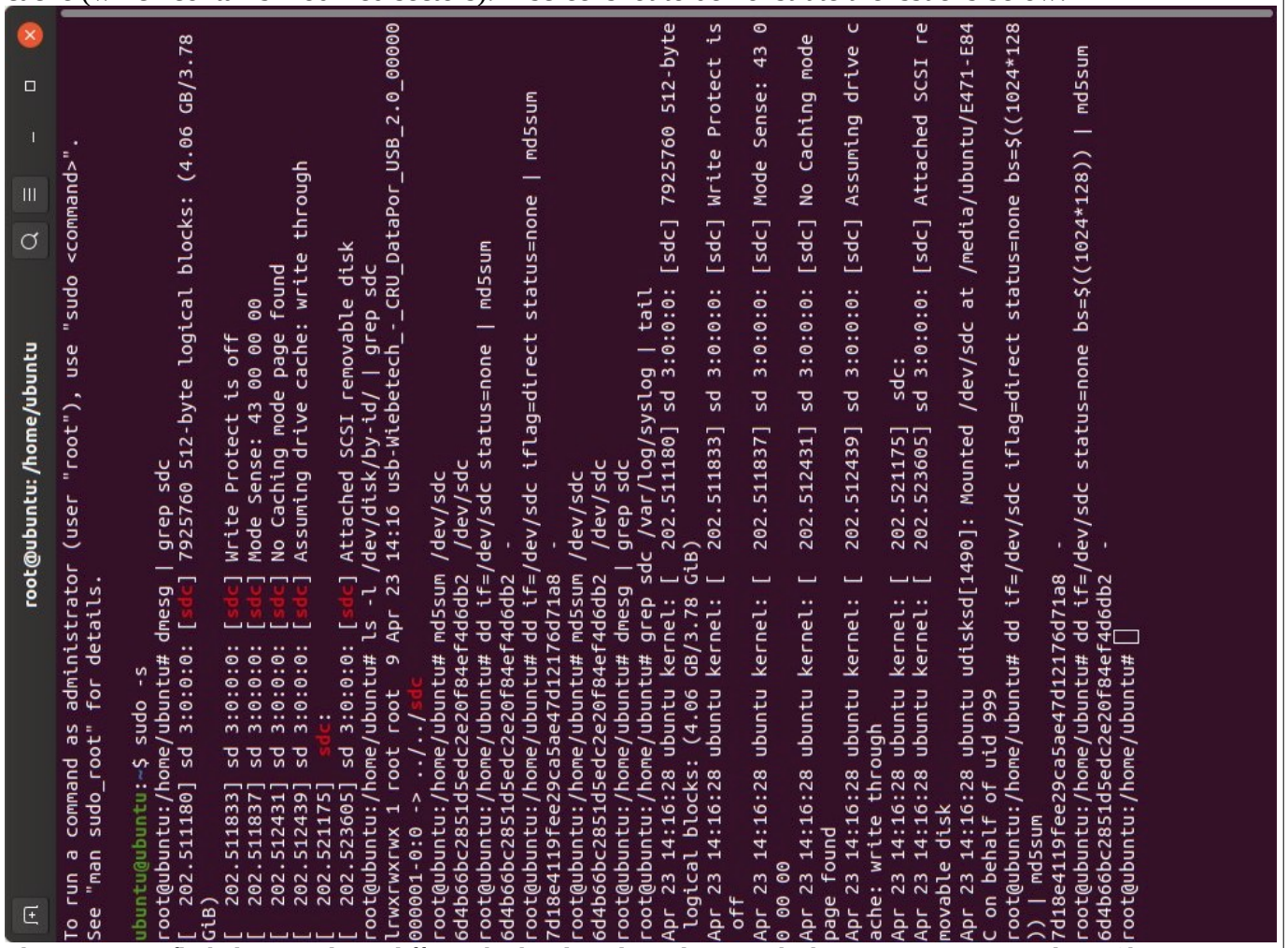


The “/dev/sda” device is not write protected, but during the second execution all writes are reported as blocked.

This greatly affects the reliability of test results for hardware write blockers.

2.1. The “read-write-read-compare” approach does not fully account data modifications performed by a hardware write blocker itself, with no command from the host (like in the Tableau TD3 case mentioned

		before). Also, this approach does not account bugs that allow overwriting a sector with the same data (i.e., when a hash value of a given sector remains the same, but its data is overwritten with a write command). Although no such issues were observed in practice. 3. The usage of hardware write blockers can introduce anomalies into acquired data, while source data (i.e., located on a source storage device) remains intact. CRU WiebeTech USB (2.0) WriteBlocker is a hardware write blocker designed for USB flash drives and other USB storage devices. It reports an attached storage device as writable and silently discards all write commands (no write error is reported to a host). An image created using this hardware write blocker can include sectors with data that differs from located on a storage device if the following conditions are met: - a Linux distribution is used to acquire the storage device; - this storage device contains a file system supported by the Linux distribution (e.g., FAT32); - this file system starts at LBA of 0 (no partition table is used); - this file system is mounted in the “read-write” mode before the acquisition (e.g., the automount feature is enabled); - the disk imaging tool does not use direct I/O requests. Under these conditions, the operating system serves read requests through its cache. Since the drive is reported as writable, and the file system is mounted in the “read-write” mode, and write errors are suppressed, write attempts successfully modify the cache but do not reach the storage device. Only read requests marked as direct bypass the cache and give data from the drive (but many disk imaging tools do not mark their I/O requests as direct: e.g., the dd tool does not use them by default). This means that modifications (like timestamp updates) made by the operating system touch the cache but not the storage device. Attempts to read data from the storage device are likely to produce data from the

		cache (which contains modified sectors). A screenshot to demonstrate the issue is below:  The same USB flash drive produces different hash values depending on whether I/O operations are direct (the "iflag=direct" argument is given) or not.
51	**KEY TAKEAWAY #4.7:** Extensive	It is not clear which anomalies are considered as minor. Since the root cause of many anomalies is often

	tool testing of over 250 widely used digital forensic tools showed that most tools can perform their intended functions with only minor anomalies.	not revealed, it is unclear how to differentiate between minor and major ones. For example, if a disk imaging tool modifies some data on a source storage device, this modification can be considered as minor when only several sectors are affected. However, further tests can demonstrate that the same issue can modify thousands of sectors (depending on the file system state found on the storage device). A more specific example is the ASR Data SMART live CD (version 2011-01). According to the report (https://www.dhs.gov/sites/default/files/publications/508_Test%20Report_NIST_Digital%20Data_ASR_Data_SMART_September%202012.pdf): The execution environment, the SMART Linux live CD version 2011-01, not the tool, modified the source drive in test cases DA-02-F12, DA-02-F32, and DA-06-ATA28. The source drive, 01-IDE, contained an NTFS and several other file systems. In each case 88 sectors belonging to the NTFS file system journal were changed. Since the execution environment's changes were limited to the NTFS partition, the accuracy of the DA-02-F12 and DA-02-F32 acquisitions (acquisitions of the drive's FAT 12 and FAT 32 partitions) were not affected. However, in DA-06-ATA28 this resulted in 88 sectors differing between the image file created by the tool and the original unaltered source. When the test cases were rerun with the source attached via hardware write block (DA-02-F12-WB, DA-02-F32-WB and DA-06-ATA28-WB), the tests completed without anomaly. It should be noted that in testing SMART, other drives that contained NTFS file systems were imaged but were not modified by the SMART Linux environment. This behavior of SMART Linux changing the source was only seen with the NTFS file system on drive 01-IDE. According to another research (https://dfir.ru/2018/07/25/a-live-forensic-distribution-writing-to-a-suspect-drive/), this issue was caused by the live environment mounting NTFS file systems in the "read-write" mode during the boot. This results in the $LogFile journal being wiped if the file system was not unmounted properly before the test. So, the issue affects more than just 88 sectors on an IDE drive. And since the root cause was not analyzed, the same issue existing in other live distributions tested (e.g., SUMURI PALADIN 6.09) was not discovered.

55	A real-world data set also has disadvantages: […]	Privacy concerns and copyright issues can be mentioned as disadvantages of real-world data. Even simulated data can include unwanted personal information. And even simple file system images can include copyrighted material (like boot code embedded into the NTFS file system).

Chapter 5

Page	Text	Comments
56	Currently, digital forensics labs are each testing the same tools causing redundant work.	Again, it is essential to test the same tools in different environments. Even the same version number of a tool does not guarantee that this tool does not produce different output for the same input data. *More information about this topic is provided above (see page 9 of this document).*
56	Better analysis of how digital evidence is used and whether there have been incorrect or misleading conclusions. Having this information centrally collected would benefit the field.	It should be noted that many issues with digital forensic tools are discovered outside of the NIST tool testing programs and projects. Such issues are often reported directly to vendors. In many cases, especially with proprietary software, these reports and related change log entries are never made available to the public (i.e., bugs are fixed quietly). For example, one popular mobile forensic tool always reported the inode changed timestamp as the file created timestamp when parsing the F2FS file system (this raised questions about the origin of a photo in one case). This issue was reported to the vendor, the vendor fixed it (and before confirming the bug, the vendor wrongly insisted that all timestamps are reported correctly). However, no change log entry (and no announcement in any form) that somehow describes the fix is available. (The vendor and the tool are not named here for legal reasons.) A good idea is to create a public issue tracker for bugs discovered in digital forensic tools, so all practitioners could report an issue in any digital forensic tool. The issue tracker can be useful when dealing with vendors who do not accept bug reports for some reason.
56	While developing this report, we encountered many areas that need further research and improved	There is another important area: - Better understanding of how forensic examiners rely on tools.

	processes, including: *[…]*	Forensic examiners rely too much on tools. But can forensic examiners effectively uncover previously unknown errors in tools they use? Especially, when tools produce expected (but incomplete or even incorrect) results. How many tool errors are mistakenly documented as proper behavior?

PC6

From: Matt Bergin
Subject: Comment on NISTIR 8354-DRAFT
Date: 29 May 2022 20:16
To: "ScientificFoundationReviews" <ScientificFoundationReviews@nist.gov>

The NISTIR 8354-DRAFT document covers many of the topics that I would expect but fails to discuss flaws in forensic software which would allow attackers to generate false positives or negatives in reports used by law enforcement. I am an information security researcher who has published examples of how this can be done and have presented my work at multiple conferences around the world. I am not the only researcher to have done this. The content in forensic reports must be indisputable and that is clearly not the case. I believe NIST has a role to play in evaluating forensic tools for such flaws and certifying that none exist prior to their use in court cases.

Further, forensic software manufacturers often rely on encryption to prevent peer review of the methodologies they've implemented. It requires significant experience with breaking encryption before methods used to perform forensic analysis can be truly peer reviewed. It is not acceptable for forensic methodologies to be hidden behind encryption as it creates an unfair burden on someone accused of a crime to disprove forensic report content.

In short, forensic tools should be evaluated for flaws which have the potential to introduce false positives and negatives in reports, and methodologies implemented by forensic software should be open to peer review and not hidden behind encryption.

Respectfully,
Matt Bergin

PC7

Questions/Comments Received During June 1, 2022, Webinar on Digital Investigation Techniques Report

Public Comment Period for NIST Review on Digital Investigation Techniques
Event Date: June 1, 2022 – 01:00 PM to 03:00 PM

Posted Questions* with Time Received
*Names listed were self-created by attendees, thus some may be aliases

[01:31 PM]
Humaira S. asked : If a criminal can change just one bit in a photo to change the hash and evade detection of illegal photos/files, what can be done to remedy this issue?

[01:38 PM]
Matt Bergin asked : The draft document doesn't cover tests for vulnerabilities in forensic software which allow forensic report content to be manipulated. I believe NIST has an important role in preventing this. Should standards include testing to minimize criminals' ability to manipulate forensic reports?

[01:40 PM]
dan mares asked : why not add complete file copy/inventory of evidence tool test in addition to "image". it may not be possible to "image" large server for single folder

[01:44 PM]
Humaira S. asked : What steps have been taken by digital forensics labs to automate their techniques? And what processes are not feasible to automate?

[01:49 PM]
Saba Mylvaganam asked : Very interesting overview of the take-aways from the NIST report. Have you tried manipulating data sets in different operating systems for testing and validating forensic techniques/tools? It should be possible with the availability of resources in the 11 000 test labs in the USA.

[01:47 PM]
Joseph Nicholls asked : Is there a set of definitions as to the differences between a true "Expert" (e.g. engineering degree) verses a technician that has simply been trained in how to use a specific tool? And thus related to how they can testify in the court room?

[01:50 PM]
Eran Salfati asked : Do you discuss the prerequisites that should be in the system in order for it to be "forensicable".

[01:51 PM]
Preston Coleman asked : In relation to the size of the DF community described by the capture/recapture, what is being defined in the use of the term 'lab' or 'laboratory'?

[01:53 PM]
Joseph Nicholls asked : Any discussion on destructive testing techniques for digital forensics, such as "chip-off" for phones, thumb drives and SSD?

[02:00 PM]
Nada Khatib asked : Companies tend to use free forensic tools to capture evidence. Should they do so or use court approved tools?

[02:01 PM]
Mark Broecker asked : what's about cloud forensics and there specific "open source INT", there isn't the possibility to say "we have found/analized everything"

[02:05 PM]
Humaira S. asked : Do you see any significant differences (i.e. tools, processes, standards) between digital forensics labs in academia, corporations, and government labs?

[02:05 PM]
Courtney asked : Can you please again slowly say where destructive testing techniques have been examined/considered? Thanks!

[02:09 PM]
Courtney asked : Are there typical forms, notes, or other types of human-created records that litigators should be aware of when digital forensics are at issue in a case?

[02:10 PM]
Humaira S. asked : Is understanding criminal psychology useful when performing capture and analysis of data?

[02:10 PM]
James Matey asked : Encrypted disks, phones, ... ?

[02:13 PM]
Brianna asked : In regards to the paper and testing of forensics tools, "Currently, digital forensics labs are each testing the same tools causing redundant work. A more structured approach could increase efficiency."

What may be possible structured approaches? I agree with the redundant work amongst labs

[02:17 PM]
Jeff Yan asked : would you consider admissibility in a court of law part of the scientific foundation?

[02:18 PM]
John Loscheider asked : Can data be recovered from a hard drive after a low-level format?

[02:18 PM]
Colin asked : How does Cloud based infrastructure differ in terms of forensics?

[02:20 PM]
Colin asked : How much responsibility would be on the service provider of the cloud service in the event that its a paid service or data storage?

[02:20 PM]
Matt Bergin asked : Many forensic software vendors use encryption to prevent their methodologies from being inspected and peer reviewed. Do forensic software vendors provide source code for NIST to review?

[02:22 PM]
Colin asked : in relation to data forensics

[02:23 PM]
J Nunan asked : How difficult is it to recover Data from a RAID drive (e.g. RAID 5 spread out over 5 hard drives)?

PC8

From: Maxim Suhanov
Date: Thu, June 09, 2022 7:15 AM -0400
To: ScientificFoundationReviews <ScientificFoundationReviews@nist.gov>
Subject: Additional comments on NISTIR 8354-DRAFT

Hello.

Please find attached my additional comments on NISTIR 8354-DRAFT.

The Q&A section of the webinar had some interesting questions about protection of digital forensic tools against reverse engineering and about vulnerabilities in digital forensic tools.

So, here are additional comments.

1. Many proprietary digital forensic tools are distributed in the obfuscated form. Sometimes software protection systems (like Oreans Technologies Themida) are used.

This makes reverse engineering to locate weak spots for further validation hard or nearly impossible (while this does not affect black-box tests, such tests have significant limitations – this was discussed before). Similarly, discovering vulnerabilities (already known and previously unknown) in protected software becomes harder.

2. Many digital forensic tools bundle third-party libraries to provide “stable” functions across different environments. If these libraries are outdated, they can become a source of vulnerabilities.

The number of bundled third-party libraries and tools is high for mobile forensic tools.

3. Some mobile forensic tools deliberately bundle outdated libraries and tools to support old mobile devices.

For example, some Android phones are incompatible with current versions of the ADB tool, so an outdated version of this tool can be packaged (along with a newer version for other Android phones) to run on the host computer.

A vendor can backport patches to fix known vulnerabilities if the source code is available... But who knows? In one mobile forensic tool (which is not named here), a very old (from 2013) “vanilla” (i.e., not patched and not even recompiled with additional protections) version of the ADB tool is shipped along with newer versions.

This practice significantly increases the attack surface.

4. As with any software, users can run outdated versions to work around known (but not yet fixed) bugs. Vendors often suggest software downgrades as a workaround.

5. In general, vulnerabilities in digital forensic tools (and in linked general-purpose libraries and tools) can be divided into two categories:

- Not affecting digital forensic functions (at least directly).

  For example, DLL search order hijacking issues that result in local privilege escalation (e.g., an attacker having local or remote access to a guest account on the computer can gain administrator privileges). Another example is insecure permissions for files, directories, and registry keys that allow unprivileged users to do something that is not normally expected or allowed.

  Usually, such vulnerabilities cannot be exploited without prior access (remote or local) to the examiner's computer (but they can be used in exploit chains covering multiple vulnerabilities). Additionally, some vulnerable executables (e.g., susceptible to DLL

hijacking) can be extracted from digital forensic tools and used to covertly launch malicious code on other computers (a vulnerable executable, which is digitally signed and considered as known good by anti-malware products, can be tricked into running custom code – this allows attackers to bypass security features like application allowlisting).

- Directly affecting digital forensic functions.

  This includes code execution issues affecting the examiner's computer (or digital forensic devices like hardware write blockers and hardware forensic imagers). After successful exploitation, data stored on devices being examined (or acquired) can be deleted or altered (unless these devices are truly write-blocked and this protection cannot be turned off programmatically) or data stored on the examiner's computer (like disk images and reports) can be modified. If the examiner's computer is not air-gapped, a backdoor can be installed to enable remote access.

  For example, many live Linux distributions (including those having forensic capabilities) can execute code stored on an internal drive automatically (without any user input) during the boot (https://dfir.ru/2018/07/21/a-live-forensic-distribution-executing-malicious-code-from-a-suspect-drive/). Another example is a vulnerability found in Cellebrite UFED 4PC, which can be exploited during the acquisition of a mobile device (https://signal.org/blog/cellebrite-vulnerabilities/).

6. Usually, vulnerabilities in digital forensic tools are underrated.

A common point of view can be summarized as follows:

Theoretical scenarios and proof-of-concept exploits are not enough to render digital evidence inadmissible in a particular case. It must be demonstrated that a particular vulnerability has been successfully exploited in this case, leading to specific changes of data (or other consequences).

As with any software, digital forensic tools have bugs. Some of these bugs are vulnerabilities. And some of them are not yet disclosed (also known as zero-day). It is okay when vendors patch vulnerabilities in a timely manner. It is not okay when vendors ignore known vulnerabilities and do not fix the software (unless these vulnerabilities are in the news). The absence of legal consequences in a particular case (like rendering digital evidence inadmissible) is not an excuse to ignore security issues.

While examiners can reduce the risks (https://www.sans.org/blog/arbitrary-code-execution-on-examiner-systems-via-file-format-vulnerabilities/), it is important for vendors to reduce the attack surface, introduce mitigations (measures that make successful exploitation of vulnerabilities harder) and controls (like security sandboxing), and implement the vulnerability management processes.

6.1. Let us assume that an examiner's computer is infected with a malicious program that places illicit material into some devices being acquired.

According to one blog post (http://cyberlaw.stanford.edu/blog/2021/05/i-have-lot-say-about-signal%E2%80%99s-cellebrite-hack), a different tool can be used to demonstrate that no exploits spoiled the digital evidence (by performing another acquisition and comparing its results to the previous one, which was allegedly affected by an exploit).

This is not going to work in many scenarios, especially with mobile devices (because they cannot be truly write-blocked) – when original data is spoiled or faked (modified in some way during the first acquisition), comparing its copies is meaningless, because “spoiled or faked bytes” make their way into both copies (two tools are going to read the same overwritten blocks of data).

There are some other ways to verify the reliability of digital evidence:

- use different types of evidence (e.g., a witness testimony or a confession) to confirm that the data in question existed before the digital forensic tool was used (so, this data was not planted during or after the acquisition);

- if multiple devices have been seized and some of them were examined (or acquired) using a different tool, find references to the data in question, or copies of that data, on those devices (to prove that the data in question exists on other devices, which were not affected by a vulnerable tool);

- similarly, more sources of digital evidence, like warrant returns and cloud data, can be used to find references to the data in question (or copies of that data);

- provide evidence that a user knew about the data in question before the seizure of the device (e.g., a user viewed or edited that data; but such traces could be planted as well);

- provide evidence that the data in question was written to the device memory before the seizure (e.g., relevant files and database entries were not backdated or timestomped);

- examine the machine used to acquire the data and demonstrate that there were no traces of exploits and active malware (this is not going to work if the acquisition tool used is, or is a part of, a live distribution, because all relevant data was stored in the random-access memory, which is lost when the computer is powered off).

If no different evidence is available, no alternative sources of digital evidence (other devices, warrant returns, cloud data, etc.) are available (or they do not contain relevant information), and the examiner’s machine is unavailable (e.g., it was decommissioned or reimaged a long time ago, or it cannot be examined due to legal reasons), the only source left is the device itself. And such situations are not unusual.

In these situations, can we rely on examiners to identify backdated or timestomped files and database entries? Is it enough to prove that evidence was not spoiled or planted? If yes, why do we implement (and, in many situations, require the use of) additional measures like tamper-evident bags when dealing with physical access to evidence?

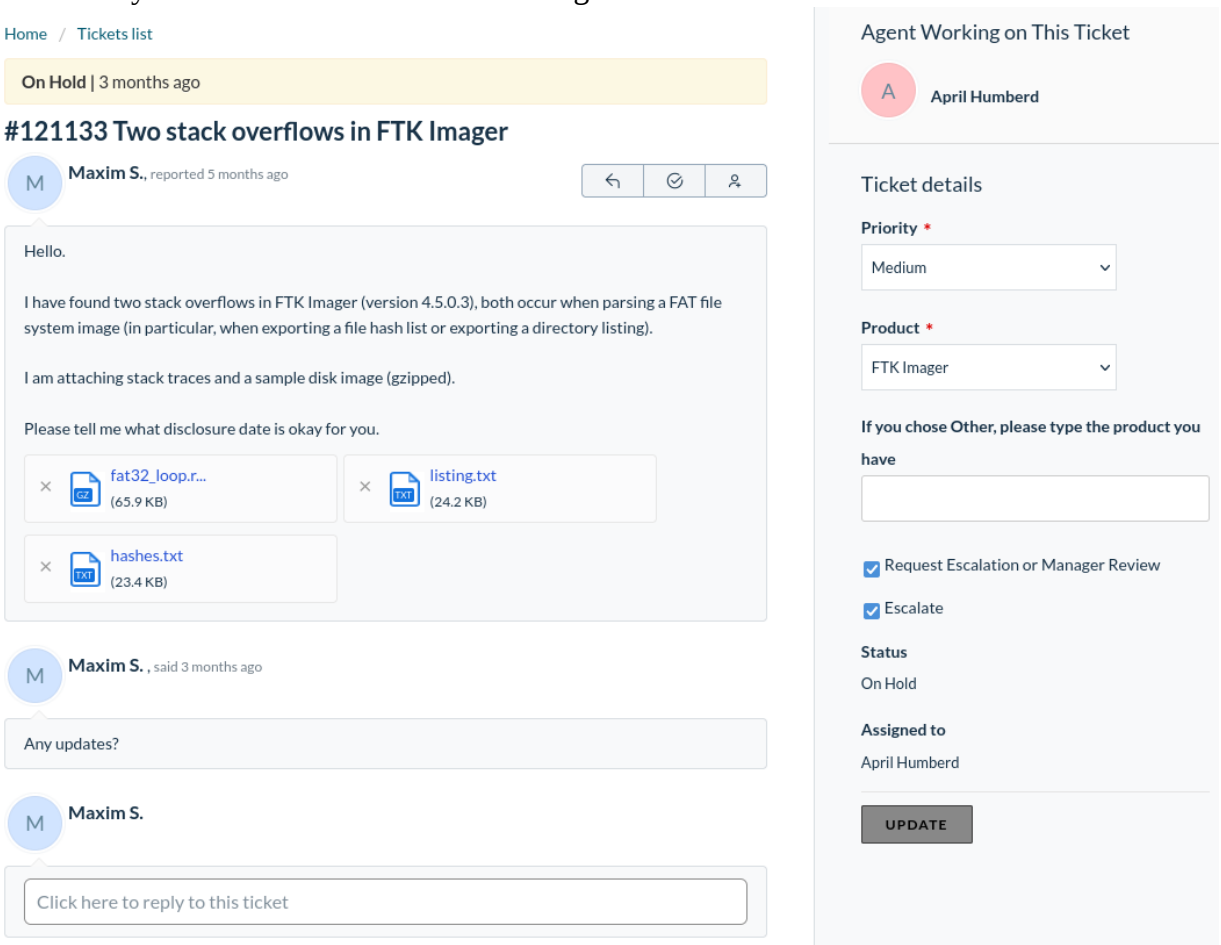
Or do we need additional protections here? If so, why do we rely on tools from vendors who do not care about security? Can we compare the *ability* of gaining unauthorized and undetected physical access to evidence (e.g., when no tamper-evident bags or similar packages are used to store and transfer evidence) with attaching evidence to a computer running a digital forensic tool having known vulnerabilities (or to a computer running a digital forensic tool that could be successfully exploited before) in terms of reliability? (We cannot overcome zero-day vulnerabilities, but we also have known ones, which can be around in digital forensic tools for years.)

7. Some vendors and maintainers of digital forensic tools occasionally ignore vulnerability reports.

Here are two recent examples:

- Multiple security issues with The Sleuth Kit: https://github.com/sleuthkit/sleuthkit/issues/2641.

- Two security issues with the Exterro FTK Imager:



(As of 2022-06-08, these stack overflows are not fixed.)

8. Vulnerabilities in digital forensic tools can affect other software products.

For example, a vendor can decide to reuse vulnerable file system parsing code taken from a digital forensic tool in an enterprise product (e.g., in an EDR tool). These new software products can be deployed in large-scale corporate networks, delivering vulnerabilities to their workstations and servers.

PC9

From: Ryan, Dawn, M
Sent: Thursday, June 23, 2022 12:07 PM
To: ScientificFoundationReviews <ScientificFoundationReviews@nist.gov>
Subject: NISTIR 8354-DRAFT comments

Hello,

Thank you for the opportunity to comment on NISTIR 8354-DRAFT. Overall, I must say this is fantastic work and a much-needed publication.

My comments focus on *Section 2.8.9 Academic Course work.* My first thought is this section could benefit from expansion. My second is specific to line #2 and the term "basic education." I would suggest removing the word basic. In support of this, please consider the following:

1. Colleges and universities offer much more than a basic education in some instances. For example, there are programs designated as Centers for Digital Forensics Academic Excellence through the Defense Cyber Crime Center. This is a thorough articulation process recognizing college programs that align with the rigor of the DC3 Cyber Training Academy. This process includes the evaluation of the curriculum not only at the fundamental level, but also the demonstrated practice level.
2. Many colleges offer very intense hands-on training in addition to the expected academic theory and fundamentals. For example, many offer a wide array of experience in all elements of the digital forensics process model including deep dive analysis with a myriad of tools ranging from raw hex analysis to advanced certified level use of commercial and open-source tools.
3. There are programs like the DC3 and the NSA/DHS Centers for Excellence that recognize the rigor and advanced level of education these schools provide. Links below for your reference.

https://www.dc3.mil/Missions/Cyber-Training/National-Centers-of-Digital-Forensics-Academic-Excellence-CDFAE/

https://www.nsa.gov/Academics/Centers-of-Academic-Excellence/

Again, thank you for accepting input. I am happy to discuss further if needed.

Dawn

Dawn Ryan
Associate Professor | Cybersecurity, Networking and Digital Forensics
Anne Arundel Community College

PC10

From: Graeme Horsman
Sent: Friday, July 1, 2022 1:42 PM
To: ScientificFoundationReviews <ScientificFoundationReviews@nist.gov>
Subject: Comments - Digital Investigation Techniques: A Scientific Foundation Review

Good evening,

First, I really enjoyed reading the work - Digital Investigation Techniques: A Scientific Foundation Review.

As part of your consultation period, I just had the following few comments.

Fig4-1. I would suggest that there is a stage of 'handling' that maybe requires mentioning for completeness. Following on from this, the model sits in 2 sections - collect and interpret. I understand that you are not trying to produce a new model but I would argue that there is an 'examine' stage that is distinct from interpretative processes. For example, processing data is something I wouldn't consider to be an interpretative process, as interpretation of the data occurs after. The 'navigate' stage may also cause a little confusion as Im not sure what you mean here in terms of a stage. If it refers to navigation of the system, then I would consider that an examination process, where findings might then be interpreted.

Finally, I think the process misses out the stage of peer review of the work undertaken.

Your statement - 'The collection steps ensure the integrity of the acquired data to provide a stable source for the analysis of the data.' - I think handling is an important part of this.

You refer to 'computer forensic tools' - should this not be a broader term to encompass all tools, not just computer ones.

'Acquire digital data. This is accomplished by copying data to make an image file of the acquired digital data.' - I think techniques are more frequently now producing an 'extraction' as a product which is not always comparable to what we know as an image.

Page 33. - Points 5 and 6 feel very much the same or require more distinction to be drawn.

To follow on from the point above, your numbers list of important tasks does not include any interpretation and evaluation stages, testing, validation of hypotheses or peer review processes.

'Before any identified data can be copied (acquired) from a storage device, the device must be attached to a computer' - what above any acquisition methods over a network or to a specific form of hardware?

Pg34 - ‘In these situations, digital data is acquired imperfectly in that there are small differences between the actual data present on the device and the acquired data.’ I would argue that using the term ‘small’ might be misleading and subjective. Jailbreaking processes could create big changes.

‘Since most operating systems do not overwrite deleted data,’ - does this statement need ‘by default’ added to the end?

Pg 37 - ‘The tool must recognize and interpret’ - Is the tool interpreting anything? Or is it identifying something based on its rules and then presenting data in accordance with those rules. Im not sure that’s an interpretation.

*Reading through these sections at this point - I wonder if you are missing a stage of ‘developing an investigative strategy’.

4.7 Analysis of Results - should alternative hypotheses be considered and evaluated?

4.7.2 Anti-Forensics - do AF tools need clarifying as not all processes that could cause an issue to a practitioner are AF - some are legitimately privacy preserving. The intention of the user may need to be considered?

Thanks
Graeme

Dr Graeme Horsman
Cranfield University

PC11

From: Constantinos Patsakis
Date: Fri, July 08, 2022 4:34 AM -0400
To: ScientificFoundationReviews <ScientificFoundationReviews@nist.gov>
CC: Tantalus X
Subject: Comments on NIST.IR 8354-DRAFT

Dear all,

Many thanks for the very intuitive and helpful review which we believe benefits the whole community. To further improve the review, we append at the end of this email some comments for your consideration.

Kind regards,

Constantinos Patsakis, University of Piraeus, Greece
Nikolaos Mantas, Independent DFIR Researcher, Greece

Although Chapter 4.2 focuses on residual data on devices, we believe that a sub-chapter on data acquisition from the cloud should also be addressed. Indeed a traditional forensic acquisition is not feasible as, most of the time, the data will reside in database warehouses; however, the web interfaces provide a means to access them. It is infeasible to acquire all the media that hold these data, as they might exist in different spatial regions with different jurisdictions and laws. Hence, the investigator should focus on identifying and extracting relevant artefacts (logs, telemetry, files) rather than the acquisition.

Arguably the data integrity lies on the owner of the database warehouse to examine their identification, and the authentication is the list of users that have access to these data. It can range from a single suspect or a list of them to workgroups with explicit privileges and administrators and even the owner of the cloud services. To combat non-repudiation claims again and verify the integrity of these files, one must acknowledge the owner of the cloud service as the "trust anchor" of the data to be acquired. Acquisition of data from the cloud should be accompanied by the persons, groups, etc., that have access to them and their respective privileges (View, Edit, etc.) and be distinguished from the access rights of the cloud services provider. For example, in Microsoft Azure cloud services, several actions are performed by the underlying cloud mechanisms themselves, leaving a trail of logs that can be valuable for an investigation. In the case of online storage providers like Dropbox, Mega NZ, etc., the cloud service only provides the storage, and it is up to the user to decide and manage the access rights. Arguably from a concrete forensic standpoint, the acquisition of evidence/data from cloud services may not be made in the strict essence of "forensic soundness", yet, the amount of data residing within and the global adaptation of cloud services is paramount to a forensic investigation.

In Section 4.2, it must be acknowledged that acquiring the forensics data may require alteration of the original evidence due to, e.g., inherent security, anti-forensics, anti-tampering mechanisms or even lack of forensics readiness. These aspects are very relevant in the cases of mobile, IoT, and drone forensics. For instance, if the owner of the device does not cooperate to unlock the device, because of unwillingness or lack of power to do so (the owner of the phone is dead and no one knows the PIN), the means to access it imply alteration of the OS, filesystem, and memory as the device must often be rooted and rebooted. (See also 6.6.1 CWA 17865 https://www.cencenelec.eu/media/CEN-CENELEC/CWAs/RI/cwa17865_2022.pdf which is for mobile forensics). Therefore, proper documentation of the process and possibly other means to verify the process, e.g., recording of the process, might be relevant.

In subsection 4.3, the resistance to the second pre-image attack must be also stated to conform to the cryptographically secure hash function definition.

In subsection 4.5, we propose the inclusion of file system metadata (e.g. last access/modification, access data from MFT) and relevant features, e.g. shadow copies, alternate data streams) since they may provide a wealth of additional information to the investigator.

Minor issues:

- On page 11, add Windows 11
- Use the term "file system consistently", there are two uses of "filesystem"
- Check spaces before citations.

--

Constantinos Patsakis

Department of Informatics
University of Piraeus

PC12

From: Erica Wissolik
Date: Mon, July 11, 2022 9:56 AM -0600
To: ScientificFoundationReviews <ScientificFoundationReviews@nist.gov>
CC: Marc Canellas, Jeanna Matthews, Carlos Ignacio Gutierrez , Russell Harrison, Bruce, Jyotika Athavale, Karen McCabe, Kayla Henneberry, Karen Mulberry
Subject: IEEE's comments on "Digital Investigation Techniques: A Scientific Foundation Review"

Please find attached, IEEE's - the combined effort of IEEE-USA, IEEE Standards Association, and IEEE Computer Society - comments on NIST's draft, Digital Investigation Techniques: A Scientific Foundation Review.

Thank you,

Erica Wissolik
Sr. Program Manager, Government Relations
IEEE-USA

11 July 2022

National Institute of Standards and Technology (NIST)
U.S. Department of Commerce
100 Bureau Drive
Gaithersburg, MD 20899
Via Email: scientificfoundationreviews@nist.gov

Re: *RFC Response: Digital Investigation Techniques: A NIST Scientific Foundation Review (NISTIR 8354-DRAFT)*

IEEE is pleased to submit these comments on the above-captioned, Request for Comment on NIST's *Digital Investigative Techniques: A NIST Scientific Foundation Review* (8354-DRAFT, "the Review"). These comments are the combined effort of IEEE-USA, the IEEE Standards Association (IEEE SA), and the IEEE Computer Society.

IEEE-USA is the American component of the IEEE, representing approximately 150,000 engineers, scientists, and allied professionals in United States, many of whom are actively conducting research and development into digital forensics, privacy, cybersecurity, artificial intelligence, software engineering, and advanced computing, as well as other foundational and emerging technologies.[1]

The IEEE SA is a globally recognized standards-setting body within IEEE.[2] The IEEE standards development process is rooted in consensus, due process, openness, right to appeal and balance.[3] It adheres to and supports the principles and requirements of the World Trade Organization's (WTO) Decision on Principles for the Development of International Standards, Guides and Recommendations. In particular, the IEEE operates in active agreement with the WTO principle that standards should not create unnecessary obstacles to trade, and whenever appropriate, should specify requirements in terms of performance rather than design or descriptive characteristics.[4]

The IEEE Computer Society (Computer Society) is the premier source for information, inspiration, and collaboration in computer science and engineering.[5] The Computer Society has over 373,000 members from 168 countries; hosts 215 technical conferences annually; maintains 230 active global technical standards; and, publishes 47 peer-reviewed journals and magazines including those that the Review[6] lists as leading publications on digital forensics and computer security (e.g., IEEE Transactions on Cloud Computing, IEEE Transactions on Information Forensics and Security, and IEEE Transactions on Software Security).

Our society sits at the intersection between enabling new levels of technological efficiency and enabling technology to become a force for good that goes beyond efficiency. We have a critical opportunity to use technology to make society more equitable, inclusive, and just; make government operations more transparent and

[1] https://ieeeusa.org/
[2] https://standards.ieee.org/
[3] "Developing standards: Who Oversees The Process?" IEEE Standards Association. https://standards.ieee.org/develop/develop-standards/govern/ (defining Due Process as "having highly visible procedures for standards creation and following them;" Openness as "ensuring that all interested parties can participate and are not restricted to a particular type or category," Consensus as requiring a supermajority of a group to approve a draft standard (75% of the ballots must be returned, with 75% of them voting yes); Balance as "ensuring that voting groups include all interested parties.
[4] IEEE, "IEEE Adherence to the World Trade Organization Principles for International Standardization," August 19 2020. Available at: https://globalpolicy.ieee.org/wp-content/uploads/2020/08/IEEE20013.pdf
[5] https://www.computer.org/about
[6] Review, § 2.8.8

IEEE-USA | 2001 L Street, N.W., Suite 700, Washington, D.C. 20036-4928 USA
Office: +1 202 785 0017 | Fax: +1 202 785 0835 | E-mail: ieeeusa@ieee.org | Web: http://www.ieeeusa.org

accountable; and encourage public participation and increase the public's trust in government. When used according to these objectives, technology can help reaffirm and protect our democratic values.

If we miss the opportunity to use these technologies to ensure protection of human values and trustworthiness, we risk reinforcing disparities in access to goods and services, discouraging public participation in civic life, and eroding the public's trust in government. Put another way: responsible development and use of technology to further safeguard human values and ensure trustworthiness is an approach that leads to a sustainable ecosystem of innovation. It is this type of approach that our society will trust and accept.

To achieve the stated objectives we draw upon our collective expertise and provide the following recommendations to address the needs specified in the draft of the scientific foundation review.[7] To summarize,

1. Digital forensics and their results should be deemed reliable based on objective information gathered through independent verification and validation as informed by standards.
2. Digital forensics and their results should be deemed reliable based on objective information gathered through independent verification and validation standards.
3. Digital forensics should be tested against and governed by standards adhering to principles of due process, openness, consensus, balance, and right of appeal.
4. To ensure digital forensics' effectiveness, competence, awareness, accountability, and transparency in operation, there must be standards and certifications for digital forensics and their operators, and recurring benchmarking exercises and independent studies.
5. Determining the reliability and trustworthiness of forensic technologies like digital forensics requires evaluating them in their operational environments, their use in legal proceedings and how fit the technology is for those uses.
6. Stakeholders of digital forensics must have appropriate access to the software necessary to assess the degree of reliability and validity of information, and whether that information is fit-for-purpose.
7. Trustworthiness is determined by more than reliability, and therefore, requirements to determine trustworthiness must include assessments of the processes and procedures where these systems are deployed.
8. To ensure digital forensics are reliable and trustworthy, governments should provide sufficient funding for testing, evaluation, certification, and investigation of digital forensics.
9. NIST's own Text Retrieval Conference (TREC) program should be used as an example to inform NIST policy in otherareas such as digital forensics.

We expand on each of these 9 recommendations in detail below. In particular, we make specific and substantial suggestions for changes to KEY TAKEAWAY #2.5, #4.2, #4.4, #4.5, #4.6 and #4.7 in the document.

We thank NIST for considering these comments in the agency's revisions to the Request for Comment on NIST's Digital Investigation Techniques: A NIST Scientific Foundation Review (NISTIR 8354-DRAFT). We would welcome any further discussions with the agency on these matters. If you have questions, please do not hesitate to contact Erica Wissolik at (202) 530-8347 or e.wissolik@ieee.org.

[7] For a similar statement, see IEEE-USA "Letter to the National Institute of Standards and Technology (NIST), responding to request for comments on NIST Internal Report 8351-DRAFT *DNA Mixture Interpretation: A NIST Scientific Foundation Review*," November 18, 2021. Available: https://www.nist.gov/document/nist-ai-rfi-ieee001pdf [Accessed July 7, 2022].

RECOMMENDATIONS

1. **Digital forensics and their results should be deemed reliable based on objective information gathered through independent verification and validation as informed by IEEE 1012(™), Standard for System, Software, and Hardware Verification and Validation. Without that objective information, neither these systems, nor their results can be considered reliable or trustworthy.**

The use of digital forensics in criminal court can result in catastrophic failures through false imprisonment and the deprivation of people's rights. Scientists and engineers have long demanded that safety-critical software and hardware be the right systems built the right way. Therefore, digital forensics should be independently verified and validated (IV&V) prior to deployment, or prior to informing decisions in the legal system, law enforcement, governance, and related compliance. Specifically, digital forensics ought to be independently verified and validated in accordance with technical standards such as IEEE 1012 Standard for System, Software, and Hardware Verification and Validation,[8] and be subject to recurring post-deployment audit, including with respect to their operators. We encourage NIST to uphold these same requirements.

IEEE 1012 provides a universally applicable and broadly accepted process for helping to ensure that a product is correctly built for its intended use. For example, it is used to verify and validate Department of Defense nuclear weapons systems and NASA manned space systems and critical space exploration probes, among many others.[9]

IV&V are interrelated and complementary processes that build quality into any system. Verification is focused on a product, providing objective evidence for whether the product conforms to requirements, standards, and practices. Validation is focused on customers and stakeholders, providing evidence for whether a product is accurate and effective, solves the right problem, and satisfies the intended use and user needs in the operational environment. In short, verification ensures that a product is correctly built, while validation ensures that the right product is built.

In the context of digital forensics, IV&V answers the following types of questions: Is the analysis used by the digital forensics software the best available, coded as designed, and appropriate for the problem? Does digital forensics systematically favor including or excluding certain types of information? How likely are false negatives and false positives? Would outside experts agree with the software's results at each stage of analysis?

To help appropriately perform IV&V, IEEE 1012 requires that each software and hardware component be assigned an integrity level that increases depending on the likelihood and consequences of a failure: negligible, marginal, critical (causing "major and permanent injury, partial loss of mission, major system damage, or major financial or social loss," referred to as integrity level 3), and catastrophic (causing "loss of human life, complete mission failure, loss of system security and safety, or extensive financial or social loss," referred to as integrity level 4).[10] As the integrity level increases, so too does the intensity and rigor of the required IV&V tasks.

Digital forensics analysis tools should undergo IV&V according to its integrity level as defined by IEEE 1012. Because a thorough and public conversation is yet to take place, there is presently no consensus on such an integrity level. However, the likelihood of digital forensics and other forensic techniques to cause wrongful convictions in the criminal legal system clearly constitutes catastrophic failure, and therefore should be held to the highest integrity level, the level where IV&V should be performed independently.

---

[8] IEEE Standard for System, Software, and Hardware Verification and Validation, IEEE Standard 1012-2016, Sept. 2017 (hereinafter referred to as IEEE 1012) (available at https://standards.ieee.org/ieee/1012/5609/).

[9] For example, NASA's Independent Verification and Validation Technical Framework, https://www.nasa.gov/sites/default/files/atoms/files/ivv_09-1_-_ver_p.doc. [Accessed 7 July 2022.

[10] IEEE 1012, p. 196.

The IV&V process must be independent to avoid conflicts of interest that could lead to catastrophic failure. To this end, IEEE 1012 requires technical, managerial, and financial IV&V when testing software and hardware where catastrophic consequences could occasionally occur and where critical consequences will probably occur.[11] Moreover, letting developers certify their own software is a clear conflict of interest, and the IEEE/Association for Computing Machinery Code of Ethics for Software Engineers is clear about the obligation of developers to manage competing aims.[12] Full definitions of technical, managerial, and financial independence from IEEE 1012 are below, but, in brief, the following must all be independent from the group that oversees the design and building of software: personnel, problem formulation, test and analysis tools for IV&V (technical), responsibility for IV&V (managerial), and control of the budget for IV&V (financial).[13]

Specifically, technical independence "[r]equires the IV&V effort to use personnel who are not involved in the development of the system or its elements. The IV&V effort should formulate its own understanding of the problem and how the proposed system is solving the problem."[14] "Technical independence means that the IV&V effort uses or develops its own set of test and analysis tools separate from the developer's tools."[15] And if sharing tools is necessary, "IV&V conducts qualification tests on tools to assure that the common tools do not contain errors that may mask errors in the system being analyzed and tested."[16] This independence requires the exclusion of parties with a stake in the outcome, which for forensic technologies includes forensic labs and law enforcement agencies who, while not financially dependent on developers, have a shared interest in software's acceptance.

Managerial independence "[r]equires that the responsibility for the IV&V effort be vested in an organization separate from the development and program management organizations. Managerial independence also means that the IV&V effort independently selects the segments of the software, hardware, and system to analyze and test, chooses the IV&V techniques, defines the schedule of IV&V activities, and selects the specific technical issues and problems to act on."[17] The IV&V effort must be "allowed to submit to program management the IV&V results, anomalies, and findings without any restrictions (e.g., without requiring prior approval from the development group) or adverse pressures, direct or indirect, from the development group."[18]

Financial independence "[r]equires that control of the IV&V budget be vested in an organization independent of the development organization. This independence prevents situations where the IV&V effort cannot complete its analysis or test or deliver timely results because funds have been diverted or adverse financial pressures or influences have been exerted."[19]

It is clear from these definitions that peer-reviewed publications, while a priceless tool for scientific inquiry, are not a substitute, nor a valid approximation of IV&V when determining reliability or trustworthiness of a deployed system. Peer-reviewed publications form the foundation of scientific advancement, but peer reviewers of scientific publications are not tasked with answering questions like "Should the digital forensics software or results be admissible in court? Is the digital forensics software fit for the evidence in this legal case?" Peer reviewers do not have access to the system itself and are not tasked with assessing its reliability. Peer reviewers are assessing whether a publication deserves the attention of the scientific community, whether the results described deserve the

[11] IV&V with "rigorous" (the highest level) technical, management, and financial independence "is generally required for integrity level 4 (i.e., loss of life, loss of mission, significant social loss, or financial loss) through regulations and standards imposed on the system development" - where "IV&V responsibility is vested in an organization separate from the development organization." IEEE 1012, p. 199.
[12] D. Gotterbarn, K. Miller, and S. Rogerson, "Computer society and ACM approve software engineering code of ethics," Computer, vol. 32, no. 10, pp. 84–88, 1999. doi: 10.1109/MC.1999.796142.
[13] IEEE 1012, p. 198.
[14] IEEE 1012, p. 198.
[15] IEEE 1012, p. 198.
[16] IEEE 1012, p. 198.
[17] IEEE 1012, p. 198.
[18] IEEE 1012, p. 198.
[19] IEEE 1012, p. 198.

attention of other scientists. With respect to specific legal cases, any individual case could go well beyond the bounds of the published studies.

We acknowledge the fact that not every digital forensic technique must undergo peer review, formal testing, or error rate analysis. However, it is unacceptable for any system influencing decisions with potentially catastrophic consequences - especially forensic techniques used in the criminal legal system[20] - to not be managerially, technically, and financially independently verified and validated.

Therefore, we recommend that NIST:

- Expand Sec. 4.8's discussion of validation vs. verification to cover other important distinctions among types of tests, e.g., local vs general testing, quantitative vs qualitative analysis of results.
- State that digital forensics software should undergo IV&V according to its integrity level as defined by IEEE 1012. If NIST does not recommend that digital forensics should be independently verified and validated in accordance with IEEE 1012, it should articulate why digital forensics are an exception to this international practice among professionals.
- Strike KEY TAKEAWAY #2.5 as informal review is an insufficient method for determining the reliability and trustworthiness of high-risk systems like digital forensics. The appropriate method is to use IV&V.

2. Digital forensics should be tested against and governed by standards adhering to principles of due process, openness, consensus, balance, and right of appeal.

We are concerned that this Review understates the limitations of digital forensics. For example, KEY TAKEAWAY #4.7 states that "[e]xtensive tool testing of over 250 widely used digital forensic tools showed that most tools can perform their intended functions with only minor anomalies." The Review does not provide sufficient analysis to justify the statement that "most tools" will only produce "minor anomalies." Especially as the Review explains that "tool testing" is explicitly not validation or verification (p. 49).

Specifically, as it relates to the typical steps in digital forensics investigations in Chapter 4, discussions of the limitations must be expanded. Sec. 4.6 "Identification and Extraction of Artifacts" and accompanying KEY TAKEAWAY #4.2 should be expanded to provide better coverage of limitations and risks of targeted collection. All techniques (e.g., keyword/string search) have inherent limitations (e.g., deriving, e.g., from the limited knowledge of the operator) which are further complicated by data issues (e.g., incomplete or inaccurate metadata values). The limitations can have very important consequences; however, as it is often the case that, once a collection is made, analysts do not return to the uncollected data. Given the limitations and risks, it is important that local validation of the collection techniques used be conducted. This section would benefit from an amplified discussion of these issues.

Section 4.7 Analysis of Results should also be expanded. Once data has been collected, it is crucial to extract the relevant information from the collected data. There are a number of different types of analysis that may be done to extract relevant information, from straightforward information retrieval to network analysis, and so on. Each of these techniques has limitations and so their effectiveness should be locally validated.

We recommend NIST:

[20] See, for example, IEEE-USA "Letter to the National Institute of Standards and Technology (NIST), responding to request for comments on NIST Internal Report 8351-DRAFT *DNA Mixture Interpretation: A NIST Scientific Foundation Review*," November 18, 2021. Available: https://www.nist.gov/document/nist-ai-rfi-ieee001pdf [Accessed July 7, 2022].

- Expand Sections 4.6, 4.7, and 4.9 and KEY TAKEAWAY #4.2 to better address the limitations of digital forensics.
- Strike KEY TAKEAWAY #4.7 and replace it with a TAKEAWAY that better synthesizes the limitations of digital forensics.
- Add to Section 4.9 that the key requirement for at least some types of testing is the ability to draw random samples from specific subsets of a data population, and that not all search tools have the ability to do so.

3. **Understanding the reliability and trustworthiness of digital forensics requires an accounting of all types and sources of error (including random, systematic, human factors, and organizational) and a meaningful characterization of what can go wrong, how likely is that, what are the consequences, and what can be done about it.**

We are concerned with some of the Review's discussion of errors and error rates in Sec. 4.10.

First, there is no evidence that digital processes "tend to have systematic rather than random errors" as conclusory stated in KEY TAKEAWAY #4.4. A digital process may have random errors or systematic errors depending on the type of digital process or digital technique. The article by Lyle referenced in justification of KEY TAKEAWAY #4.4 itself explains that while for some forensic tools like string search, an error rate may be of limited value, for other tools like file recovery and carving, statistical error rates based on random errors can be meaningful.[21] There should be no blanket statement that any specific type of error is automatically more applicable to digital processes or that error rates are only useful for random errors. Errors of any kind, from random to systematic, organizational to human factors, should be considered in-scope unless its limited value is clearly justified by the context and context and type of process being evaluated.[22]

In the context of error rates, the Review should not be so focused on error rates, and instead focus on the key questions of what can go wrong, how likely is that, what are the consequences, and what can and should be done about it.[23] These are the questions central to the fields of reliability engineering and system safety and the theories of probabilistic risk analysis and defense-in-depth. Lyle understood this broader perspective that just because there is not a formal statistical error rate, does not mean there is no meaningful error to be discussed. As he explained, "[t]ools and techniques without a meaningful statistical error rate should have the types of failures and triggering conditions characterized."[24]

Second, it is generally untrue that "[e]rrors in computer science techniques tend to be so small as to be negligible" as alleged in KEY TAKEAWAY #4.5 Such a statement could be interpreted as a license to do nothing and is not acceptable. This KEY TAKEAWAY seems to imply that industry standards in computer science - such as IEEE 1012 - should not apply to digital forensics technology. We strongly disagree with any conclusion of that kind.

We recommend that NIST:

- Expand Sec. 4.10 to use the appropriate reliability engineering and system safety approach to errors which discusses what can go wrong, the odds of something going wrong, the consequences, and what can and should be done about it.

---

[21] J. R. Lyle, "If error rate is such a simple concept, why don't I have one for my forensic tool yet?," *Digital Investigation*, vol. 7, pp. S135–S139, Aug. 2010, doi: 10.1016/j.diin.2010.05.017.

[22] J. H. Saleh, K. B. Marais, E. Bakolas, and R. V. Cowlagi, "Highlights from the literature on accident causation and system safety: Review of major ideas, recent contributions, and challenges," *Reliability Engineering & System Safety*, vol. 95, no. 11, Nov. 2010, doi: 10.1016/j.ress.2010.07.004

[23] Saleh *supra* n. 20.

[24] Lyle *supra* n. 19.

- Rewrite Key Takeaway #4.4 to state "*For some* digital processes, *error rates may be of limited value*. In those cases, an error mitigation analysis provides more information and is the correct way to manage uncertainty. *Tools and techniques without a meaningful statistical error rate should have the types of failures and triggering conditions characterized."* If NIST does not adopt our recommendation, it should include support for this key takeaway, especially as it pertains to the definition of "some digital processes" and how it came to the conclusion that "error rates may be of limited value." These statements appear to be offered without supporting evidence.
- Strike the language in KEY TAKEAWAY #4.5 that "Errors in computer science techniques tend to be so small as to be negligible." If NIST insists on keeping this in, it should explain its supporting evidence.

4. **To ensure digital forensics' effectiveness, competence, awareness, accountability, and transparency in operation, there must be standards and certifications for digital forensics and their operators, and recurring benchmarking exercises and independent studies.**

IEEE believes that trustworthy systems must adhere to principles including effectiveness (system creators and operators shall provide evidence of the effectiveness and fitness for purpose), transparency (that the basis of a particular decision should always be discoverable), accountability (systems shall be created and operated to provide an unambiguous rationale for all decisions made), awareness of misuse (system creators shall guard against all potential misuses and risks in operation), and competence (system creators shall specify and operators shall adhere to the knowledge and skill required for safe and effective operation).[25]

Fully identifying and characterizing the limitations, failure modes and sequences, and error rates is critical to the trustworthy and reliable use of digital forensics. Therefore, rather than concluding in KEY TAKEAWAY #4.6 that it is infeasible to test all combinations of tools and digital evidence sources, IEEE believes the finding should communicate a reasonable standard for testing. While it is true that exhaustive testing is often not possible, we would encourage establishing benchmarks and would like to see NIST encourage and promote standards for reasonable testing and best practices for clear documentation of what testing has been done and what the results were.

We recommend that NIST strike KEY TAKEAWAY #4.6.

Critically, this testing must include the operators and individual forensic laboratories and law enforcement agencies. The Review makes clear each individual operator's significance. We agree with the Review that "[e]ach lab should ensure that personnel understand the basic capabilities and limitations of a tool, especially the relationship between the tool and the fast-changing IT environment." (Sec. 4.8) We also agree with KEY TAKEAWAY #2.4 that **"**[t]he forensic examiner needs to be aware of key changes in computing technology relevant to the examination being performed. Frequent changes in digital technology introduces the possibility for incomplete analysis or for misunderstanding of the meaning of artifacts."

We believe, and NIST should recommend, that governments should make the reports documenting the required IV&V and audits of forensic techniques public. Furthermore, we believe that governments, including NIST, should encourage, develop, and update standards and certifications for digital forensics and their operators, and fund recurring benchmarking exercises and independent studies to ensure their effectiveness, competence, inclusiveness, accountability, and transparency in operation. Specifically, we believe these standards, certifications, exercises, and studies should address:

[25] IEEE Ethically Aligned Design at 18.

- The requirements for informed trust by the general public in digital forensics (see Recommendation #6 below) and the development of metrics that are immediately and easily accessible by experts and non-experts alike;
- The existence or absence of reliable and unbiased underlying scientific principles and methods in digital forensics;
- The requirements for recurring testing and auditing of the operation of digital forensics, including the operators, field conditions, testing data, environments, methodologies, and performance metrics;
- The requirements for publicly available documentation by developers and testers of digital forensics, and of the use of digital forensics in individual and aggregate cases and decisions;
- The requirements for certification or loss of certification of operators and digital forensics, and for their validation for digital forensics already in use;
- The requirements for individuals to be able to access, review, contest, and correct the data about them, to review and contest the decisions that affect them, and to request human review of such data and decisions;
- The requirements for operation in an ethical manner; and,
- The requirements for identifying and addressing vulnerabilities and threats to security, safety, and privacy such as spoofing, evasion attacks, transfer learning attacks, and data poisoning.

5. **Determining the reliability and trustworthiness of forensic technologies like digital forensics requires evaluating them in their operational environments, their use in legal proceedings and how fit the technology was for those uses.**

IV&V is predicated on the value of testing technology in operational environments. No software or hardware is "generally" reliable -- any technology is only fit for certain purposes. Even technologies that are widely considered to be reliable have known failure modes. For example, cellular telephones are widely considered to be reliable but are not classified as "generally" reliable because they do not work effectively in tunnels or underground. Further, the desire for a technology to be classified as "generally " reliable rather than to consider its reliability in a particular case is misguided. A core premise of labeling a product or process as "well-engineered" is that these operating conditions are specifically defined, tested against pre-defined standards, and accompanied with estimated rates of failure. Systems like digital forensics are engineered products incorporating scientific models and therefore require not only the perspective of researchers who have published proofs-of-concept but also engineers who have used product trials and operational testing and evaluation to demonstrate system performance in operating conditions, against predefined standards, and estimated rates of failure.

Therefore, a scientific foundation review of the reliability and trustworthiness of forensic technologies cannot be effective if detached from an analysis of how the technology is used in legal proceedings, in the forensic technology's operational environment -- yet that is exactly what this Review is purporting to do. The Review examines the peer-reviewed and laboratory studies but does not compare that to any of the criminal cases where digital forensics has been used. Notwithstanding the concerns over peer-reviewed studies discussed above, if the types of files, types of encoding, etc. analyzed in legal proceedings are not similar to the samples used in the peer-reviewed or laboratory studies, the studies have little value.

To determine the reliability of digital forensics, we recommend that NIST catalog and evaluate how digital forensics are being used in legal proceedings and how fit the technology is for those uses.

6. **Stakeholders of digital forensics include far more than forensic scientists, attorneys, judges, and juries. They include the public upon whom these systems are used, litigants, academics, journalists, and other researchers. For those users to assess the degree of reliability, validity, and whether that information is fit-for-purpose, they need appropriate access to the software.**

Users are too often inappropriately denied access or forced to overcome improper and unnecessary barriers to access digital forensics in order to determine the degree of reliability, validity, and whether that information is fit-for purpose. There are many more users of digital forensics than merely forensic scientists, judges, or juries. Independent testing of proprietary or government digital forensics by litigants, academics, journalists, and other researchers is needed to ensure that digital forensics are properly vetted and held accountable. NIST should recommend governments clarify whether and how proprietary digital forensics may be reverse engineered, modified, and evaluated under laws such as the Computer Fraud and Abuse Act and the anti-circumvention provision of the Digital Millennium Copyright Act, and rules of procedure and evidence. More broadly, NIST should recommend governments take steps to affirmatively promote awareness, access, research, and testing including:

- Ensuring accountability and transparency in government procurement and contracting for digital forensics;
- Identifying and disclosing the digital forensics used by the government;
- Adopting clear procedures relating to collection, usage, storage and sharing of personal information collected and used by digital forensics;
- Providing constituents notice about digital forensics decisions, explanations for those decisions, and processes for challenging decisions or data; and,
- Specifically, in legal disputes, tribunals should permit disclosure under appropriate protective orders of intellectual property related to digital forensics when necessary to obtain evidence in compliance with other judicial requirements, including constitutional requirements, discovery laws, or subpoenas.

7. Trustworthiness is determined by more than reliability, and therefore, to determine trustworthiness, one must assess the processes and procedures where these systems are deployed.

Technical assessments of reliability as surveyed in the Review are not the sole determination of trustworthiness. There are eight principles for creating and operating systems that further human values and ensure trustworthiness:[26] (i) human rights: systems shall be created and operated to respect, promote, and protect internationally recognized human rights; (ii) well-being: system creators shall adopt increased human well-being as a primary success criterion for development; (iii) data agency: system creators shall empower individuals with the ability to access and securely share their data, to maintain people's capacity to have control over their identity; (iv) effectiveness: system creators and operators shall provide evidence of the effectiveness and fitness for the purpose of systems; (v) transparency: the basis of a particular system decision should always be discoverable; (vi) accountability: systems shall be created and operated to provide an unambiguous rationale for all decisions made; (vii) awareness of misuse: system creators shall guard against all potential misuses and risks of systems in operation; and, (viii) competence: system creators shall specify and operators shall adhere to the knowledge and skill required for safe and effective operation.

Therefore, we recommend that NIST evaluate more than the technical assessments of reliability to determine trustworthiness. Below we list additional requirements for ensuring the trustworthiness of systems in general which includes the automated decision systems such as digital forensics and many of the forensic technologies used today. If those providing or using digital forensics or any other forensic technologies do not adhere to these requirements, then they should not be deemed trustworthy or fit for their use in determining or affecting people's rights and liberties.

To ensure trustworthiness of AI systems, digital forensics, and other forensic technologies, we believe that governments, forensic laboratories, and law enforcement agencies should be required to:

[26] IEEE Ethically Aligned Design.

- Ensure awareness, access, and research on the existence, fairness, safety, security, privacy, and ethical and societal impacts of digital forensics.

Governments should: (i) publicly identify and disclose the digital forensics used by the government; (ii) conduct and publicly disclose a methodological validation study that establishes the value of using new digital forensics in place of existing practices prior to deploying digital forensics; (iii) adopt clear procedures relating to the collection, usage, storage, and sharing of personal information in the context of developing, using, and validating a given digital forensics in a privacy-preserving manner; and (iv) prevent intellectual property, confidentiality claims, lack of funding, or lack of an designated independent body within government to monitor compliance from impeding duly limited independent validation and verification and publicly disclosed review of the fairness, safety, security, privacy, and ethical and societal impacts of digital forensics. Digital forensics ought to be submitted voluntarily to the agency performing validation and verification thereof, and the agency using related private intellectual property or proprietary data in its evaluation must adopt rules to protect such private rights from misappropriation.

Users and the public should be allowed to (i) request and receive an explanation of how a government determination using digital forensics was reached; (ii) determine whether the digital forensics used in government decision-making disproportionately impacts a protected class; and (iii) rectify, challenge, or complete inaccurate or incomplete personal data that is part of the digital forensics system or decision.

- Commit to removing barriers to parties' access to information needed to ascertain relevant evidence about and from digital forensics in legal disputes.

Specifically, in legal disputes where judges, juries, and lawyers are the users of digital forensics results, barriers to parties' access to information needed to ascertain relevant evidence about and from digital forensics should be eliminated.[27] Intellectual property protections should not be used as a shield to prevent duly limited disclosure of information needed to ascertain whether digital forensics meets acceptable standards of effectiveness, fairness, and safety. Specifically, in legal disputes, tribunals should permit disclosure under appropriate protective orders of intellectual property related to digital forensics necessary to obtain evidence in compliance with other judicial requirements, including constitutional requirements, discovery laws, or subpoenas. Furthermore, laws, procedures, and public funding should not make it more difficult for non-government parties in legal disputes to develop, obtain expertise regarding, or gain access to evidence from digital forensics than for government parties to do so.

- Ensure accountability and transparency in procurement and contracting for digital forensics.

To support awareness, access, and research on the existence, fairness, safety, security, privacy, and ethical and societal impacts of digital forensics, there must be accountability and transparency in government procurement and contracting for digital forensics. The government should not procure digital forensics that (i) require the governmental entity to indemnify vendors for any and all negative outcomes; (ii) do not adhere to the eight principles in IEEE's Ethically Aligned Design for creating and operating digital forensics that further human values

[27] For example, when source code is ordered to be provided, "information needed" requires providing sufficient information for the recipient to build, run, and test the software themselves including, at minimum:

- All software dependencies including third-party code libraries, toolboxes, plugins, frameworks, and databases;
- Software engineering and development materials describing the development, deployment, and maintenance of the version(s) of the software system used in the instant case, including software engineering documents and build instructions;
- All records of software glitches, crashes, bugs, or errors encountered during the developmental validation study;
- Software version numbers of the components of the system used for the developmental validation study; and,
- All records of unexpected results, including false inclusions, false exclusions and the conditions under which the unexpected results were achieved.

When source code is ordered to be provided, "access" requires, at minimum, that the source code be made available for inspection, in a format allowing it to be reasonably reviewed, searched, and tested, during normal business hours or at other mutually agreeable and reasonable times, and at mutually agreeable and reasonable locations.

and ensure trustworthiness (as may be reflected in articulated guidelines, standards, certifications, audits, and other sound documentation);[28] (iii) do not comply with federal, state, and local anti-discrimination laws; or, (iv) are shielded from independent validation and verification, and public review.

8. **To ensure digital forensics are reliable and trustworthy, governments should provide sufficient funding for testing, evaluation, certification, and investigation of digital forensics.**

Throughout the document, the Review highlights the value of government funding in the development of research on digital forensics (e.g., the tool specifications, test plans, and data sets for tool testing in Sections 4.10.4 and 4.10.5). We believe NIST should go further by including a KEY TAKEAWAY recommending that governments provide sufficient funding for testing, evaluation, certification, and investigation of digital forensics. The adoption and acceptance of digital forensics requires developing and sustaining public confidence in their quality, reliability, and compliance with regulations and social norms. Increased government funding for government and independent third-party evaluation and certification of digital forensics is essential to ensure efficacy, transparency, traceability, accountability, and competency. Development of design requirements, methods, metrics, and environments so that digital forensics can be tested and evaluated for interactions with different systems is critical in the adoption and acceptance of digital forensics. To this end, mechanisms must be developed for identifying and accounting for the features of digital forensics that could cause current testing, evaluation, certification, and investigation methods to misinform decision makers or the public about the risk of system deployment or the causes of system malfunction.

9. **Take inspiration from NIST's own Text Retrieval Conference (TREC) program.**

It appears that NIST may not be taking advantage of the agency's extant work. NIST's TREC is a strong model that could provide inspiration for the level of rigor in evaluation, measurement, and reporting that NIST seems to be aiming for in this document. While TREC is not directly applicable - the system focuses on the information retrieval phase that comes after the information collection phase - it could provide an example to follow.

[28] IEEE Ethically Aligned Design.

PC13

From: Nicola Morrow
Date: Mon, July 11, 2022 10:35 AM -0600
To: ScientificFoundationReviews <ScientificFoundationReviews@nist.gov>
Subject: Public Comment from NACDL's Fourth Amendment Center on NIST's "Digital Investigation Techniques: A Scientific Foundation Review"

Hello,

Attached please find a public comment from the National Association of Criminal Defense Lawyers' Fourth Amendment Center regarding NIST's draft report on digital forensics.

Sincerely,

Nicola Morrow
Summer Legal Intern
NACDL, 4th Amendment Center

NACDL
FOURTH
AMENDMENT
CENTER

July 11, 2022

National Institute of Standards and Technology
U.S. Department of Commerce
100 Bureau Drive
Gaithersburg, MD 20899

INTRODUCTION

The National Association of Criminal Defense Lawyers (NACDL) is the preeminent organization in the United States advancing the goal of the criminal defense bar to ensure justice and due process for persons charged with a crime or wrongdoing. NACDL serves as a leader in identifying and reforming flaws and inequities in the criminal legal system and ensuring that our members and others in the criminal defense bar are fully equipped to serve all accused persons at the highest level.

NACDL insists that any guidance regarding forensic searches of digital devices should, at the very least, acknowledge the legal and constitutional concerns that such searches present. As our lives migrate increasingly to the digital realm, computers and cellphones have become maps of our political and religious affiliations, our sexual preferences, our networks of social and professional relationships, and more. While the National Institute of Standards and Technology (NIST) is not responsible for resolving the complex legal and constitutional questions that digital forensic investigations raise, it would be irresponsible for NIST's technical guidance on this subject to ignore the privacy interests—and the attendant legal and constitutional concerns—that are implicated by sweeping intrusions into a person's digital life.

The purpose of this comment is not to challenge the technical guidance presented in NIST's report on digital forensics, but rather to urge NIST to include a disclaimer regarding the contested legality and constitutionality of the techniques outlined in its report. As technology evolves, so does Fourth Amendment jurisprudence on digital device searches. It is therefore imperative that any guidance on such a sensitive issue—particularly from a government agency—acknowledge the legal limits of its recommendations.

COMMENT

Computers, cellphones, and other digital devices are repositories of deeply personal information. In the pre-digital age, people conducted their financial business at the bank, their medical affairs at the doctor's office, their political activism in community centers, and their romantic dalliances on landlines and via letters. Now, all that information and more is consolidated on their digital devices. A computer or cellphone is less like a file cabinet than it is like the inside of a person's brain.[1] As the Supreme Court noted in *California v. Riley*—and again in *United States v. Carpenter*—digital devices contain "the privacies of life."[2] Because the data and metadata stored on—and accessible from—these devices are so revealing, courts are contending in real time with rapid technological evolution and the migration of quotidian private activities to the digital sphere.[3]

The Fourth Amendment of the United States Constitution protects against "unreasonable searches and seizures" and requires that all warrant applications be based upon probable cause and include a "particular[] descri[ption] [of] the place to be searched."[4] These requirements were

---

[1] *Contra* Nat'l Inst. of Standards and Tech., *Digital Investigation Techniques: A NIST Foundation Review* 7 (May 2022), https://perma.cc/MC2W-QFGY (analogizing categories of "real-world" evidence to "digital-world" evidence).
[2] *Riley*, 573 U.S. 373, 403 (2014) (citation omitted); *Carpenter*, 138 S. Ct. 2206, 2210 (2018).
[3] *See Public Safety, Privacy, and Particularity: A New Approach to Search Warrants for Digital Evidence*, Bloomberg Law (June 17, 2014), https://perma.cc/5WZZ-TCBS (describing circuit split on protocols for computer searches).
[4] U.S. Const., amend. IV.

codified in the Constitution largely in opposition to the general warrants and writs of assistance that "allowed British officers to rummage through [papers and effects] in an unrestrained search for evidence of criminal activity."[5] The probable cause requirement demands that there be "a fair probability that contraband or evidence of a crime will be found in a particular place"[6] and the particularity requirement defines a valid warrant as one that limits searches and seizures to evidence that is related to a specific crime.[7] Courts have spent the past half-century interpreting these requirements largely in the analog world; now, they are in the process of translating that jurisprudence to the wild west of the digital realm.

The debate over what the probable cause and particularity requirements mean when it comes to digital device searches is focused largely on the sort of investigative techniques that NIST describes in its draft report on digital forensics.[8] Digital forensics investigations frequently involve wholesale search and seizure of digital information. Courts are increasingly concerned with the constitutionality of such limitless searches and seizures. NIST's report recommends far-reaching digital forensics techniques, including copying entire hard drives,[9] recovering deleted data,[10] and more. While it may be that NIST's report recommends the most efficient and comprehensive methods for conducting forensic investigations of digital devices, it remains an open question whether those methods are constitutional. In fact, many courts have already held that indiscriminate searches and seizures of digital device data are unconstitutional.[11]

---

[5] *Riley*, 573 U.S. at 403.
[6] *Illinois v. Gates*, 462 U.S. 213, 238 (1983).
[7] *See Andresen v. Maryland*, 427 U.S. 463, 481–83 (1976).
[8] *See supra*, n.1.
[9] *See id.* at 34.
[10] *See id.* at 36
[11] *See, e.g., United States v. Burgess*, 576 F.3d 1078, 1091 (10th Cir. 2009) ("If the warrant is read to allow a search of all computer records without description or limitation it would not meet the Fourth Amendment's particularity requirement."); *United States v. Morales*, 77 M.J. 567, 573 (A. Ct. Crim. App. 2017) (holding that there was probable cause to search text messages on defendant's phone, but not photographs); *United States v. Winn*, 79 F. Supp. 3d 904, 922 (S.D. Ill. 2015) (holding that warrant containing "unabridged template that authorized the police to seize the entirety of the phone and rummage through every conceivable bit of data" was unconstitutional and that complaint established probable cause only for "a very small and specific subset of data on [the] cell phone"); *In re*

As courts continue to identify and define the constitutional boundaries of digital device searches, the technical guidance governing digital forensics will likely have to adapt to the evolving jurisprudential reality. Because the legal standards are in flux—and indeed, are hotly contested by parties interested in the privacy rights of criminal defendants—it is incumbent upon NIST to state in unequivocal terms that its recommendations are purely technical and do not necessarily conform to legal and constitutional requirements.

CONCLUSION

Digital devices contain vast quantities of deeply personal, far-ranging, and revealing information. Consequently, digital searches are qualitatively different from their analog equivalents. The Fourth Amendment's probable cause and particularity requirements raise special concerns for searches and seizures of digital information. In providing technical guidance on digital forensics, NIST should be careful not to misrepresent the legality and constitutionality of the investigative techniques that it describes. At the very least, NACDL recommends that NIST insert a disclaimer in its report acknowledging the ongoing legal debate over the permissible scope of digital device searches. NIST's report should further state, in explicit terms, that its guidance is purely technical in nature and does not reflect contemporary jurisprudence on the constitutional limits of digital device searches.

Nextel Cellular Tel., No. 14-MJ-8005-DJW, 2014 WL 2898262, at *13 (D. Kan. June 26, 2014) ("[J]ust as probable cause to believe that a stolen lawnmower may be found in a garage will not support a warrant to search an upstairs bedroom, probable cause to believe drug trafficking communication may be found in [the] phone's [] mail application will not support the search of the phone's Angry Birds application.") (citation and quotation marks omitted).

PC14

From: Tania Brief
Date: Mon, July 11, 2022 1:17 PM -0600
To: ScientificFoundationReviews <ScientificFoundationReviews@nist.gov>, Sarah Chu
Subject: Innocence Project, Inc.'s Comments on NISTIR 8354-DRAFT

Dear NIST Scientific Foundation Review report authors,

Thank you for the opportunity to provide feedback on the NISTIR 8354-DRAFT report. The Innocence Project's public comments are attached here.

Respectfully submitted,
Tania Brief

--
Tania Brief
Senior Staff Attorney, Strategic Litigation
she/her

INNOCENCE PROJECT

Executive Director
Christina Swarns, Esq.

Co-Founders & Special Counsel
Barry C. Scheck, Esq.
Peter J. Neufeld, Esq.

INNOCENCE PROJECT PUBLIC COMMENT ON NISTIR 8354-DRAFT

Digital Investigation Techniques: A NIST Scientific Foundation Review
July 11, 2022

The Innocence Project is pleased to respond to the National Institute of Standards and Technology (NIST) call for public comments regarding the NISTIR 8354-DRAFT report, Digital Investigation Techniques: A NIST Scientific Foundation Review (the "report"). For nearly thirty years, the Innocence Project has worked to exonerate the innocent and prevent wrongful convictions through systemic reform. Nearly fifty-two percent of the individuals exonerated by post-conviction DNA testing were convicted based, at least in part, on expert forensic evidence later shown to be erroneous. To improve the integrity of convictions and reduce the risk of an innocent person being found guilty, the Innocence Project urges robust gatekeeping and works to ensure that forensic evidence is admitted at trial only when it has strong scientific support, particularly from well-designed empirical studies.

With respect to digital forensics, this report embodies an opportunity to ensure that these tools are applied with transparency and proper safeguards. We commend the authors for actively disseminating information regarding their process at conferences across the country and now holding a public comment period to receive feedback.

However, we have some significant concerns regarding what we view as several oversights in the report that lead to overstated confidence in the results of digital forensic investigations and fail to account, in particular, for the role of human subjectivity and error or to recommend simple, noncontroversial strategies for the problems it identifies.

Moreover, this report does not seem to engage meaningfully with the reality that digital forensics is used within the context of the criminal legal system. It treats the application of digital forensics to life and liberty as incidental or separate from the technical issues of the digital investigative process. This lack of connection with the implications of the very serious problems of digital forensics (informal reviews, lack of validation, subjective analysis without cognitive bias guardrails, unnecessary intrusion into private information, and lack of transparency or documentation) is a major defect of this report.

We respectfully offer the following specific comments on the report.

INNOCENCE PROJECT

Comment No.	Page	Chapter	Text	Comment
1	1	Executive Summary	The overall finding of this report is that digital evidence examination rests on a firm foundation based in computer science. Several of the techniques had already been extensively studied and documented in the peer-reviewed literature. Others are documented more informally through community discussion forums	The mechanical process of manipulating computers is based on a firm computer science foundation, but the interpretation of the data may not be. How are attorneys supposed to assess the validity of techniques that are documented informally in community discussion forums? This report does not provide the public with a way to evaluate these informal techniques and seems to ask the reader to simply trust that they are sufficient.
2	1	Executive Summary	The application of these computer science techniques to digital investigations is sound, only limited by the difficulties of keeping up with the complexity and rapid pace of change in IT	Without a reference to the degree of uncertainty introduced by human examiners implementing the tools and analyzing the data, this statement is misleading. This sentence can be misapplied by parties seeking to deflect scrutiny from digital forensic examination. Please include the following statement to the last sentence to prevent misleading stakeholders regarding the validity and reliability of the entire digital forensic process: The application of these computer science techniques to digital investigations is sound, **but the reliability of the human interpretation of the gathered data is unknown. The digital forensic tool limitations include** ~~only limited by~~ the difficulties of keeping up with the complexity and rapid pace of change in IT. **The human interpretative component of digital forensic investigations includes incompleteness, inaccuracy, and misinterpretation.**

INNOCENCE PROJECT

Comment No.	Page	Chapter	Text	Comment
3	6	1	NIST also performed an interlaboratory study (Guttman et al. 2022) as part of its work on the scientific foundation of digital forensics. The study did not attract enough participants to draw meaningful conclusions but did demonstrate that digital forensic examiners could answer difficult questions related to the analysis of mobile phones and personal computers.	This passage refers to NISTIR 8412. It first states that the study did not attract enough participants to draw meaningful conclusions, but then draws the conclusion that "digital forensic examiners could answer difficult questions." While that statement is technically true, it is also misleading. Taken alone, the statement appears to reassure the reader of digital examiners' proficiency, however, Table 3 (p.8-10) of the black box study actually shows that examiners often get answers wrong. Twenty questions were asked in the study. In 9/20 (45%) of the questions, at least 27% of examiners gave wrong answers or skipped questions. In 3/20 (15%) of the questions, at least 49.4% of examiners gave wrong answers or skipped questions. Please revise this sentence to indicate that incorrect or skipped answers were frequent in this study.

INNOCENCE PROJECT

Comment No.	Page	Chapter	Text	Comment
4	6-7	1	After obtaining proper authorization and warrant for a search then a search can proceed. Digital evidence differs from physical evidence in the concept of search and seizure. For a physical search, the authorization covers searching the location and the seizure of objects of possible evidentiary value. In digital forensics an entire digital storage device, e.g., hard drive or flash drive is taken to then search it for evidence.	The analogy in pp. 6-7 to crime scene investigation is very helpful for understanding the specific digital forensics techniques covered in this report. However, this language should be revised to make clear that nothing inherent to the technology mandates that law enforcement look at every file on a digital device. Indeed, a frequent problem with warrants authorizing digital searches is that they are overbroad and insufficiently particular. While law enforcement may need to take possession of a physical device to perform a search, the warrant for a digital device—just like that for a physical search—must list the specific evidence that law enforcement has probable cause to believe is evidence of a crime, and there is no technological reason that law enforcement must look at individual files that go beyond those parameters. Indeed, the software addressed in this report allows for just such a targeted search. Moreover, it is essential that a defendant have a clear understanding of exactly how and by whom any search was undertaken, and this report should recommend further transparency with respect to disclosure and discovery.

INNOCENCE PROJECT

Comment No.	Page	Chapter	Text	Comment
5	7	1	In like manner, a digital investigation generates hypotheses, and the investigator searches for data artifacts, e.g., files, logged events with a time stamp, emails, etc., that can be used in evaluating observed evidence in light of alternative (opposing) hypotheses.	Using a hypothesis to drive a digital investigation presents a real risk that an investigator will seek out data to confirm that hypothesis—that is, a danger that cognitive bias will affect the investigation's outcome. Moreover, this approach exposes an individual's private information to unnecessary and unlawful exposure. This report does not sufficiently take these factors into account. Please include a recommendation that examiners be shielded from unnecessary biasing information and that the non-responsive personal information to which law enforcement has no legal right be protected from view. For example, the initial extraction of the entire contents of a digital device and culling down to items responsive to the terms of a search warrant should be undertaken by an examiner wholly unrelated to the investigation.
6	23	2	The capture-recapture method yielded a lower bound estimated population size of 11,000 with a 95% confidence interval of (9,900, 12,600). Due to the overlap between the lists and the fact that some of the total population has a zero probability of being selected in any list, the final value is interpreted as a lower bound estimate, rather than an absolute population size. This value of 11,000 US digital forensics organizations contrasts with the 409 publicly funded crime labs reported by the Bureau of Justice Statistics (Burch, Durose, and Walsh 2016). The decentralization of the digital forensics community in the United States is apparent in where digital forensics labs are found; they are not only in federal, state, and local crime labs, but also in prosecutor's offices, private consulting firms, and corporate cybersecurity operations.	This report does not communicate the urgency made apparent by the fact that there are so many digital forensics units in the U.S., but so little oversight in the form of accreditation, commissions, or laws to ensure the accurate and high-quality operation of these units. Moreover, the report fails to take into account the fact that it is private, for-profit companies that create the technology used in digital forensics and that these companies are incentivized to inflate the capabilities of their technology and disincentivized to be transparent.

INNOCENCE PROJECT

Comment No.	Page	Chapter	Text	Comment
7	20	2	KEY TAKEAWAY #2.4: The forensic examiner needs to be aware of key changes in computing technology relevant to the examination being performed. Frequent changes in digital technology introduces the possibility for incomplete analysis or for misunderstanding of the meaning of artifacts.	There should be a further recommendation concerning oversight and accreditation to address what this report identifies as necessary ongoing technical education. Moreover, it is important that any new understanding of a discredited approach or analysis be made transparent to the defense.
8	21	2	An examination of a mobile phone seized from a suspected drug dealer might begin by the examiner looking at contacts (possible customers and collaborators) and messages (setting up illegal transactions). To investigate a suspected espionage case the examiner might look for contraband (classified documents), removable device history (moving the contraband around), geolocator information (places the suspect has visited), contacts (identify collaborators), messages (extraction of planned actions) and deleted documents (hiding activity).	These searches would be appropriate only if specifically authorized by a valid search warrant. While it is true that all the searches described in this sentence might yield useful information, this sentence incorrectly implies that they are all automatic and legally proper in every investigation of a suspected drug dealer.

INNOCENCE PROJECT

Comment No.	Page	Chapter	Text	Comment
9	30	2	KEY TAKEAWAY #2.5: Not every digital forensic technique undergoes a peer review, formal testing, or error rate analysis. In general, the digital forensic community performs an informal review by providing feedback about the usefulness of techniques. This general acceptance process allows for techniques to be quickly evaluated and revised.	As we have learned through the widespread acceptance of now discredited forensic techniques such as bite mark analysis and hair microscopy, consensus and general acceptance does ensure valid techniques. This report does not communicate the urgent concern that much of digital forensic practice is informal and documented in fora that are outside the purview of the legal actors who must litigate and assess these techniques. This takeaway essentially asks criminal legal system stakeholders to simply put their trust in examiners without any way to verify their techniques. Please add language that communicates the problems that can occur in the criminal legal system if a discipline is built largely on informal reviews. Additionally, it is unclear why this Takeaway would not also include a recommendation to adhere to *IEEE 1012-2012: IEEE Standard for System, Software, and Hardware Verification and Validation* (IEEE 1012), a robust industry standard that already exists to set requirements for formal verification and validation of digital systems. The authors should definitively address in this Takeaway or in this section why they do or do not recommend the use of IEEE 1012.
10	33	4	4.1 Steps in a Digital Investigation	Please add a step to address the concerns regarding bias and unnecessary exposure of private information addressed in Comment 5, above. Moreover, the report should recommend that each of these steps be transparent to individuals whose information is being searched.

INNOCENCE PROJECT

Comment No.	Page	Chapter	Text	Comment
11	36	4	Cryptographic hashing is used to detect inadvertent or deliberate changes. Cryptographic hashing is a robust technique used in multiple high security applications. NIST publishes hashing standards as part of its cryptography program (NIST 2015a, 2015b).	This text comes from **4.3 Integrity Verification**. The integrity of the digital crime scene is essential to the validity of the examiner's analysis. If NIST publishes hashing standards that can help examiners or defense experts identify if inadvertent or deliberate changes were made to the extracted digital device data, why wasn't a recommendation made for digital examiners to use NIST's hashing standards? Please include a recommendation for digital forensic examiners to use NIST hashing standards as part of their digital forensic investigations.
12	37	4	KEY TAKEAWAY #4.1: When using techniques to recover deleted or hidden artifacts the examiner must determine the relevance of the recovered information as it may be incomplete or improperly merged with irrelevant information.	This takeaway makes clear how dependent the recovery of deleted data is on the judgment of the examiner. However, there is no discussion of how this process is documented, how examiners might make these judgments, and what protections can be made to insulate the examiner from cognitive biases. Since this process is so subjective, it will be essential for the defendant to have their own expert evaluate the forensic examiner's analysis. Please include language in this section regarding the need for documentation, transparency, cognitive bias protections, and defense expert access as mitigation for this very subjective process.
13	38	4	The main assembly of a narrative to describe the events of interest of an investigation or answering questions that arise during an investigation involves identifying, finding, and extracting relevant artifacts. A question of interest might prompt an examiner to select a specific artifact for examination. The examiner then tries to locate the selected artifact and then extract the artifact for examination.	This process has great potential to promote confirmation bias in the examiner, and this section does not describe any steps to mitigate confirmation bias. Please include language in this section regarding the need for documentation, transparency, cognitive bias protections, and defense expert access as mitigation for this very subjective process.

INNOCENCE PROJECT

Comment No.	Page	Chapter	Text	Comment
14	39	4	KEY TAKEAWAY #4.2: Searching tools have limitations based on the multiple ways that computers store information. Limitations include the type of files, types of encoding, and many other parameters. In general search tools are very effective at finding information, but there is a possibility that data will be missed because a tool does not have the capability to find it.	It is understandable that in conducting a digital search, not all data will be captured. However, in a criminal prosecution, missing data can have severe consequences. To mitigate any recovery problems, this section should recommend defense expert access to digital data. The Takeaway also references the limitations of the searching tools without offering parameters for assessing their conditions or impact. The limitations of any tool should be defined through validation processes. Please include language about the importance of defense expert access to digital devices given the fact that not all data may be captured in a search, as well as a requirement that technological limitations be documented through validation testing.
15	41	4	Artificial Intelligence (AI) tools use a technique called deep learning that can be used to uncover unseen relationships between case elements or search through data to recognize relevant items. Some AI applications have been controversial because of the introduction of unexpected, unintentional bias. Examples include facial recognition software exhibiting poor or misleading results for racial minority subjects (Grother, Ngan, and Hanaoka 2019). [...] AI tools are powerful, but not perfect and should be used with caution due to unexpected behaviors. What comes out depends on the data set used to train the AI and may not be relevant to the data at hand, and any results could be misleading and should be verified or confirmed. As with other techniques, examiner must use caution and check that AI based finding are used in the appropriate context.	The problems introduced by AI tools are serious, and the only recommendation that the authors offer is for the examiner to "use caution and check that AI based findings are used in the appropriate context." This recommendation is beyond insufficient and does not consider how cognitive biases may interfere in the "checking" process. Please include language in this section recommending the need for documentation, transparency, cognitive bias protections, and defense expert access as mitigation for these severe problems generated by AI tools.

Comment No.	Page	Chapter	Text	Comment
16	42	4	KEY TAKEAWAY #4.3: If someone has taken steps to change information in digital evidence to mislead an examiner, it may be difficult to detect the changes. Depending on the sophistication of the manipulation, identification of the changes relies on the skill of the examiner.	First, this Takeaway states that digital evidence can be manipulated, is difficult to detect, and that the criminal legal system is reliant on the skill of the examiner to identify this problem. However, this report does not include any language regarding how to proficiency test an examiner or determine their competence. NISTIR 8412 has demonstrated that examiners make mistakes and also indicated that 71-75% of participants passed a mobile or hard drive proficiency test in the past five years (see pp. 19 and 31). If we are to give examiners this much trust, there must be a way to demonstrate their capacity. Please include language in this section recommending the need to develop robust competency or proficiency testing programs and the frequency with which they should be given. Second, this Takeaway asserts that information can be changed to mislead a forensic examiner without any assessment of the technical capabilities necessary to undertake such measures or which kinds of files or devices would be most susceptible to such manipulation. The report should clarify what kind of technological capabilities would be necessary to, for example, disguise one type of file commonly found on a phone as another one. Such a scenario is often used by law enforcement to justify unnecessarily overbroad searches of digital devices.

INNOCENCE PROJECT

Comment No.	Page	Chapter	Text	Comment
17	41	4	For a forensic technique or method to be considered validated it should be shown to be fit for purpose otherwise defined as "the process of providing objective evidence that the method is good enough to do the job required by the end user". Validation can give a false indication of "fitness for purpose" that becomes apparent later.	This language flags an important problem, but does not provide more information about what happens if a technique or method is later determined to not be fit for purpose. What are the consequences of this? How might it impact an analysis? What happens to the people who have been adversely impacted by the flaws in the technique or method?
18	43	4	There have been several papers published on validation of digital forensics methods (Regulator 2020; Arshad, Jantan, and Abiodun 2018; Beckett and Slay 2007; Brunty 2011; Casey 2011a; Craiger et al. 2006; Guo, Slay, and Beckett 2009; Horsman 2018; Horsman 2019; Marshall and Paige 2018; Risinger 2018; SWGDE 2014; Wilsdon and Slay 2006). Some of these papers seem to confuse validation of a method and verification of a software tool and try to fold the two activities together instead of keeping them separate. The guidance from the UK Forensic Science Regulator (Regulator 2020) seems the most clear and includes consideration of risk assessment of the method, documentation of acceptance criteria and possible outcomes.	Given the central role of validation in the forensic science process, it would be important for the authors to assist the readers in better understanding which publications among those mentioned in this section properly or improperly describe validation and verification. The manner in which these reports are described here skirts an important issue, and the criminal legal system audience needs these experts to point out which publications we can rely upon rather than leaving it to a lay audience to debate. Additionally, if the UK FSR is the resource with the best guidance, then the authors should explicitly make that clear. It is also unclear why the authors do not include IEEE 1012 on this list. Please either recommend the UK FSR guidance if it is the best resource for validation and verification information for digital forensics methods or list the publications among those listed that provide accurate guidance on the topic. Additionally, the authors should definitively address why they do or do not recommend the use of IEEE 1012.
19	43	4	The general validation and verification for a given version of a tool can be done once. It does not need to be performed by every lab.	Is this true for every digital forensics tool? How do we know which digital tools require verification by specific labs and which do not?

INNOCENCE PROJECT

Comment No.	Page	Chapter	Text	Comment
20	45	4	Another problem is that the properties and characteristics of digital data changes with the software environment as the technology evolves over time and an error rate valid at one point in time might not apply at any other point in time. KEY TAKEAWAY #4.4: Digital processes tend to have systematic rather than random errors. Therefore, an error mitigation analysis provides more information and is the correct way to manage uncertainty. Asking for an error rate is only useful where there are random errors. KEY TAKEAWAY #4.5: When error rates are provided, it is important for the user to understand the context of the numbers. Errors in computer science techniques tend to be so small as to be negligible. For some forensic techniques, the error rates may vary significantly based on attributes of the technology and usage patterns.	The focus on error rates here is a bit of a deflection of the real problem. While the technical application of tools may not incur error, the interpretation of collected data is not well studied (see Black Box Study). It is this interpretive phase that introduces very serious biases and can mislead investigations. The statement in takeaway #4.5 that "Errors in computer science techniques tend to be so small as to be negligible" can be misunderstood and misused by examiners seeking to boost their credibility if it is not also paired with a statement that explicitly states that errors in the interpretive portion of the process are unknown. Lastly, this section of the report does not provide evidence for the statement "errors in computer science techniques tend to be so small as to be negligible." Even small errors have been demonstrated to have serious consequences, and courts have not been an effective arbiter of this risk.[1] Please provide evidentiary support in this section for the statement regarding negligible error rates. If no supporting evidence can be found, the statement must be removed from the Takeaway. Please also add language to Takeaway #4.5 to state the following: TAKEAWAY #4.5: When error rates are provided, it is important for the user to understand the context of the numbers. Errors in computer science techniques tend to be so small as to be negligible. For some forensic techniques, the error rates may vary significantly based on attributes of the technology and usage patterns. **Errors in the human component, the interpretation of the data gathered through the computer science techniques, is unknown. Therefore, a complete analysis of digital forensic techniques requires an evaluation of the application of the tools and the interpretation of that data. There is a dearth of research in the latter.**

INNOCENCE PROJECT

Comment No.	Page	Chapter	Text	Comment
21	46	4	4.10.2 Observed Errors	This section describes the problems with forensic tool implementations, namely incompleteness, inaccuracy, and misinterpretation. It does not offer a way to detect these errors, nor does it recommend that these errors be corrected when identified. This section is incomplete without addressing these two issues.
22	56	5	5 Conclusions	This section does not suggest cognitive bias protections, nor does it recommend a mechanism for identifying and remediating errors and notifying impacted parties when these errors occur. Please include language referencing these recommendations in this section.
23	56	5	The application of these computer science techniques to digital investigations is sound, only limited by the difficulties of keeping up with the complexity and rapid pace of change in IT.	Please see Comment #2, above.

[1] *See, e.g.*, https://www.pbwt.com/second-circuit-blog/second-circuit-oks-use-of-now-defunct-dna-testing-method

PC15

From: Daniel Kahn Gillmor
Date: Mon, July 11, 2022 5:34 PM -0600
To: ScientificFoundationReviews <ScientificFoundationReviews@nist.gov>
Subject: review of NIST IR 8354-draft (forensic technology)

Hi good folks at NIST--

Thanks very much for your work on IR 8354-draft on the scientific basis and fundamental requirements of digital forensic technology.

I'd like to submit the attached review of the draft, which opens with some substantive feedback and concludes with some minor nit-picking of the sort I always appreciate myself on public documents I work on. I'd be happy to talk more about either IR 8354 or my feedback, if you have any questions or followup suggestions.

I hope I'm not too late in submitting this feedback!

Regards,

--dkg

--
Daniel Kahn Gillmor
Senior Staff Technologist
Speech, Privacy, and Technology Project
American Civil Liberties Union

Feedback on `NIST IR 8354-draft`: "Digital Investigation Techniques: A NIST Scientific Foundation Review"

Daniel Kahn Gillmor*

2022-07-11

Contents

This is a brief review of "Digital Investigation Techniques: A NIST Scientific Foundation Review"[1].

Major points

Problematic Analogies and Procedural Risk

The analogies described in the introduction sections to physical searches may not be entirely accurate. For example, Table 1-2 suggests that "on site records such as a filing cabinet or desk" are comparable to "Files stored on the computer hard drive, removable media". But a desk and filing cabinet together can only hold a relatively small amount of data compared to the information that might be stored on a computer hard drive. Even more troubling, a digital device's

*Individual submission, organization affiliation indicated for identification purposes only.

[1]NIST IR 8354-draft, https://www.nist.gov/forensic-science/digital-investigation-techniques-scientific-foundation-review

storage device might contain records of a person's life collected incidentally in the course of daily activity, such as a browser cache, geolocation history, call records, etc., whereas a desk and filing cabinet typically only contain what the user has deliberately chosen to physically store.

Even where the analogies are plausible, the document doesn't spell out some basic legal risks involved with handling the data in question. For example, a reasonable narrowly-tailored warrant in a case involving wage theft might permit search of a business for purposes of seizing employment and payment records. Such a warrant would *not* cover collection or retention of data related to which newspapers are delivered to the business, personal communications about unrelated matters, or industrial processes.

However, a single seized digital device could contain all of this information. If the seized data is not selected carefully, handled diligently, and disposed of properly, it would be easy for an investigator to overstep the bounds of a reasonably-scoped judicial order. The consequences of an investigator overreaching the powers granted to them by warrant may vary across jurisdictions, but such overreach is never appropriate. Despite the document describing several different abstract goals of digital investigators (for example, on page 8), no text in the draft discusses what this kind of careful selection or diligent handling should look like.

While the beginning of §4 ("Scientific Foundations of Specific Tasks") acknowledges potentially thorny legal requirements, the numbered list of "main digital forensic tasks" in the same section fails to mention them.

A forensic tool should be able to filter data—for example by category, date, and keyword—to limit both data seizures and data searches. The tool should have an audit trail of searches executed and returned data in order to enable judicial oversight of searches. Data relevant to the investigation should be segregated and the non-responsive data should be sequestered and ultimately returned or deleted.

Clearly, this document cannot provide comprehensive legal analysis, but it should at least acknowledge the risks of overcollection and overretention, and highlight responsible behavior as a requirement for any serious investigator.

Recovering Deleted Files on Systems that use Encryption

Deleted file recovery is significantly more problematic than described in the draft due to potential for encryption on either a per-volume or per-file basis. This is comparable to someone shredding or burning their paper files, rendering them similarly inaccessible to forensic analysis.

In particular, §2.6.3 ("Deleting Files") should probably note the impact of per-file encryption on deletion. In per-file encryption systems, the metadata describing a particular file contains a key, and the key encrypts the file contents. Rapid secure deletion is possible on such a system if the metadata can be completely

wiped as the content blocks are added back to the unallocated store; because they're encrypted and the key is unrecoverable, the file itself is unrecoverable.

Rapid secure deletion should be relevant for at least two widely-available file-based encryption systems: Apple's "Data Protection"[2] on APFS, and Linux's `fscrypt`[3] on ext4, F2FS, and UBIFS.

Similarly, the bullet point about "physical acquisition" in §4.2.2 ("Mobile Device Acquisition") probably also needs adjustment. It should not claim that physical acquisition "allows recovery of deleted data" unequivocally, as cryptographic key material might be missing, leaving the acquired data in an unreadable state.

Chain of Custody and Reactions to Failed Tooling

The document does a good job of highlighting forms of failure and error that digital forensic tools may exhibit, as well as acknowledging the ever-shifting technical and informational landscape. These are daunting prospects for any digital investigator.

However, the document fails to draw some of the straightforward conclusions that should follow. For example, while the document discusses verification and validation of tools, and chain of custody for specific *data*, it does not describe comparable precautions for the *tools* involved in a given investigation.

Given that a particular version of a tool may be found to be flawed after an investigation is complete, a responsible investigator must keep track not only of the data analyzed, but the specific versions and tools used for the analysis. In the event that a version of a specific tool is found to be flawed, the investigators should promptly flag any resulting conclusions as problematic both internally and externally. For example, a forensic examiner should formally and promptly notify any court and prosecutor that may have relied on representations made as the result of the flawed tool. If any underlying data are still available, and a fixed version of the tool has been produced, the investigators may also choose to re-run their analysis on the remaining data with the updated tool.

This concern is not just about specific versions of specific software, but also about specific devices. For example, a forensic analysis machine typically contains a large amount of diverse software to enable it to read and inspect a wide range of file formats. But the more software the machine includes, the more likely there are to be bugs lurking somewhere in it, including potentially serious security vulnerabilities that can be triggered when parsing arbitrary input from a seized digital device. The forensic analysis machine itself could be compromised when it scans any particular device that has a weaponized file in it.[4]

[2] Apple Platform Security, "Data Protection Overview", https://support.apple.com/guide/security/data-protection-overview-secf6276da8a/1/web/1

[3] Linux Kernel Source, "Filesystem-level encryption (fscrypt)", https://github.com/torvalds/linux/blob/master/Documentation/filesystems/fscrypt.rst

[4] Signal Private Messenger blog, "Exploiting vulnerabilities in Cellebrite UFED and Physical Analyzer from an app's perspective", https://signal.org/blog/cellebrite-vulnerabilities/

While the document includes a wide range of capabilities a forensic examiner should consider for their tools, and it encourages re-testing tools after an upgrade, it doesn't recommend re-testing a complex and potentially-vulnerable tool after exposure to arbitrary data. Can a specific analytic machine be audited to ensure it is operating as expected and that it has not been compromised?

If a given investigation unit has several of the same devices from a given manufacturer, and they are all running the same version of the device software, one of the devices could be compromised independently of the others. If that happens and is detected, then investigations that involved that specific device (as well as any other machines that may have connected to the compromised device) need to be considered as potentially tainted. Again, a responsible investigator should flag tainted investigations both internally and externally when such a circumstance arises, and may want to re-run any tainted analysis.

Statistical Failures in Digital Forensics

§4.10.1 ("Error Rates") rightly observes classes of error in digital forensic tooling: for low-level tools, systematic error as opposed to statistical or stochastic error predominates. However, statistical and stochastic errors do exist, and even predominate in certain classes of digital forensic activity.

For example, §4.6.2 ("Locating Contraband") describes some forms of probabilistic matching of contraband data. An advanced forensic tool that creates an index for text search over a variety of files may use optical character recognition on images, or speech-to-text on audio recordings to be able to add those files to the index. Likewise, forensic scanners tasked with finding images of specific places, objects, or people (e.g. searching an archive of seized videos for the appearance of a specific individual) might use sophisticated image analysis, including object or facial recognition.

These advanced digital forensics not only have statistical error modes, but the advanced models they use may be subject to significant bias[5] or inappropriate training[6] which could could have a significant impact on the conclusion drawn by the forensic analysis. Any forensic analysis that draws on tools of this level of sophistication needs to record and report not only the specific software and version used for analysis, but also the range of training datasets and model parameters used during the investigation.

[5] Allison Koenecke et al., "Racial disparities in automated speech recognition", https://doi.org/10.1073/pnas.1915768117

[6] Kate Crawford and Trevor Paglen, "Excavating AI: The Politics of Images in Machine Learning Training Sets", https://excavating.ai/

Minor Clarifications

Scope

While §5 ("Conclusions") and §1.1 ("Scope") both try to scope the draft to digital forensics related to recovery of data from seized hardware, the boundaries of what is in-scope for the draft aren't tightly held. For example, some of the sample courses and journals listed appear to be focused on the out-of-scope network forensics. And §4.2.5 ("Social Media Acquisition") covers data from live, proprietary network services.

It be better if the document were to at least explicitly identify out-of-scope topics when they do come up, to avoid the appearance of incomplete analysis.

Steganography

The steganography discussion in §4.7.2 ("Anti-Forensics") focuses on image-based steganography without taking into account that the most popular image representation formats (like JPEG) are lossy, and do not have stable low-order bits in any pixel channel. In other words, compressing a raw image with JPEG will typically render steganographically-injected data unrecoverable if the steganography is applied in the form described here. So this particular approach for steganography doesn't make sense with any comparable image format. It's possible to use the low-order bits in JPEG's DCT blocks (as described in the Rodriguez and Peterson 2007 citation), but that's not the same thing as storing the data in the least-significant bits of a pixel channel itself. The same concerns apply to using comparable techniques to hide data in low-order bits of audio samples, when those samples are lossily compressed (e.g., MP3 is also lossy).

The steganography discussion also makes a vague representation that an unmodified picture would have "color values of the pixels [that] cluster around the dominate colors in the picture". It's unclear to me what this phrasing means, and there is no reference for the claim. Surely it does not apply to all types of image data. For example, the distribution of colors in graphically designed objects like logos differ significantly from the distribution of colors in photographs taken under natural light.

If the goal of this section is to describe key aspects forensic work on potentially-steganographic datasets, it should probably:

- Acknowledge the interaction between lossy media compression algorithms and steganography.
- Consider the not-uncommon combination of steganography and cryptography—data can be encrypted before steganographic injection, making steganographic detection and recovery even more difficult.
- Describe the use of steganography in processes like digital watermarking, which may be related to attribution and provenance for digital investigations.

Nitpicks

- page i: "(operating systems and applications are revised)" should be "(operating systems and applications) are revised"
- page 15: the "international standard for dates" should explicitly refer to ISO 8601.
- page 19: "HSF+" should be "HFS+"
- page 42: discussion of "16-bit pixels" and their least-significant bits seems to assume that each pixel is a single channel (that is, contains a single number), but most modern image formats have at least three channels (red, green and blue) per pixel. The description in this section is presented as a concrete example, but it does not align with practical reality.
- page 42: the discussion of "left most" and "right most" bits appears to assume a particular endianness of the integer representation. However, some architectures and some file formats may use big-endian representations of integers, and others may use little-endian representations. It's probably better to just say "least significant" or "most significant" when referring to specific bits here.
- page 43: the list of "requirements [...] needed for a hash algorithm to be fit for purpose in a forensic context" includes "Original message cannot be recovered or reconstructed from the hash value." This is indeed a requirement for cryptographic hashes (and it is included appropriately elsewhere in the document where cryptographic hash requirements are listed). And it's a good idea to use a cryptographic hash for forensic hashing due to extensive applied analysis of these algorithms. But this specific property is probably *not* needed for forensic hashing purposes. If there is a forensic requirement for this property, it needs more explanation in the document. Otherwise, it should be omitted from this list.