Scientific Working Group on Digital Evidence

Best Practices for Research in Digital Forensics

SWGDE 26-F-001-1.0

The version of this document is in draft form and is being provided for comment by all interested parties for a minimum period of 60 days.

Disclaimer Regarding Use of SWGDE Documents

SWGDE documents are developed by a consensus process that involves the best efforts of relevant subject matter experts, organizations, and input from other stakeholders to publish standards, requirements, best practices, guidelines, technical notes, positions, and considerations in the discipline of digital and multimedia forensics and related fields. No warranty or other representation as to SWGDE work product is made or intended.

SWGDE requests notification by email before or contemporaneous to the introduction of this document, or any portion thereof, as a marked exhibit offered for or moved into evidence in such proceeding. The notification should include: 1) The formal name of the proceeding, including docket number or similar identifier; 2) the name and location of the body conducting the hearing or proceeding; and 3) the name, mailing address (if available) and contact information of the party offering or moving the document into evidence. Subsequent to the use of this document in the proceeding please notify SWGDE as to the outcome of the matter. Notifications should be submitted via the SWGDE Notice of Use/Redistribution Form or sent to secretary@swgde.org.

From time to time, SWGDE documents may be revised, updated, deprecated, or sunsetted. Readers are advised to verify on the SWGDE website (https://www.swgde.org) they are utilizing the current version of this document. Prior versions of SWGDE documents are archived and available on the SWGDE website.

Redistribution Policy

SWGDE grants permission for redistribution and use of all publicly posted documents created by SWGDE, provided that the following conditions are met:

1. Redistribution of documents or parts of documents must retain this SWGDE cover page containing the Disclaimer Regarding Use.
2. Neither the name of SWGDE nor the names of contributors may be used to endorse or promote products derived from its documents.
3. Any reference or quote from a SWGDE document must include the version number (or creation date) of the document and also indicate if the document is in a draft status.

Requests for Modification

SWGDE encourages stakeholder participation in the preparation of documents. Suggestions for modifications are welcome and must be submitted via the SWGDE Request for Modification Form or forwarded to the Secretary in writing at secretary@swgde.org. The following information is required as a part of any suggested modification:

DRAFT

a) Submitter's name
b) Affiliation (agency/organization)
c) Address
d) Telephone number and email address
e) SWGDE Document title and version number
f) Change from (note document section number)
g) Change to (provide suggested text where appropriate; comments not including suggested text will not be considered)
h) Basis for suggested modification

Intellectual Property

All images, tables, and figures in SWGDE documents are developed and owned by SWGDE, unless otherwise credited.

Unauthorized use of the SWGDE logo or document content, including images, tables, and figures, without written permission from SWGDE is a violation of our intellectual property rights.

Individuals may not misstate and/or over represent duties and responsibilities of SWGDE work. This includes claiming oneself as a contributing member without actively participating in SWGDE meetings; claiming oneself as an officer of SWGDE without serving as such; claiming sole authorship of a document; use the SWGDE logo on any material and/or curriculum vitae.

Any mention of specific products within SWGDE documents is for informational purposes only; it does not imply a recommendation or endorsement by SWGDE.

DRAFT

Scientific Working Group on Digital Evidence

Best Practices for Research in Digital Forensics

Table of Contents

1. Purpose and Scope 2
2. The Imperative for Research 2
3. Disclaimer 2
4. Limitations in Performing Digital Forensic Research 3
5. Addressing Bias 3
6. Research Process - The Scientific Method 4
   - 6.1 Identify Research Gaps 5
   - 6.2 Specify Scope of Research 5
   - 6.3 Draft Hypothesis 5
   - 6.4 Design 5
   - 6.5 Perform Baseline Acquisition 8
   - 6.6 Test Data 8
   - 6.7 Data Analysis 9
   - 6.8 Conclusion Phase 9
   - 6.9 Peer and Administrative Review 10
   - 6.10 Report and/or Publish Results 10
7. Peer Review Process 11
   - 7.1 Hypothesis 12
   - 7.2 Evaluation 12
   - 7.3 Provide Feedback to Author 13
8. Sample Templates 13
9. References 20
10. Additional Resources 21
11. History 22

DRAFT

1. Purpose and Scope

The purpose of this document is to provide best practices for understanding the meaning of a digital forensic artifact and providing evidence to support that conclusion. This process can be used to support both general research about artifacts and to support case specific testing as described in *SWGDE 25-Q-001-1.0 Quality Management System* [QMS] *for Digital and Multimedia Evidence Units*. The QMS provides a framework for tool and technique reliability which includes both a tool selection process and more specific testing which may be needed to support specific casework. This document also supports the NIST Interagency Report 8589 *Validation in Forensic Science: Guiding Principles for the Collection and Use of Validation Data* [1], which states that case specific uses of forensic techniques should have data that is pertinent to that case.

The guidelines in this document apply to any individual or organization that is conducting forensic artifact research, including digital forensic examiners and digital forensic tool developers.

A digital forensic examiner or researcher is not required to conduct additional research if they determine existing research was performed properly and/or reach a proper conclusion addressing the digital forensic artifact.

2. The Imperative for Research

There are many reasons a forensic examiner may conduct research including, but not limited to:

- Support the forensic examiner in reporting and testimony
- Support the admissibility of expert testimony that is subject to review in accordance with admissibility rules [2]
- Academic pursuits
- Analysis of new or updated hardware, software, artifacts, or systems
- Analysis of poorly understood or contested artifacts
- Testing of toolset artifact interpretations
- Identification/clarification of forensic toolset limitations and errors
- Development and update of forensic toolsets and methodologies
- Laws, regulations, or training materials
- Lack of prior peer-reviewed research

3. Disclaimer

This document is not intended to be used as a step-by-step guide for researching and verifying digital forensic artifacts, nor should it be construed as legal advice. The best practices in this document reflect a consensus among examiners and professionals in digital forensics. Those who find that they lack the competency or equipment to conduct research in line with these best practices are encouraged to seek the assistance of qualified experts in lieu of performing research to a lower standard.

4. Limitations in Performing Digital Forensic Research

Original evidence (e.g., mobile device) shall not be used when performing research if it could result in destruction or alteration of data on the evidence device.

Digital forensic research cannot address all possible variations in software, systems, and devices. This does not nullify the need for research. Examiners shall verify results. Verification includes, but is not limited to, conducting additional research or leveraging additional tools. Document known limitations which may include, but are not limited to:

- Tool and/or testing dataset availability
- Issues replicating the source software, systems, or conditions
- The passage of time may result in the inability to create an optimal testing environment
- Timespan of testing data vs evidence data
- Proprietary source code

5. Addressing Bias

Biases in digital forensic artifact research can significantly affect the accuracy, reliability, and objectivity of findings. Examiners conducting digital forensics research must be mindful of potential biases that can influence the validity of results and/or their interpretation of data. Limiting overall case involvement to necessary information and implementing safeguards is essential to address this risk.

The selected bias descriptions outlined below are not all-encompassing but intended to provide an awareness for the digital forensic examiner to address.

Bias in the context of digital forensics and this paper is categorized into two separate categories: cognitive and research. Cognitive bias is personal bias, whether implicit or explicit, the researcher brings to their analysis and interpretation of results, e.g. ignoring data that doesn't fit the desired outcome. Research bias occurs when the reliability and validity of test results are impacted by limitations in the design, sample, methods, or tools.

To mitigate these biases, examiners shall implement strategies such as the following:

- Training on bias recognition – Training makes researchers more aware of common cognitive and methodological biases. Being aware of these biases, researchers can better identify and mitigate them in their own work [3].
- Establishing communication boundaries – The order in which information is received can shape how evidence is interpreted and the conclusions reached. First impressions are remembered more strongly and can shape the entire analysis [4]. Limiting outside discussions insulates researchers from external influence.
- Maintaining detailed documentation – This may include explaining how the methodology was developed, providing the rationale for tool selection, and documenting the results

which may include any contradictory information observed. This allows others to trace how conclusions were reached and spot areas where subjective choices may have influenced outcomes.

- Blind analysis procedures – Design and execution of the test methodology prior to viewing and analyzing the results reduces confirmation bias.
- Peer review protocol – Independent reviewers provide an external check on assumptions, methodology, and interpretations. This helps identify overlooked biases that arise when researchers are too close to their own work.

6. Research Process - The Scientific Method

Documenting each digital forensic artifact research step should be completed using an organization-approved, standardized, artifact research template, (refer to example research template). Figure 1 is a visual flow chart for suggested workflow of the research process.

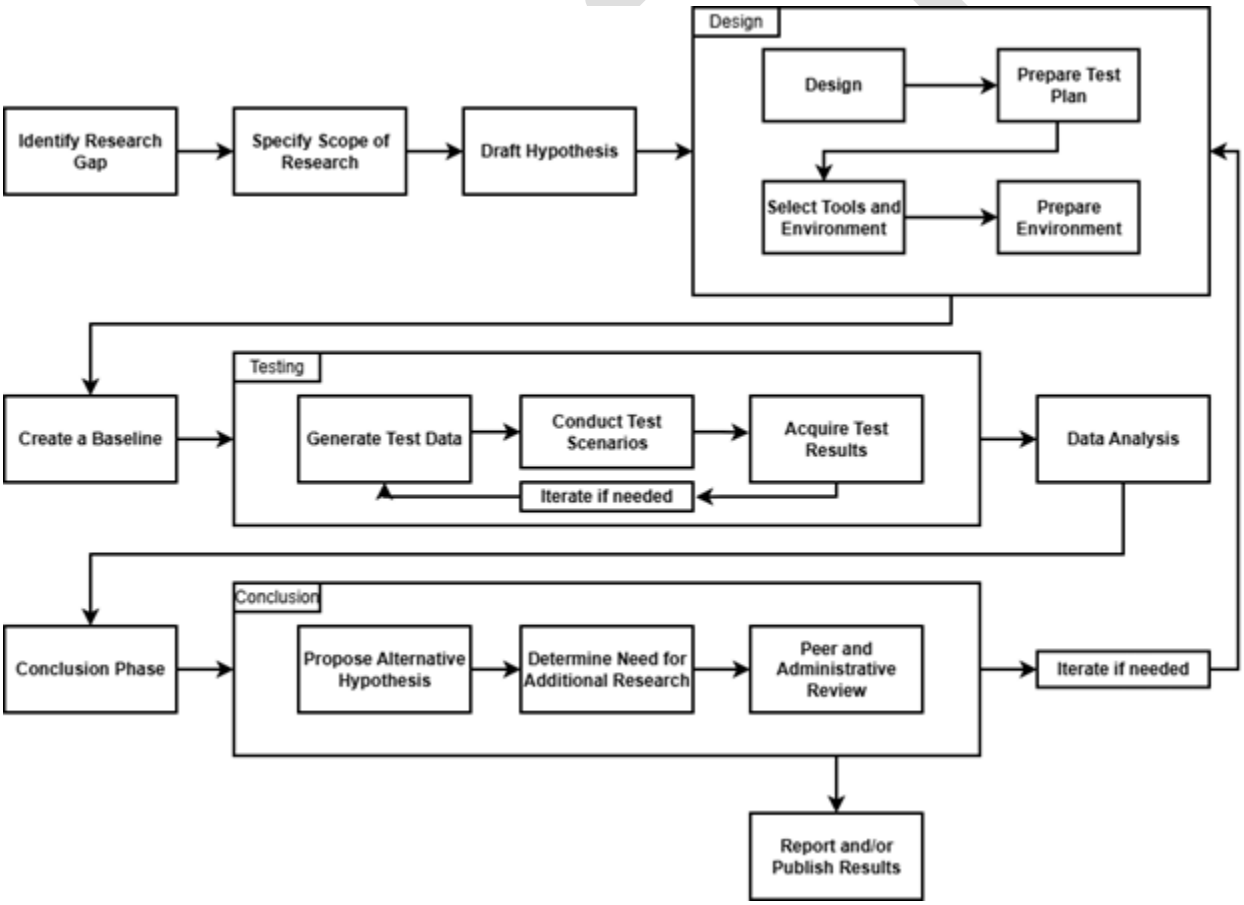


Figure 1. Suggested workflow of the research process (Figure credit: SWGDE)

6.1 Identify Research Gaps

A research gap becomes evident when there is a need to address a lack of established knowledge regarding an artifact that has not been covered by existing research or studies in the field. This commonly occurs during digital forensic investigations where the origin or behavior of an artifact is not widely understood as addressed in Section 2.

A literature review is used to determine what research has already been completed, and what methodologies or analyses have been utilized. Literature review shall be included in the final report documentation to clearly show the similarities, differences, and previous research gaps addressed.

6.2 Specify Scope of Research

The scope of research shall be defined based on the specific, testable question(s) that address the identified gap, including

- the forensic artifact(s) at issue
- the method used to create or interpret the artifact(s)
- the decision context in which the results will be relied upon

Narrowing the scope of research can help to improve the specificity of the findings and mitigate limitations.

6.3 Draft Hypothesis

The hypothesis proposes an expected outcome about the source or behavior of digital artifacts which is then tested through research [5]. It establishes a testable prediction, connects to pre-existing research, and guides methodological decisions to be used to measure variables [6,7].

The hypothesis should be drafted with a narrow scope and make a definite claim that can be tested based on observable evidence. The hypothesis should adhere to scientific principles, including falsifiability, meaning the evidence could confirm or refute the hypothesis [5].

6.4 Design

The design phase outlines a plan to be followed in testing the hypothesis. This plan takes the research gap, scope, and limitations into consideration.

Digital artifacts can result from the interoperability of hardware components, operating system processes, software applications processes, and user inputs. Examining an artifact through source code may reveal the condition of the artifact (e.g., a true or false indication) but may not reveal how an artifact is created. To understand an artifact's behavior, an examiner needs to test the conditions that trigger the creation of the artifact.

Access to source code is not needed to determine the cause of an outcome and may not advance forensic research. Providing an input to a control will result in an output. If that output is consistent and repeatable, then the output can be conclusive even outside of the understanding of how the output occurred through reviewing of source code.

6.4.1 Prepare Test Plan

Preparing a test plan is a component of the design phase that outlines the structured approach for evaluating the hypothesis. The primary goal is to achieve reliable, repeatable, and defensible results while minimizing bias. The examiner should identify sources of potential errors and develop steps in the test plan to mitigate them [8]. For sample test plan templates, please see the *SWGDE Recommended Guidelines for Validation Testing* [6]. For error mitigation techniques, please see *SWGDE 12-Q-001-2.0 Establishing Confidence in Digital and Multimedia Evidence Forensic Results by Error Mitigation Analysis* [8].

Key Elements of the Test Plan:

- Define the purpose, scope, and hypothesis
- Outline the requirements
- Establish procedures for documentation
- Establish the methodology (e.g., testing procedures, specific user actions, and environmental conditions)
- Identify variables and controls
- Prepare appropriate datasets (including known values)
- Document and establish the testing environment (e.g., tools, platforms, software versions, system configurations)
- Describe the test data
- Establish acceptance criteria

6.4.2 Select Tools and Environment

When conducting research, examiners shall consider the environment in which the testing occurs. Original evidence shall not be used when performing research if it could result in destruction or alteration of data on the evidence device. The device model, operating system, and application versions all form part of this environment and can influence the reliability of test results.

6.4.2.1 Establishing a Testing Environment

The establishment of a testing environment will be defined by four tiers. Tier 1 is optimal research conditions; through Tier 4 being a research environment that has little to no similarity to the original case study conditions.

Tier 1: Analogous: For a testing environment to be considered analogous, the examiner shall use the same device model, operating system, security patch level, and application version in the test environment as the environment that prompted the research question. The test environment may also include the physical environment such as location, weather, interaction with other nearby devices and/or assorted other environmental factors.

Tier 2: Sufficiently Analogous: For a testing environment to be considered sufficiently analogous, the examiner shall test with a device model and/or software version that came before the target environment and version that came after and confirm that the results are consistent.

The examiner should obtain application versions as close as possible since the extrapolation of the results become less reliable as the differences in versions increase. The examiner shall also confirm that any changelogs for incremental versions do not indicate changes that would affect the artifact. For example, if the research gap concerns how the Telegram iOS app associates images with messages in version 12.2.2, and the examiner cannot acquire that particular version for testing, they can test version 11.15 and 12.3.1.

Tier 3: Reasonably Analogous: Devices that are close to the original evidentiary media's operating system versions or security patch levels in that there is confidence that artifact findings are consistent between the two. Additionally, to meet this standard, an examiner shall have other supporting evidence to indicate that the changes in these versions would not cause different results in the testing.

If it is not possible to test a version before and after, such as in the case of an artifact on a Software as a Service (SaaS) platform where the examiner can only access the current production version, consideration should be taken for how much deviation from the original environment (if any), such as newer OS versions or updated applications, can be tolerated while still producing results that remain valid [7]. Examiners shall carefully evaluate and document these environmental differences, explain whether they are acceptable, and demonstrate how the underlying data remains trustworthy despite such changes. This transparency ensures that results can withstand scrutiny and that the validity of findings is not undermined by unacknowledged environmental shifts.

Tier 4: Non-Analogous: Test environments that do not meet the criteria of Tier 1-3 are non-analogous. When an analogous test environment cannot be generated due to a lack of resources, tools, or external factors, alternative methodologies should be considered.

The examiner should consider the tools employed, the accuracy of the tools, and address that in their findings. When considering possible sources of errors, error mitigation analysis is more appropriate than error rates for digital forensic research. For more information on applying error mitigation analysis, please see the NIST publication *NISTIR 8354: Digital Investigation Techniques: A NIST Scientific Foundation Review* [7], *SWGDE 12-Q-001-2.0 Establishing Confidence in Digital and Multimedia Evidence Forensic Results by Error Mitigation Analysis* [8], and the SOLVE-IT framework.

6.4.3 Prepare Environment

Establishing a consistent testing environment is essential for reliable digital artifact research. Virtualization, emulation and physical hardware are all viable options, each with its own considerations. Ensuring that the device model, operating system version, security patch level, and application version align with the original environment is crucial, as variations can significantly influence artifact behavior. Documenting all configurations and any discrepancies helps maintain the integrity and reproducibility of your research. [7, p.48-49]

Key Considerations of the Environment:

- Device hardware, operating system, application, and networking
- Credentials and user profile data used for the testing [9]

- Sanitizing the environment prior to testing
- Operational security (e.g. avoid exposing unique device and network identifiers)

6.5 Perform Baseline Acquisition

A baseline captures the state of the device before testing occurs and includes all the data necessary to support or refute the analysis being conducted. It establishes a reference point that allows researchers to observe changes that occur during and after the testing occurs [10]. When feasible, a baseline acquisition shall be created. Without a complete and well-documented baseline, it becomes harder to assess the integrity of findings, weakening their defensibility under peer review or reliability standards.

6.6 Test Data

Test data is the examiner generated inputs necessary to conduct the experiment. The Organization of Scientific Area Committees for Forensic Science (OSAC) document *Guidelines for Dataset Development* provides further information for creating test data [9].

6.6.1 Generate Test Data

Detailed documentation shall be created to record test data populated to the environment. The data generated should be within the scope and relevant to the proposed hypothesis. The detailed documentation of generated test data allows for the testing environment to be reproduced.

Care shall be taken that the dataset does not contain sensitive information.

6.6.2 Conduct Test Scenarios

Test scenarios are structured situations where the examiner interacts with the test data. The test scenarios shall be executed in accordance with the written test plan and within the defined toolset and environment, so results are reliable, repeatable, and defensible.

Each scenario shall be performed as a series of discrete, documented actions (e.g. inputs and interactions) while maintaining appropriate controls and minimizing opportunities for bias (e.g., designing and executing the methodology prior to viewing results).

Where possible, examiners should execute scenarios iteratively to isolate causality, varying one factor (or as few as possible) at a time, while preserving the accuracy of controls.

Any deviations from the planned procedure, shall be documented so limitations are transparent and the findings remain defensible.

6.6.3 Acquire Test Results

Acquiring test results should be treated as a planned, documented, repeatable component of the methodology rather than a single end-of-test activity. Examiners should conduct one or more iterative data collections, after any intentional or unintentional change to test parameters and for anomalous behavior. Differences observed across iterations can then be attributed to the test actions, not to changes in the collection process. These iterative acquisitions create a chronological record of the system under test, allowing examiners to capture when relevant artifacts first appear, how they change, and whether they persist, migrate, or are overwritten as

additional actions occur. Complex workflows and long-duration tests, where artifact creation and modification may occur asynchronously or in multiple iterations, are important situations to consider.

Any deviations or reasons for deviations from the planned acquisition method shall be documented so limitations are transparent and the findings remain defensible.

6.7 Data Analysis

Data analysis is the structured assessment of test outputs to normalize the resulting data. Once normalized, the output is compared against the other experiment results to determine what findings are within scope and can be used to draw the conclusion.

Examiners shall compare each iterative acquisition to the baseline and prior iterations to isolate the impact of specific actions and identify unanticipated behaviors. Special attention should be made to artifact changes that occur over time or in stages. It is important to understand potential limitations of automated forensic tools that may impact the analysis of novel artifacts. Examiners should distinguish between findings related to data provenance (e.g., source, processing history, or encoding lineage) and findings related to content interpretation. These represent different analytical domains and may require different approaches and levels of support. Manual parsing may be warranted. For additional guidance on best practices for data analysis, see *SWGDE 20-F-005-2.0 Best Practices for Mobile Device Forensic Analysis* and *SWGDE 18-F-001-2.0 Best Practices for Computer Forensic Examinations*.

6.8 Conclusion Phase

A conclusion is made based on the analysis of data produced by confirming or refuting a hypothesis. The examiner shall evaluate the limitations identified earlier and their potential impact on any conclusions reached in the research. These limitations may impact the potential applicability of the research.

The examiner shall be explicit in their conclusion as to their degree of confidence that their results can be extrapolated from the testing environment to the dissimilar environment.

The examiner should express their conclusion with a standardized level of confidence with supporting contextual information for the confidence of the findings. The examiner should not express their conclusion as a percentage or on any numeric scale.

In the conclusion phase, the examiner shall:

- assess whether their analysis results confirm or refute the hypothesis
- prepare and test an alternative hypothesis
- determine whether there is a need for further research
- assesses the sufficiency of the research goals

DRAFT

6.8.1 Propose Alternative Hypothesis

The scientific method is iterative. After concluding whether the results support the hypothesis, the examiner shall take into consideration viable alternative hypotheses. Testing alternative hypotheses helps avoid confirmation bias [11].

If the actual outcome of the research does not confirm the hypothesis, then the initial hypothesis shall be changed and an alternative hypothesis proposed. Depending on the discrepancy between the expected and actual outcome, the design of the test and resulting equipment setup and test scenarios shall be reconfigured.

6.8.2 Determine the Need for Additional Research

Additional research may be required when initial testing yields inconclusive or unexpected results, or when findings do not sufficiently support the original hypothesis. Additional research may also be necessary when tool limitations, environmental differences, or replication challenges prevent the examiner from reaching a defensible conclusion. If no results are produced, or if results fail to align with either the hypothesis or an alternative hypothesis, further testing should be conducted to refine or reframe the research question. Ultimately, the decision to pursue further testing should be based on whether existing findings provide adequate clarity, reliability, and reproducibility to withstand peer review and legal scrutiny. If further research cannot be conducted, the examiner should document as to why they are unable to conduct that research.

6.9 Peer and Administrative Review

The examiner shall submit their research for peer review. The purpose of peer review, often referred to as a technical review, is critiquing the entire research process to ensure repeatability and reliability. Peer reviewers shall have competence and practical experience in digital forensics relevant to the artifact or technique. See *SWGDE 12-F-006-2.0 Core Competencies for Digital Forensics*.

Administrative review is conducted for quality control purposes rather than technical evaluation. The administrative reviewer may be a non-technical individual. Depending on circumstances, the same individual may be completing peer review and administrative review.

The examiner will review and analyze any feedback received during technical and administrative reviews to determine if the research requires further iterative testing and alternate hypotheses. This ensures the interpretation of results is accurate and supported by established forensic principles.

6.10 Report and/or Publish Results

Research findings should be publishable only when they represent the outcome of a testable hypothesis evaluated against a defined dataset using a documented, repeatable method. Research that contributes to the body of knowledge should be considered and is encouraged for publication in scholarly journals to allow for appropriate review and validation. Examiners should consider making their dataset available for release should the community wish to peer-review their work.

This encouragement of publishing peer-reviewed research is not intended to discourage the blogging, publishing, or presenting of current work or work-in-progress.

6.10.1 Considerations for Online Content and Blog Posts

Examiners frequently share techniques they have developed through blog posts, aiming to disseminate methods that may assist the community in establishing a process for acquiring or interpreting artifacts pertinent to specific investigations.

Non-scholarly research, including content from blogs, forums, and videos, can inform the development of research topics. These sources can provide valuable initial insights and spark ideas for further investigation. However, such sources often lack formal peer review, posing risks of being misleading, incomplete, or inaccurate. Therefore, examiners must rely on their expertise and may need to conduct research to assess the material's veracity. This process serves as an informal peer review, providing constructive feedback to the original author.

6.10.2 Considerations for Digital Forensics Tools

The digital forensic community includes developers who release tools to be used by examiners to analyze artifacts in electronic evidence. These tools may range from scripts that parse a specific artifact to large digital forensic platforms that parse categories of artifacts from an entire drive or device.

Digital forensic tool developers may not be able to release proprietary information about their tools. Tools shall be tested in accordance with *SWGDE 18-Q-001-2.1 Minimum Requirements for Testing Tools Used in Digital and Multimedia Forensics* [12]. Digital forensic tool developers are encouraged to make information about the specific versions of the forensic artifacts they have tested available to their users, e.g., a forensic provider has tested version 4.5 of their forensic tool and has validated that it can associate image attachments with the corresponding messages for versions 7.9.1 through 12.3.2 of the Telegram iOS app.

However, in practice, the level of transparency provided by tool developers may vary. This expectation is a recommended practice, as developers may not always be able, or have an obligation, to disclose the full scope or methodology of their testing.

With this documentation, an examiner can assess whether the tool has been tested in an analogous environment. If not, the examiner shall conduct further research.

An examiner shall not rely upon any research where they do not know if the results are analogous and can be extrapolated to the artifact they are investigating.

7. Peer Review Process

Figure 2 is a visual flow chart for a suggested workflow, based upon DFIR Review organization’s process for performing a peer review [13].

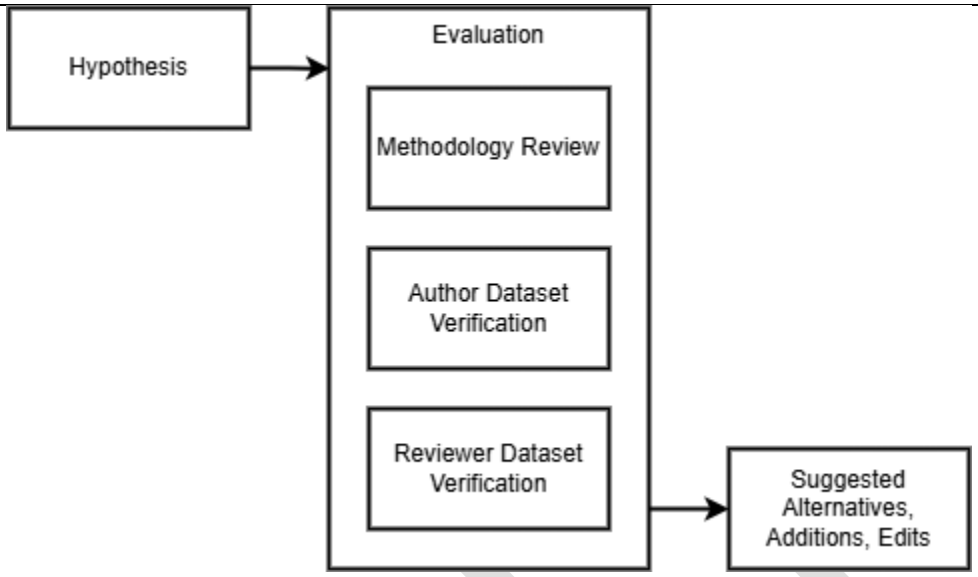


Figure 2. Suggested workflow for peer review (Figure credit: SWGDE)

7.1 Hypothesis

The researcher's hypothesis needs to be clearly defined and be testable, that is, capable of being confirmed or refuted through experimentation [5]. The reviewer should determine if the researcher's hypothesis has been clearly defined and the answer to the question has been thoroughly tested through experiments or observations [14]. Consideration should be given to other potential explanations that may result in the same digital forensic artifacts [7].

7.2 Evaluation

There are multiple methodologies for the evaluation of digital forensic artifact research. Some approaches, like the seven-level peer review hierarchy proposed by Horsman and Sunde, focus on peer review of active digital forensic investigations [15]. Another approach for digital forensic research is the DFIR Review framework.

This section follows the DFIR Review framework. Its methods include Methodology Review, Author Dataset Verification, and Reviewer Dataset Validation [13]. One or more of these methods may be used in the Evaluation phase as appropriate.

7.2.1 Methodology Review

The reviewer shall examine the methodology and verify that the concepts and methods used in the research are sound [15]. This is the minimum acceptable level of review. More extensive peer review, such as data verification, should be employed in addition to methodology review when possible.

7.2.2 Author Dataset Verification

If performing an Author Data Verification, the reviewer shall verify the technical details of the research using the author's dataset. The reviewer shall confirm whether, through independent examination of the dataset, they reach the same conclusion as the author [13].

7.2.3 Reviewer Dataset Verification

In Reviewer Dataset Verification, the reviewer also verifies the technical details of the research, but not with a dataset provided by the research's author. Instead, the reviewer uses a dataset that they generated. The reviewer should document the hardware, operating system, and app versions used in the dataset [13]. If performing Reviewer Dataset Verification, the reviewer shall verify the technical details of the research using the reviewer's dataset.

7.3 Provide Feedback to Author

The reviewer shall provide the researcher with feedback based upon their review: suggesting edits to the author's write-up, and identifying alternative methods or additional testing needed to establish whether the author's conclusion is valid [13].

8. Sample Templates

DRAFT

Dataset Development Documentation Template

This resource is available for download from Dataset Development Documentation Template.docx[10]

Intended usage(s) for dataset:

☐ Training exercise ☐ Competency/Proficiency/Certification test

☐ Capture the Flag exercise/test ☐ Tool testing

☐ Verification of artifact/finding ☐ Product/Tool development

☐ Other:

Development method:

☐ Hardware-Based/Physical Device ☐ Virtualization|Emulation

Make/Model: Virtualization|emulation environment:

OS/Version: Version/Configuration:

Device Identifier/Serial Number: Applications tested:

Phone Number (if applicable):

ICCID1:

ICCID2:

Applications tested:

Device date/time settings: Choose an item.

Date and time format used (include time zone): *YYYY/MM/DD HH:MM (24-hour UTC) Note that date formats are often confusing since there are many commonly used variations. The key point is to be clear what format is being used.*

User Account Information:

Platform/Service: Click or tap here to enter text.

User Account/Username: Click or tap here to enter text.

Password: Click or tap here to enter text.

Account Creation Date: Click or tap to enter a date.

Account Creation Time: Choose an item.

Repeat section as needed based on the number of configured user accounts.

DRAFT

Application Information:

Application: Click or tap here to enter text.

Version: Click or tap here to enter text.

Installation Date: Click or tap to enter a date.

Installation Time: Click or tap here to enter text.

Username: Click or tap here to enter text.

Password: Click or tap here to enter text.

Permissions:

Default:

Requested by app:

Granted by user:

Notes:

Repeat section as needed based on the number of installed applications.

Operating System (Device/Virtualized|Emulated System):

Name:

Version:

Installation Date:

Installation Time:

Username:

Passcode/Password:

Administrator/Root Privilege? ☐ Yes ☐ No

Passcode/Password:

Administrator/Root Privilege? ☐ Yes ☐ No

Operating System (Host, if using virtualization|emulation):

Name:

Version:

Installation Date:

Installation Time:

Username:

DRAFT

Test Activities:

Test Procedure:

Include brief paragraph or list outlining test prep, setup, activities, and extraction info.

Detailed log: *Add table rows as needed*

Date	Time	Application	Action	Content/Details/Location

Acquisition Information: (if applicable)

Acquisition Tool: *Include tool name and version number*

Acquisition Date: Click or tap to enter a date.

Acquisition Time:

Acquisition Method/Type: Choose an item.

Acquisition File Name(s):

Acquisition Notes:

Hash Value(s): *List hash type and value for each extraction file*

Filename: Click or tap here to enter text.

Hash Type: Choose an item. Hash Value: Click or tap here to enter text.

Known Limitations: *Include brief paragraph or list outlining limitations re: dataset creation or usage. Might want to include some typical examples – like variability among app/OS versions, underlying hardware, etc. How was the test data populated*

DRAFT

Test Results for Mobile Device Acquisition Tool [9]

Tool Tested: {Tool Name}

Software Version: {Version#}

Supplier: {Supplier}

WWW: {Web address}

Results Summary

{Tool Name} {Version#} was tested for its ability to extract and report data from supported cloud-based applications.

Except for the following anomalies, the tool acquired and reported all supported cloud-based application data.

Note that tools tested are reporting what is contained within cloud-based applications. Cloud-based applications often modify data (e.g., compressing the file, changing the file name) which results in an inconsistent file names, file sizes and/or hashes compared to the original file uploaded by a user.

(Remove this insert before publishing your report.

Note: Bullet items are only needed when more than one item/statement is going to be listed. Otherwise, indentation without the bullet should be used (Section 508 compliance guidance.))

***Storage service data*:**

{Enter Test Result findings}

***Email service data*:**

{Enter Test Result findings}

Productivity data:

{Enter Test Result findings}

Social Media and Messaging data (Facebook):

{Enter Test Result findings}

Social Media and Messaging data (X):

{Enter Test Result findings}

Social Media and Messaging data (WhatsApp):

{Enter Test Result findings}

Social Media and Messaging data (Instagram):

{Enter Test Result findings}

NOTE: Some social media applications will compress files as they are uploaded, resulting in inconsistent file size, file names and hash values compared to the original uploaded data files, resulting in different file sizes and hashes. This is reported "as expected" behavior and highlighted with an asterisk.

Test Results

This section provides the test cases results reported by the tool. Section 3.1 identifies the cloud-based service and data artifacts within each service used for testing {Tool Name} {Version#} .

The *Test Cases* column in sections 3.1 are comprised of two sub-columns that define a particular test category and individual sub-categories of cloud services that are verified when testing. The results are as follows:

As Expected: the CDX tool returned expected test results.

Partial: the CDX tool returned some of the data.

Not As Expected: the CDX tool failed to return expected test results.

Not Applicable (NA): the CDX tool does not provide support.

NIST Computer Forensics Tool Testing Federated Testing templates available at - https://www.nist.gov/itl/ssd/software-quality-group/computer-forensics-tool-testing-program-cftt/federated-testing

DRAFT

9. References

[1] NIST Interagency Report, NIST IR 8589, *Validation in Forensic Science*, December 2025. Accessed at: https://doi.org/10.6028/NIST.IR.8589.

[2] United States Supreme Court. *Daubert v. Merrell Dow Pharmaceuticals, Inc*. 509 U.S. 579. 1993. https://supreme.justia.com/cases/federal/us/509/579/case.pdf.

[3] Kim, Chulyoung, and Chansik Yoon. "Adversarial Bias and Court-Appointed Experts in Litigation." SSRN, Aug. 2018, https://ssrn.com/abstract=3240472. Accessed 2 Nov. 2026.

[4] Rabin, Matthew, and Joel L. Schrag. "First Impressions Matter: A Model of Confirmatory Bias." The Quarterly Journal of Economics, vol. 114, no. 1, 1999, pp. 37-82. https://doi.org/10.1162/003355399555945.

[5] Carrier, Brian. *A Hypothesis-Based Approach to Digital Forensic Investigations*. CERIAS Tech Report 2006-06, May 2006. https://www.cerias.purdue.edu/assets/pdf/bibtex_archive/2006-06.pdf.

[6] Scientific Working Group on Digital Evidence. *Recommended Guidelines for Validation Testing*. SWGDE Accessed at: https://www.swgde.org/guidelines-validation-testing-v2-archive/.

[7] National Institute of Standards and Technology. *Digital Investigation Techniques: A NIST Scientific Foundation Review*. NIST IR 8354. NIST, Nov. 2022, https://doi.org/10.6028/NIST.IR.8354.

[8] Scientific Working Group on Digital Evidence. *Establishing Confidence in Digital and Multimedia Evidence Forensic Results by Error Mitigation Analysis*. SWGDE 12-Q-001-2.0. *SWGDE,* 2018, https://www.swgde.org/documents/published-complete-listing/12-q-001-swgde-establishing-confidence-in-digital-and-multimedia-evidence-forensic-results-by-error-mitigation-analysis/.

[9] Organization of Scientific Area Committees for Forensic Science. *Guidelines for Dataset Development*. *OSAC*, 2022, https://www.nist.gov/system/files/documents/2022/12/15/OSAC-DE-Guidelines%20for%20Dataset%20Development.pdf.

[10] Geus, Julian, et al. "Systematic Evaluation of Forensic Data Acquisition using Smartphone Local Backup." Proceedings of the Digital Forensics Research Conference (DFRWS US), 2023, https://dfrws.org/wp-content/uploads/2023/07/geus-backupevaluation.pdf.

[11] Dhami, Mandeep, et al. "Effects of Task Structure and Confirmation Bias in Alternative Hypotheses Evaluation." Cognitive Research: Principles and Implications, 2024. https://doi.org/10.1186/s41235-024-00560-y.

[12] Scientific Working Group on Digital Evidence. *Minimum Requirements for Tool Testing for Tools Used Digital and Multimedia Forensics*. SWGDE 18-Q-001-2.1. *SWGDE,* 2024, https://www.swgde.org/wp-content/uploads/2024/04/2024-03-07-SWGDE-Minimum-Requirements-for-Testing-Tools-Used-in-Digital-and-Multimedia-Forensics-18-Q-001-2.1.pdf.

[13] DFIR Review. "Review Guidance." *DFIR Review,* https://dfir.pubpub.org/review-guidance. Accessed 2 Nov. 2026.

[14] Horsman, Graeme. "Part 2: Quality Assurance Mechanisms for Digital Forensic Investigations: Knowledge Sharing and the Capsule of Digital Evidence (CODE)." Forensic Science International: Digital Investigation, vol. 2, 2020, article 100035. https://doi.org/10.1016/j.fsir.2019.100035.

[15] Horsman, Graeme, and Nina Sunde. "Part 1: The Need for Peer Review in Digital Forensics." Forensic Science International: Digital Investigation, vol. 35, 2020, article 301062. https://doi.org/10.1016/j.fsidi.2020.301062.

10. Additional Resources

- Bernstein, David Eliot. "Expert Witnesses, Adversarial Bias, and the (Partial) Failure of the Daubert Revolution." Iowa Law Review, vol. 93, no. 451, 2008, SSRN, https://ssrn.com/abstract=963461.
- Camilleri, Andew, et al. "A Risk-Based Approach to Cognitive Bias in Forensic Science." Science & Justice, vol. 59, no. 5, 2019, pp. 533-543.
- Death Penalty Information Center. "Bias in the Courtroom." Death Penalty Information Center, https://deathpenaltyinfo.org/policy-issues/biases-and-vulnerabilities/lgbtq-people/bias-in-the-courtroom. Accessed 15 Sept. 2025.
- *DFIR Review,* https://dfir.pubpub.org/. Accessed 2 Nov.2026.
- Dror, Itiel E., and Jeff Kukucka. "Linear Sequential Unmasking–Expanded (LSU-E): A General Approach for Improving Decision Making as Well as Minimizing Noise and Bias." Forensic Science International: Synergy, vol. 3, 2021, article 100161.
- Eastman, Nigel, and Keith Rix. "Bias In Expert Witness Practice: Sources, Routes, to Expression and How to Minimize It." BJPsych Advances, vol. 28, 2022, pp. 35-45. https://doi.org/10.1192/bja.2021.19.
- Hart, Ariel. "In Court, Digital Evidence Can Shine or Fizzle." Crime & Public Safety, 25 Mar. 2016, https://www.ajc.com/news/crime--law/court-digital-evidence-can-shine-fizzle/dsdyFH23L3IZZaoTrLYZAO/. Accessed 4 Feb. 2021.
- Horsman, Graeme, and James R. Lyle. "Dataset Construction Challenges for Digital Forensics." Forensic Science International: Digital Investigation, vol. 38, 2021, article 301264. Elsevier, https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=931168.
- Koppl, Roger. "How to Improve Forensic Science." European Journal of Law and Economics, vol. 20, 2005, pp. 255-286. https://doi.org/10.1007/s10657-005-4196-6.
- Neal, T. M. S. "Are Forensic Experts Already Biased before Adversarial Legal Parties Hire Them?" PLoS ONE, vol. 11, no. 4, 2016, e0154434. https://doi.org/10.1371/journal.pone.0154434.

- National Institute of Standards and Technologies. "OSAC Lexicon - Scientific Method." NIST, https://www.nist.gov/glossary-term/31596. Accessed 2 Nov. 2026.
- President's Council of Advisors on Science and Technology. Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods. Sept. 2016, https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf
- Scientific Working Group on Digital Evidence. *Core Competencies for Digital Forensics*. SWGDE 12-F-006-2.0. *SWGDE,* 2024, https://www.swgde.org/documents/published-complete-listing/12-f-006-core-competencies-for-digital-forensics/.
- Scientific Working Group on Digital Evidence. *Best Practices for Computer Forensic Examinations*. SWGDE 18-F-001-2.0. *SWGDE,* 2025, https://www.swgde.org/documents/published-complete-listing/18-f-001-2/.
- Scientific Working Group on Digital Evidence. *Best Practices for Mobile Device Forensic Analysis*. SWGDE 20-F-005-2.0. *SWGDE,* 2026, https://www.swgde.org/documents/current-documents/.
- Scientific Working Group on Digital Evidence. *Quality Management System for Digital and Multimedia Evidence Units.* SWGDE 25-Q-001-1.0. *SWGDE,* 2026. https://www.swgde.org/documents/current-documents/.
- The Centre for Evidence-Based Medicine. Catalogue of Bias, https://catalogofbias.org/biases/. Accessed 15 Sept. 2025.

11. History

Revision	Issue Date	History
1.0 DRAFT	9/15/2025	Created document.
1.0 DRAFT	1/15/2026	Technical and editorial updates.
1.0 DRAFT	5/14/2026	Technical and editorial updates.
1.0 DRAFT	9/3/2026	Technical and editorial updates. SWGDE voted to approve as a Draft for Public Comment.
1.0 DRAFT	10/29/2026	Formatted for release as a Draft for Public Comment.