SANS DFIR

DIGITAL FORENSICS & AI INVESTIGATIONS FRAMEWORK

AI-Assisted. Human-Led. Trusted Investigations.
Practical framework for governing the responsible use of AI in Digital Forensics.

A LIVING FRAMEWORK | **AI-ASSISTED. HUMAN-LED.** | **DIGITAL FORENSICS**

Created by: **Heather Barnhart**

DFIR Curriculum Lead and Head of Faculty, SANS Institute | Safe AI Integration | SWGDE Best Practices for Digital & Multimedia Evidence

The human examiner is always the authority | Disclaimer: This is a living document. Guidance reflects current understandingand is subject to change as AI technology, legal precedent, and industry practices evolve. Always confirm content is current before use.

CORE PRINCIPLES
Why This Framework Exists

SANS DFIR

Why This Framework Exists

AI adoption in DFIR is already happening, with or without formal policies. Most teams lack governance models, validation standards, and defensible practices for integrating AI into critical investigative workflows.

Validation with current tools is often insufficient. As AI assertion confidence increases, the problem compounds without clear guardrails.

Augmentation, Not Replacement

AI is a force multiplier. It accelerates work; it does not own conclusions. The human examiner remains the final authority at every phase.

The analyst must independently verify AI findings, make the final determination, and accept full responsibility for all assertions.

Force Multiplier

AI is a force multiplier, not a replacement for trained analysts or examiners. It accelerates work. It does not own conclusions.

Human Accountability

Human accountability must remain central to all investigative decisions. AI systems should not assume legal or professional responsibility.

Validation Required

AI outputs must be validated before operational or legal use. Every AI-assisted finding used in an investigation requires human review.

Designed to Evolve

This framework is designed to evolve alongside AI technology, legal precedent, and industry best practices.

HUMAN OVERSIGHT MODELS
Risk Categories

SANS DFIR

1 Human IN the Loop
Maximum Control

Humans approve AI actions before execution.

Required for high-risk, legally sensitive, or irreversible tasks.

Applies to:
Validation | Attribution

2 Human ON the Loop
Monitor and Intervene

Humans supervise AI-driven processes in real time. AI acts, but humans monitor and can intervene at any point.

Applies to:
Examination | Analysis | Triage

2 Human OVER the Loop
Governance

Humans maintain governance and policy oversight. AI operates within defined, approved boundaries.

Applies to:
Identification | Reporting

RISK CATEGORY

Reversibility Risk

Activities that may alter systems, evidence, or operational environments require increased human oversight. Actions may not be easily undone.

e.g. Evidence imaging, system changes, collection

RISK CATEGORY

Accountability Risk

Attribution, validation, testimony, legal conclusions, and final decision-making require direct human ownership. Responsibility cannot be delegated to AI.

e.g. Attribution, expert testimony, final conclusions

Critical Principle: Regardless of oversight model, digital forensic conclusions require direct human validation before they are acted upon or entered into any legal or evidentiary record.

DIGITAL FORENSICS & AI FRAMEWORK
SWGDE Best Practice

SANS DFIR

	AI Recommended Low Risk	AI Optional Elevated Risk	AI Never Human Only	Human Driven
01	**Identification** Human On Loop	• AI scans environments to identify potential evidence sources • Automated device and artifact discovery across endpoints • AI-assisted scope definition and evidence mapping		**AI RECOMMENDED** FOR498 FOR508 FOR610
02	**Collection and Preservation** Human In Loop	• AI may assist with automated imaging workflows • Hash verification support can be automated • Chain of custody must remain human-controlled at all times		**AI OPTIONAL** FOR498, FOR500, FOR508, FOR509. FOR518, FOR528, FOR572, FOR585, FOR577, FOR589, FOR608, FOR610
03	**Triage** Human On Loop	• AI rapidly scores and prioritizes evidence by relevance • Automated IoC matching and threat scoring • AI-assisted keyword and pattern flagging at scale		**AI RECOMMENDED** FOR498, FOR500, FOR585
04	**Examination and Analysis** Human On Loop	• AI-assisted artifact parsing and timeline building • Pattern recognition across large evidence datasets • Human must validate all AI conclusions to mitigate inaccuracies		**AI RECOMMENDED** FOR498, FOR500, FOR528, FOR585, FOR589
05	**Validation** Human In Loop	• A human must validate AI findings at all times • Human examiner must independently verify all conclusions • Legal and evidentiary integrity requires human accountability		**AI NEVER** FOR500, FOR508, FOR518, FOR528, FOR572, FOR577, FOR585, FOR589
06	**Reporting and Presentation** Human Over Loop	• AI may assist with report drafting; human must author and sign off • Expert witness testimony is always human responsibility • All legal conclusions and actions require human authority		**HUMAN DRIVEN** FOR498, FOR500, FOR508, FOR572, FOR577

AI is a force multiplier | The human examiner is always the authority | Disclaimer: This is a living document subject to change as AI technology, legal precedent, and industry practices evolve. Content may become outdated. Always confirm guidance reflects current standards before use.

HIGH-RISK AREAS
Operational Use Cases

SANS DFIR

HIGH-RISK AND PROHIBITED

Some actions must never be delegated to AI, regardless of capability, confidence score, or efficiency gains.

No Independent Attribution

AI should not independently determine attribution. Attribution requires human expertise and legal accountability.

No Legal or Evidentiary Conclusions

AI should not make legal or evidentiary conclusions. These carry professional weight that cannot be delegated to automated systems.

Testimony and Reporting Ownership

Human experts must retain ownership of testimony and final reporting. Expert witness credibility cannot be transferred to an AI system.

Full Documentation Required

All AI-assisted actions must be documented. Transparent records of where and how AI was used are essential for defensible investigations.

OPERATIONAL USE CASES

Where AI can responsibly accelerate workflows, with human oversight at every step.

Timeline Reconstruction

AI-assisted timeline reconstruction and evidence triage accelerates identification of key events across large forensic datasets.

Triage Assistance

AI can help prioritize, score and flag evidence at scale requiring deeper investigation.

Potential Evidence Identification

AI can scan for additional sources of evidence and flag this for the examiner.

Report Drafting

AI-assisted drafting reduces analyst burden while preserving human authority over all final findings and language. Human must author and sign off.

Critical Boundary: AI CANNOT put a human behind an artifact to prove attribution.

SANS DFIR

ACCELERATE DFIR. RESPONSIBLY.

Start With Low-Risk Use Cases

Human Examiner Is Final Authority

Transparent Governance and Validation

AI can accelerate DFIR operations responsibly. The human examiner remains the final authority. Organizations should begin with low-risk use cases and build governance before expanding capability.

A LIVING FRAMEWORK FOR SAFE AI INTEGRATION IN DFIR