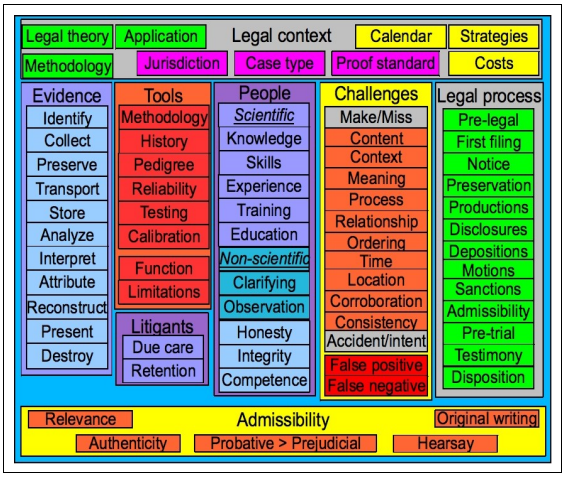
Digital Forensic Evidence Examination

The State of the Science

and Where to Go From Here

Fred Cohen – President - California Sciences Institute



Introduction:

We have some descriptions of digital forensic evidence (DFE) examination, which is where most efforts at creating a science begin. DFE is not operating as a normal science today. We don't yet have; (1) a community consensus around a body of knowledge, (2) well defined and consistently used terms, (3) a well understood epistemology, theory, and methodology, (4) a strong experimental basis, or (5) an agreed upon physics. We need to change this by creating a community consensus, defining and using terminology consistently, agreeing on an epistemology, theory, methodology, experimental basis, and physics. Unless and until we do, we will never become a true science and stunt the long-term advancement of our field.

Epistemology: Can we agree on these things?

Digital forensic evidence (DFE) is entirely sequences of bits. It has different physics than matter and energy. It has finite granularity in space and time. It can be observed without alteration. It can be duplicated without removal. It is trace but not transfer; the trace of the execution of finite state machines (FSMs). FSMs have specific properties that define the physics. Finite granularity → limits on accuracy and precision. FSMs are syntactic in nature so semantics are driven by context. DFE is latent in nature, so we must use tools to see and examine it. DFE can never speak directly to events in the physical world except in limiting what FSMs could have done. Computational complexity is a sort of speed of light. At the edge between physical and digital spaces there are assumptions.

Theory: Can we agree on these things?

Scientific theories are not casual theories, and casual theories should not be treated as part of the scientific basis for our field. Refutation can destroy any

theory, but confirmation cannot prove it, except for finite cases. Our theoretical basis should change slowly once accepted, and only do so because of dramatic changes in underlying epistemology or the nature of our understanding at fundamental levels. Accepted theories should remain useful for most cases as we come to new understandings. → We need to be careful at this level. Most theories today will stem from computer engineering, computer science, mathematics, and related fields. Such theories will limit what can be stated scientifically about DFE.

Methodology: Can we agree on these things?

We need a standard model, and one is proposed: Laws, regulations, and violations. (L, R, V); Events (E) are non-digital things like statements and declared facts. Traces (T) are sequences of bits. DFE examination identifies consistencies and inconsistencies between traces (TxT) and events (TxE). A finite set of known procedures (P) are used to do examination. There are imposed limits on resources (R) and schedule (S). This model implies sizes of P, T, (TxT), (TxE) → limits of P(R), and limits on thoroughness. The fundamental theory of DFE examination is: “What is inconsistent is not true”.

Experimental basis: Can we agree on these things?

The |corpus| of peer reviewed scientific DFE papers < 500. Most are limited applicability and not focused on fundamental understanding. Few experimental results are available and fewer meet standards of scientific rigor from other fields. Most experiments are confirmatory and not refutational. DFE is latent → we need to understand tools, their limits, and functions. We don't have a methodology to evaluate tools; (1) no error model, (2) no calibration approach, (3) no testing and verification process, and (4) no theory of measurement.

Physics: Can we agree on these things?

Only limited results, not widely accepted. Digital space converges with time: FSM: (I,O,S,m:{IxS}→{O,S'}) IF |I|>(|O|+|S|) THEN ∃(i,i')∈I:∃(o)∈O,∃(s)∈S, i→(o,s) and I'→(o,s). Time is a partial ordering. Traces are subject to Δt not now known. Time has a directional asymmetry: Given {IxS}, {O,S'} is unique and known. Given {O,S'}, {IxS} is non-unique. Precision is always finite. Some elements of meat-space physics still apply (e.g., *c* provides lower bounds). I/O with meat-space is not repeatable (e.g., print a Jpeg and rescan it. Scan a picture again and again).

What can we agree on?

[1] F. Cohen, “Digital Forensic Evidence Examination - 2nd edition”, 2010, ASP Press. ISBN 1-878109-45-6.