The Future of Digital Forensics

2012-09-21

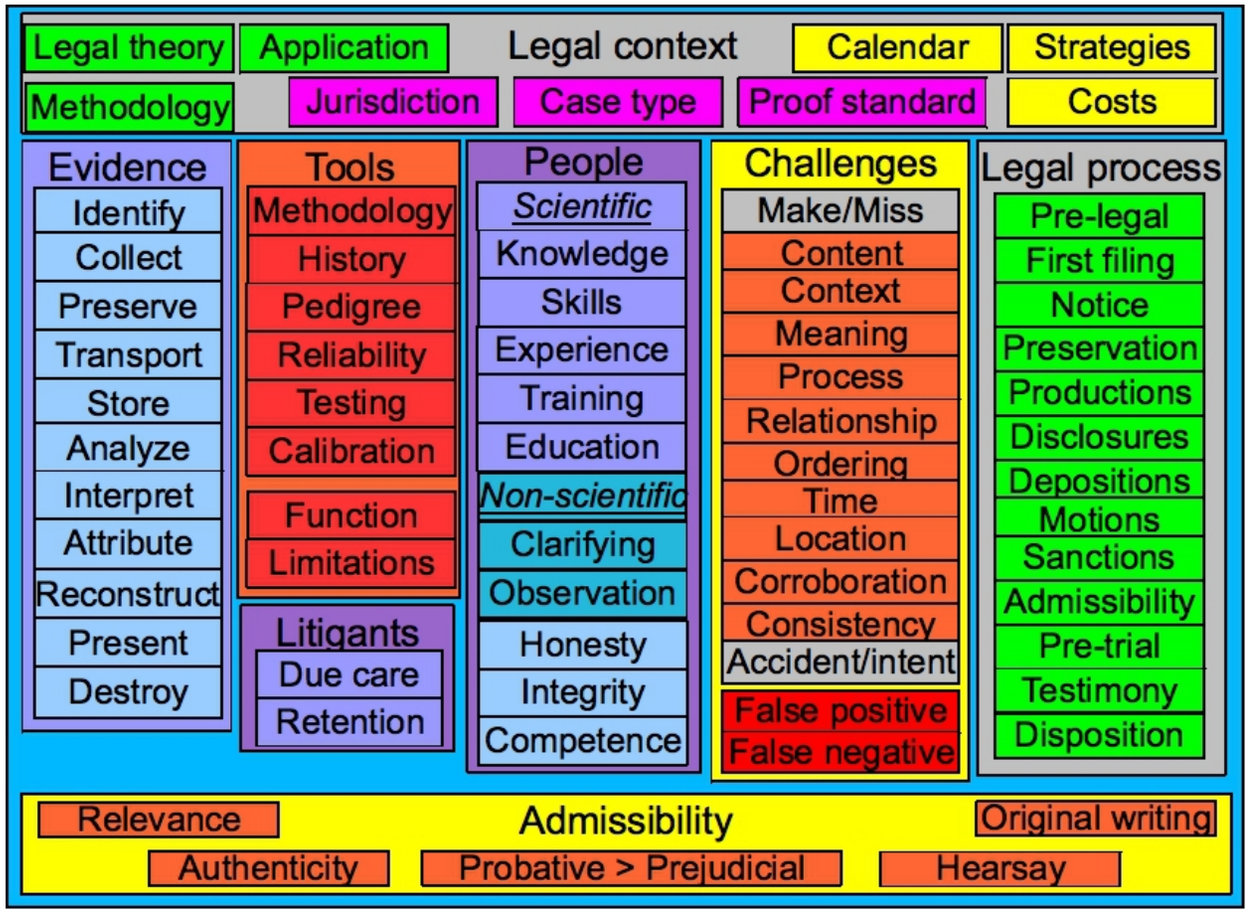
1st Chinese Conference on Digital Forensics

Keynote Address

Dr. Fred Cohen
President - California Sciences Institute
CEO – Fred Cohen & Associates

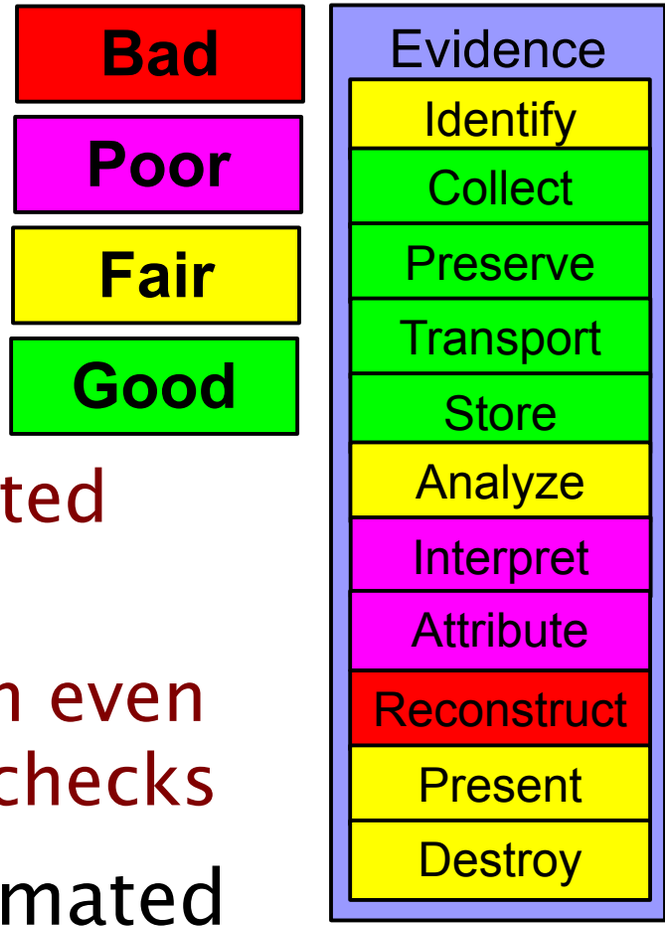
Outline

- The state of the art and what we are missing
- Where we want to go and how we will get there
- Things to look out for
- Questions / Comments / Discussion?



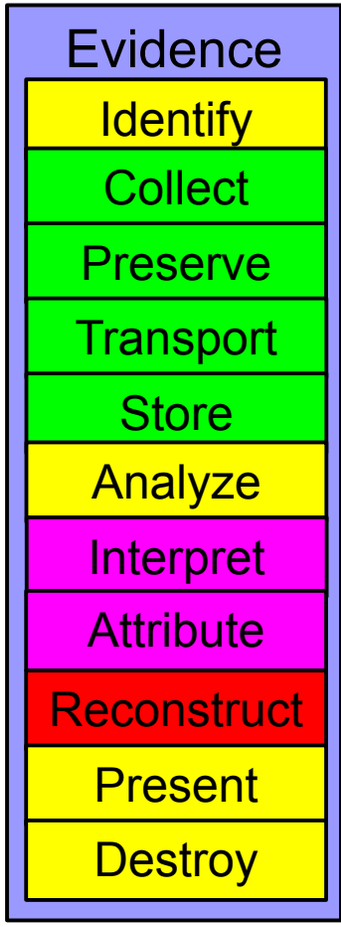
We're not as dumb as they think we are

- We can do some basic things very well
- Identification of evidence
  - Problematic in many cases
  - Obvious direct things readily identified in many cases, but redundant and related evidence is often overlooked
  - Courts often limit redundant information even though it is vital to evidence integrity checks
- Collection is well-defined and largely automated
  - Imaging media is excellent or nearly that
    - Automated tools make it very easy
  - Imaging over networks is not as good, but...
    - It is readily accepted if experts do it right



We're not as dumb as they think we are

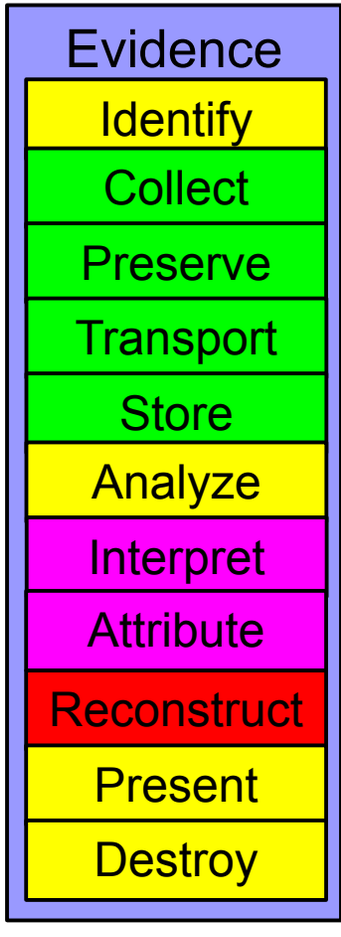
- Preservation is usually done very well
- Transport is fast, efficient and rarely a problem
- Storage can be done well and with integrity
- Analysis is complicated and poorly understood
  - Much of the research is in this area
  - Many algorithms exist but bases are limited
  - Example: trace typing is problematic
    - Embedded content, Steganography, etc.
  - Example: simple searches miss many results
    - Hypen-ated and missspellled words
    - Words spread across non-continuous blocks
    - Different character sets and representations



YES ≠ YES

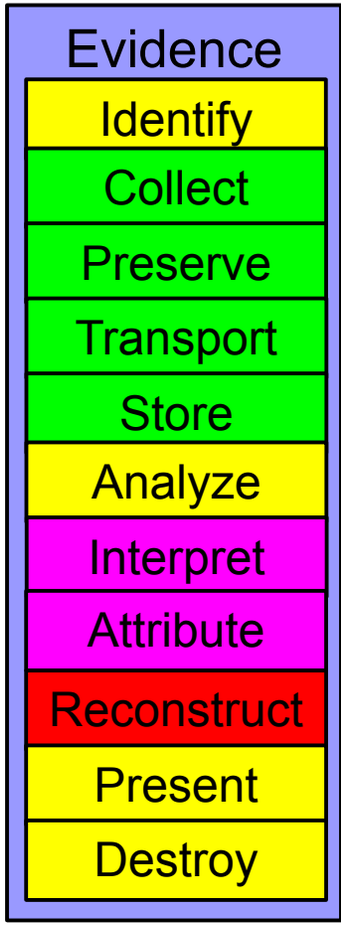
We're not as smart as we think we are

- Interpretation is often just plain wrong
  - Example: Encase interpretation of timestamps
    - Claim of wrong time and zone
    - Ignores multiple time stamps choosing one over the others
    - The user not shown underlying data and interpretation method to challenge it
  - Example: "Expert" testimony fails smell test
    - Claims implied that information exceeds the speed of light
  - Example: Claims of easy forgery
    - But the "Expert" forgery demonstration was easily detected as forged
- Attribution is usually good at level 1 (direct)



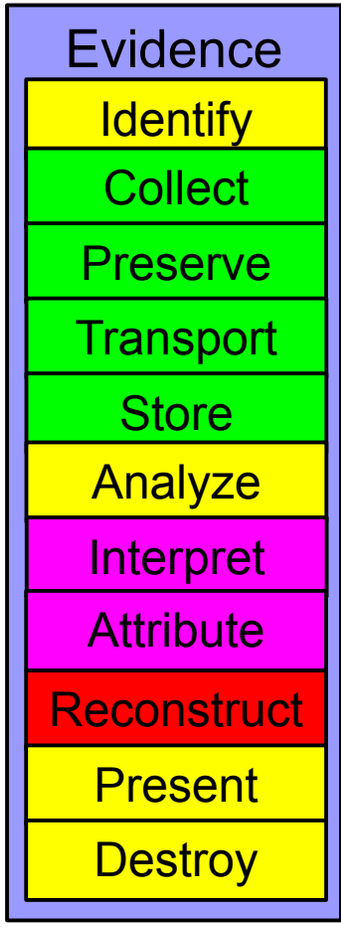
We're not as smart as we think we are

- Attribution is usually good at level 1 (direct)
  - At levels 2, 3, and 4, it is only good by nation states who rarely allow LE to use it
    - Level 1 – the direct IP address that did it
    - Level 2 – the indirect source (where the person who did it was)
    - Level 3 – the person who did it
    - Level 4 – the organization behind the person
  - Example: Anonymizers, Onion routing, etc.
    - Diffuse and confuse traceback processes
  - Example: Who launched the XXX worm?
    - Recursive infection leads to difficulty in identifying victim 0 and sourcing from there



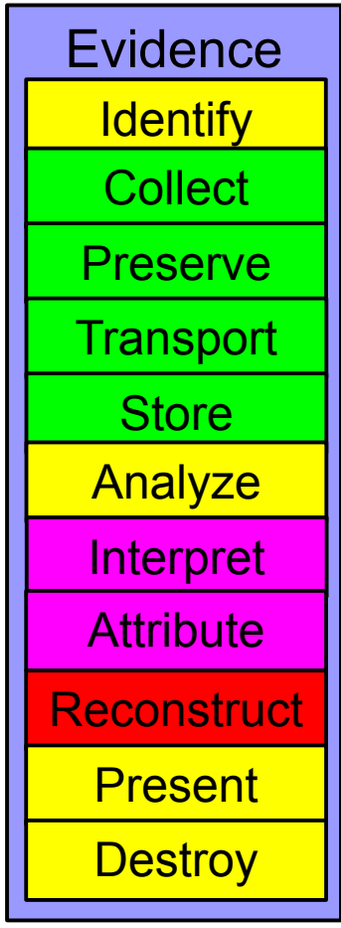
We're not as dumb as they think we are

- Attribution is only good at level 1?
  - Counterexample: Polgar v. US Chess et. al.
    - Automated analysis tool used to pick 33 postings out of 1M+ and uniquely associate them to a posting from a separate site
    - Anonymizer used to steal attorney client privileged emails – perpetrator detected by financial, timing, address, and related records
  - There are few cases and experts where this level of effort is undertaken and expertise available, but it can often be done
  - The challenge is getting more in LE who understand and more automated methods



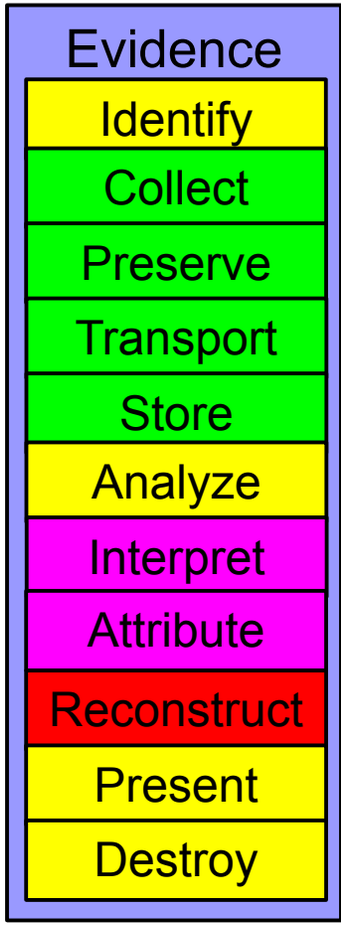
We're not as dumb as they think we are

- Reconstruction is poorly understood and not adequately studied
  - It is expensive and hard to do well
  - It takes time because of inadequate underlying support
  - Counterexample from 2001:
    - AZ v. Miller – manual reconstruction showed claims of prosecution unsupported
  - More recent automation of reconstruction:
    - (unpublished) automated reconstruction for attribution testing produces demonstrable claims in situ for causal chains
    - Automated particularization to mechanism
    - Automated individualization via related records



We're not as dumb as they think we are

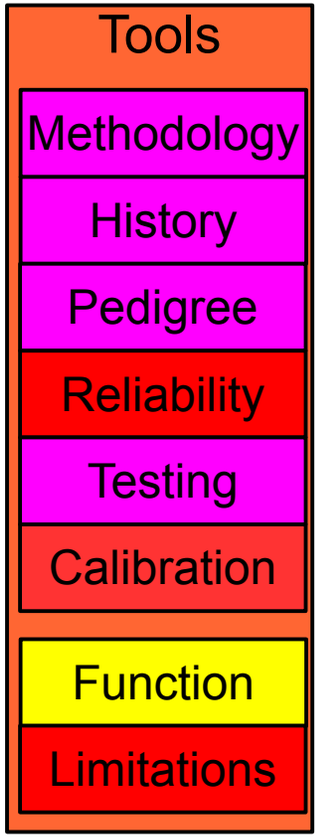
- Presentation problematic (for many)
  - Even the simplest things are poorly presented
  - But methods like Forensic Fonts resolve these
- Destruction appears settled
  - We know what can and cannot be recovered
    - Keyboard and laboratory attack example from NIST 2006 review of HD wiping by overwriting
    - Sandia studies of rapid destruction circa 2000
  - We just don't do it very well
    - Example: Chinese booth at RSA show demonstrated destruction technology that could readily be recovered in a laboratory
    - Suitable for commercial, but not against a government or for very high valued content



P	K	^c	^d	^t
50g	4Bg	03g	04g	14g
8	S	6	∅	∅
38g	53g	36g	00g	00g

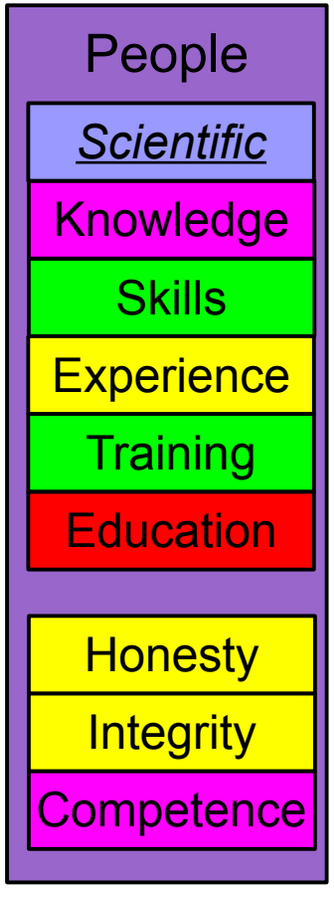
The tools are more questionable

- We lack a real scientific basis for most tools
  - The methodology we currently use limits what we can do and understand regarding
    - Analysis, Interpretation, Attribution, Reconstruction
  - Commercial tools typically lack the sorts of pedigree and quality of high grade SW
    - Trusted systems and trusted development, proof of correctness, well understood algorithms, well tested software, reliable operations, output tied to source data, etc.
    - Methodologies for calibration are just emerging, and in many cases, we don't even have well defined functions and specified limitations – in commercial products.



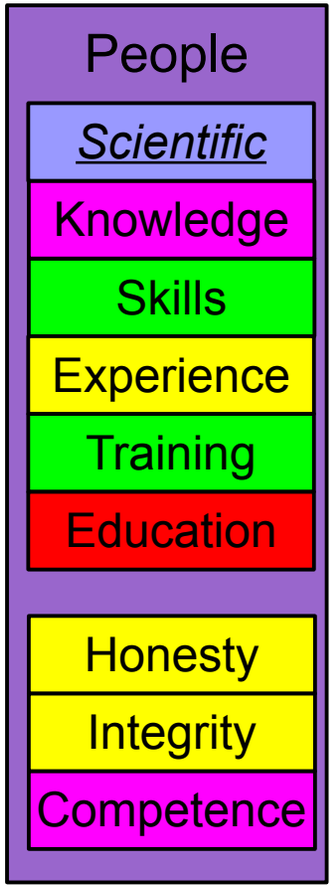
We're not as smart as we think we are

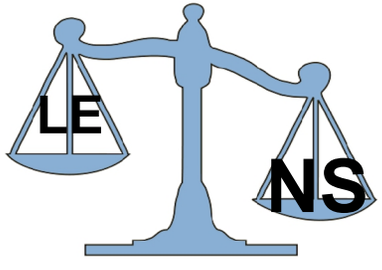
- Almost all LE I have encountered in this field are honest and have high integrity
  - Only a very few cases have been identified where this is not true
  - The “real” experts are honest and have high integrity in all areas and cases I have seen
  - Experts for hire exist (at least in the US) and are often low quality and will say anything for enough money.
  - The courts often identify them as not credible
- Competence is a different issue
  - It stems from knowledge, skills, experience, training, and education



We're not as smart as we think we are

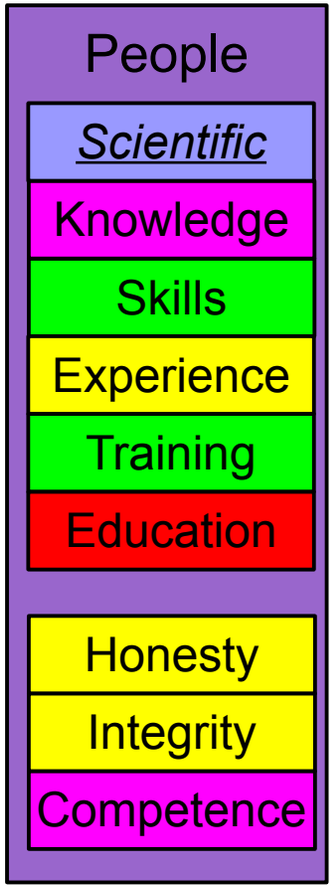
- At the national security level
  - Many very smart and dedicated people
  - Extremes in knowledge and skills
  - Outstanding experience over many years
  - Very strong training and operations
  - Limited education/ more tradecraft than science
- The legal system – LE is given short shrift
  - Much of the national security tradecraft is not available to LE because of secrecy
  - Revealing these secrets may reveal too much information and endanger national security
  - It is called the "equities" problem





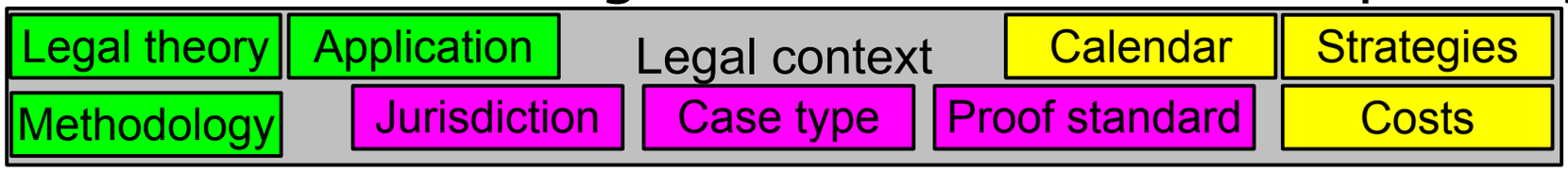
We're not as smart as we think we are

- Knowledge is limited by lack of funding for science – related to equities – but changing?
  - There is more push for science and engineering today, but still a lack of funding.
  - Example: "The physics of digital information" http:infophys.com is a chapter on this subject
- Skills through practice and training are good in areas with well-developed technology
- Training on commercial technology is good
- Experience is limited because of rapid IT changes (e.g., cloud computing, mobility, pad technology)
- Education is severely lacking (1st US Ph.D. Program in digital forensics at CalSci)



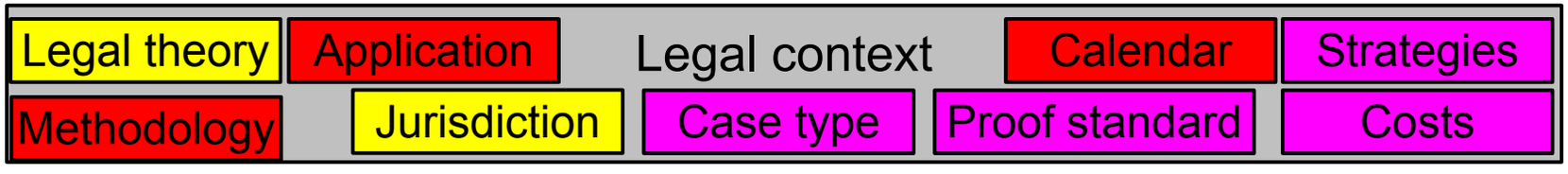
Management science is lacking

- Case management is in its very early stages
  - Hong Kong Police Force is the best example of starting in this line
    - HKPF + University collaboration in modeling cases
    - Bayesian model of evidence needed to support cases
    - Adding in prosecution history to identify needed evidence
    - Identification of what evidence to seek first / at all
    - Earlier cut-off of bad cases, clearing the innocent
    - Earlier determination of good cases (likely guilty)
    - Earlier cut-off when enough evidence is found
  - It's not just about cost
    - It's about the rights of the innocent and public good



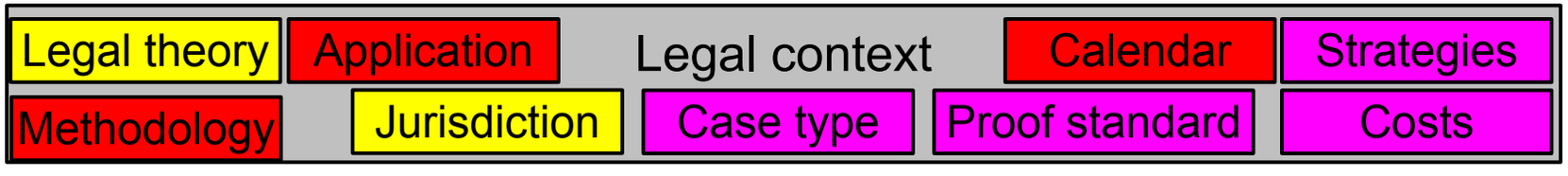
Management science is lacking

- Legal context drives the management of cases
  - Legal theories are rarely codified into frameworks
    - Counter-example: charting of laws by LE for making cases
  - Theoretical work largely ignores case experience
    - Those working on cases rarely do theory and vice versa
  - Lack of scientific methodologies to base work on
  - You cannot properly apply methodology you don't have
- Lack of models for optimization
  - Lack of automation to support processes and calendar
    - Except for Hong Kong, no real effort in strategy and cost



Management science is lacking

- Jurisdiction x Case type x Standards of proof
  - Jurisdictional issues are obviously well understood
  - A lot of cooperation exists on a global basis for hunting
  - But largely limited by political will - support for some activities and not for others by nation states
  - This requires global treaties – that will take a while
  - Proof standard is very tricky as there is a lot of judgement involved – historical data is needed – again HKPF leads
  - Limited attempts at case management help to keep track of cases and the large amounts of related information
    - But these technologies are information support only



Outline

- The state of the art and what we are missing
- Where we want to go and how we will get there
- Things to look out for
- Questions / Comments / Discussion?

A future vision

- What we all want (police and most people the World over)
  - Only criminals get arrested and prosecuted
  - Innocents are rarely investigated and never charged
  - Court cases are rock solid on a sound basis
  - Costs are low, danger to the public is low, officers are safe
  - Only the necessary information is gathered, processed, and used – and only within the legal strictures of jurisdictions
  - The evidence is overwhelming – and the courts agree
  - Corruption is almost non-existent and easily detected
  - The public has high confidence in the integrity and fairness of the system and it is transparently so
  - Detection, arrest, prosecution, and punishment are sure, fast, and just, and everyone knows it

Achieving the future vision

- Only criminals get arrested and prosecuted
  - Key factors include:
    - Reliability* of evidence (records)
    - Attribution of actions to actors (at 4 all levels)
  - I maintain that this is adequate in an otherwise just system
- Innocents are rarely investigated and never charged
  - This is an optimization issue closely related to HKPF efforts previously cited
  - If we understand what evidence makes a case, we may rapidly identify, collect, and analyze only the necessary evidence to rule out suspects as soon as possible
  - Near-real-time examination of relevant records for classes of acts is already feasible in some cases

Reliability: from Diplomatics – The record reflects the reality it purports

Achieving the future vision

- Court cases are rock solid on a sound basis
  - The science is done to provide the basis for everything produced regarding the forensic evidence
  - The basis is provided, well studied, and presented clearly
- Only the necessary information is gathered, processed, and used – and only within the legal strictures of jurisdictions
  - See the optimization issues discussed above
- The evidence is overwhelming – and the courts agree
  - This requires systematic collection, retention, analysis, and review, and tracking legal decisions for what is and is not acceptable
  - Similar to the efforts of the HKPF but on a far greater scale

Achieving the future vision

- Costs are low, danger to the public is low, officers are safe
  - Automation on a large scale reduces cost in the long run
  - Optimizations discussed earlier reduces costs
  - For some classes of crimes, automated ongoing analysis may rapidly detect possible crimes, identify and collect just the right evidence, preserve it, do analysis, attribute acts to actors, and present the case for human review
  - Some such mechanisms already exist for some classes of activities in some environments, but at a larger scale, the challenges are far greater and there is a long way to go
  - Example: In well controlled computing environments with low base rates of activities of import, indicators are rapidly detected, automatically analyzed, and presented to investigators in a form suitable to legal use.

Digital Officer Safety

- Operational security is key to effective digital officer safety
  - This is a well established area with lots of sound practice
  - But it is largely tradecraft and not widely practiced
  - Federal Law Enforcement Training Center (FLETC) in the US and Kevin Manson in particular, did an outstanding job of building this into the training program
- At a basic level, this requires that risks to the officers be considered and traded off against other equities
  - IF criminals can do attribution as well as law enforcement
  - THEN criminals can find the officers as well as officers can find the criminals
  - This makes safe undercover work infeasible?
    - A differential in capabilities or their use is necessary

Achieving the future vision

- Corruption is almost non-existent and easily detected
  - Redundant records provide the means of detecting falsifications of records today
  - The technology for automating this at large scale is now being tested and deployed in select environments
  - This is the future basis for questioned documents examination that will roll into the legal system as standard practice over time
  - The leverage is shifting to where detecting subversion is far easier and more reliable than forgeries or deceptions
- The days of easy / feasible forgeries are over
  - In case after case – where adequate expertise is applied – forgeries are detected and responsible parties punished
  - The myth still persists – but the reality does not

Achieving the future vision

- The public has high confidence in the integrity and fairness of the system and it is transparently so
  - As these changes get integrated into the legal system, public confidence will grow – but only if we do it well
  - As much as possible, transparency must be achieved
  - The press and public must know the reality of what works, how well, and why
  - The global academic community must be integrated into the public process and support the validity of the results
  - Global scientific consensus surrounding the science and methods must be broad and clear
  - The global educational and research system must support ongoing open science to build the global capacity and the science and arts involved

Achieving the future vision

- Detection, arrest, prosecution, and punishment are sure, fast, and just, and everyone knows it
  - This will follow from all the rest
- But to achieve it, we need resources and support adequate to achieve it as a global goal
  - Nation states can only go so far on their own because of the global nature of crime and policing
  - If and as we fight against this, we will throw out the information age baby with the bathwater
  - Crime is global and globally organized – crime fighting must also be global or it will fail to meet the challenge

Don't forget civil cases

- Legitimate civil disputes require similar capabilities
  - Often more resources applied by all sides
  - Different standards of proof and discovery requirements
  - Few of the permanent infrastructures in place
  - Far less repeatability in cases – each is fairly unique
- Civil justice is just as important as criminal justice
  - Confidence in a political system demands justice at all levels – from the petty thief through the corporate raider
  - Advancements in science and engineering require protection of intellectual property or investment will collapse
  - Personal motives drive many innovators, and ego is as important as money – give credit where due - attribution

Outline

- The state of the art and what we are missing
- Where we want to go and how we will get there
- Things to look out for
- Questions / Comments / Discussion?

Opportunity Ahead

Big brother and going too far

- There has to be a balance between freedom and justice
  - Unlimited freedom leads to people unjustly taking advantage of others – a human rights issue we all face
  - But perfect justice at the cost of freedom guarantees that we will all be mediocre and society will stagnate
  - A balance must be struck and adapted over time
- People are not perfect – and neither are computers
  - Whatever we do will be imperfect because we are imperfect
  - But by striving to always improve, we will lift all boats
- Resiliency and transparency are the guarantors of justice
  - The system has to self-correct to support the will of the people without supporting the tyranny of the majority

The equities issue

- The balance between secrecy (attack) and justice (defense)
  - We all recognize that secrets are necessary in a hostile world (even if we wish it wasn't a hostile world)
  - We all recognize that justice is the solution to many of the problems of the world (both criminal and social justice)
  - But we may disagree about the balance
- My personal view
  - Attack has been given too much of the resource in the information arena – and as a result, defenses suffer
  - The net effect is that crime is easy and rampant
  - The net effect is that justice is hard and hard to come by
  - We need to balance the equities toward justice

Emerging times

- Flash mobs and instant forensics
  - Flash mobs faster than police – how do we catch up? (London protests)
  - How do we attribute after the fact? (London subway bombings)
  - How and when do we intervene in advance (BART in the Bay area)
  - Technology limits from end to end (cameras to analysis to response)
- Social media and the information currency
  - Wiki-Leaks and information "needs to be free" - the insider
  - Revolutions in the Middle East – where else can they go?
  - Steganography, cryptography, and Linux – the info-weapons
- Cybernetic society and critical infrastructures
  - How do we protect and serve against Stuxnet and cyber systems?
  - If the attacks directly hit our own systems, how do we continue?
  - Are police really going to investigate cybercime? How exactly?

Research / Education / LE linkage

- There are many things we know how to do well
  - Technically, there is a great deal we can do today
    - But there is still a lot to learn
  - A few people have a lot of know-how – most have little
    - We need to share what we know as we learn more
  - The rate of progress is currently limited by who is resourced to what level for what purpose
- Partnership is the best way forward I know of
  - University researchers x Real world examiners x Students
    - Advancement of the state of the art
    - Solutions to current challenges
    - Building the capacity for future workforce and research
  - A positive feedback system to grow justice and knowledge

Summary

- As a field, digital forensics is weak in many areas
  - But there are some areas of strength we can build on
- A lot of progress has been made lately
  - Much of it is not yet available in the open market and the science is not openly and adequately supported
  - There is too little resourcing today directed in the wrong ways to achieve what we all want to achieve
- A future vision is realizable in parts and over time
  - And may be achieved in pieces where it is prioritized
- The future of justice demands open global cooperation
  - “Sunlight is said to be the best of disinfectants" - Brandeis
  - “Injustice anywhere is a threat to justice everywhere.” - Martin Luther King, Jr.

References

- F. Cohen, “Digital Forensic Evidence Examination”, ASP Press, 2009-2012
- Susan Polgar v. US Chess Federation et. al. In the US District Court, Northern district of Texas, Lubbock Division, C.A. NO. 5-08CV0169-C, Sep 15, 2009.
- Rose v. Albritton, Superior Court of the City and County of San Francisco, Case No.: FDV-09-806677, July 14, 2009
- F. Cohen, “Update on the State of the Science of Digital Evidence Examination”, Conference on Digital Forensics, Security, and Law, May 29-31, 2012.
- F. Cohen, “Forensic Methods for Detecting Insider Turning Behaviors”, IEEE Workshop on Research for Insider Threat, 2012-05-25.
- M Kwan, K P Chow, F Law & P Lai, "Reasoning About Evidence Using Bayesian Networks", Advances in Digital Forensics IV, 2008, pp.141-155. and others
- US Patents 6,813,682; 7,159,086; 7,228,379 as examples.
- F. Cohen, "Fonts for Forensics", IEEE SADFE, 2010-05-19, Oakland, CA.
- “Guidelines for Media Sanitization”, NIST Special Publication 800-88, 2006-09
- Justice Lewis Brandeis, “What Publicity Can Do”, Harpers Magazine, 1913
- Martin Luther King, Jr. “Letter from Birmingham Jail”, April 16, 1963.
- L. Duranti, “Diplomatics”, Encyclopedia of Library and Information Sciences, Third Edition DOI: 10.1081/E-ELIS3-120043454, 2010, Taylor & Francis.

Outline

- The state of the art and what we are missing
- Where we want to go and how we will get there
- Things to look out for
- Questions / Comments / Discussion?

http://all.net/ - fc at all.net

Thank You