DIGITAL FORENSICS

Parts of this material has been compiled from various open sources

Goal of this lecture

- Give a basic insight in digital forensics?
  - Computer crime
  - Some terminology
  - General approaches to digital forensics
  - Importance of tools

- Recovering data from storage media
  - File formats: example FAT
  - SSD/Flash media

- Insights in methods of information hiding
  - Steganography

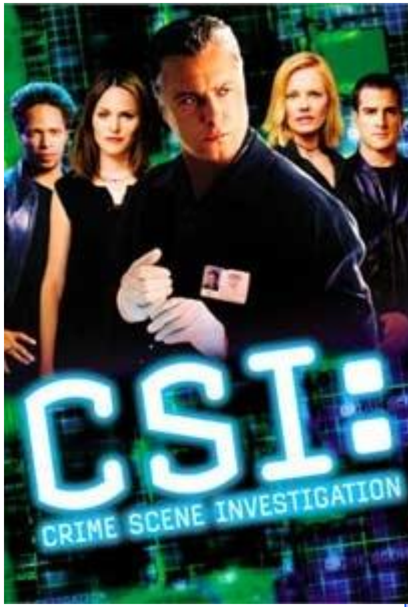
Mandatory reading

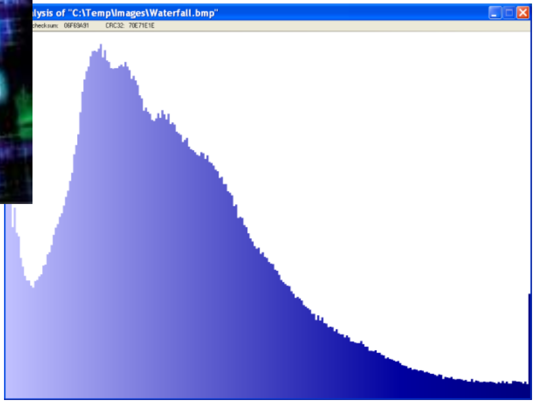
- Forensics of mobile phone internal memory: by Svein Y. Willassen. Norwegian University of Science and Technology
- A Hierarchical, Objectives-Based Framework for the Digital Investigations Process, Nicole Beebe, Jan Clark, DFRWS 2004
- Altheide Video The death of computer forensics

Overview of this lecture

- What is Digital Forensics?
- Terminology
- Incident handling
- Organizational Roles & Responsibilities
- Detection and Correlation Tips
- Tools
- Recovering data from storage media
- Steganography

What is digital forensics?





Origin/history

- Forensic investigations of computers is basically as old as computers started to be common in information and financial processing systems.

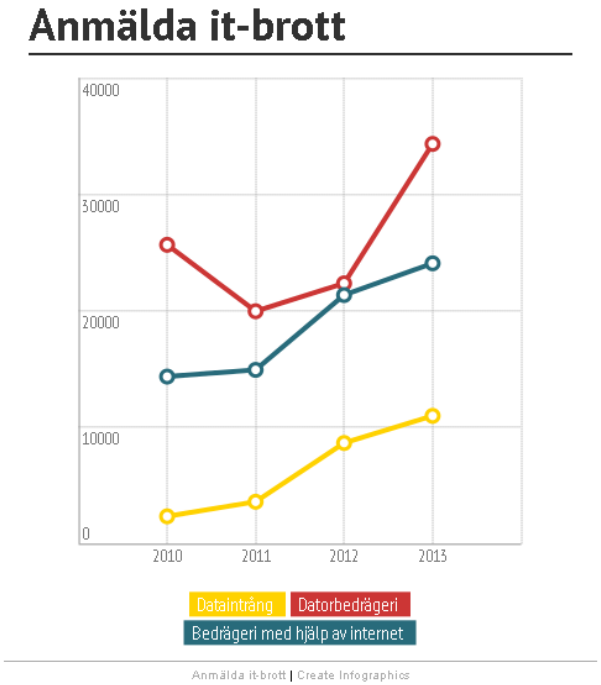
- Computer/data fraud and data forensics go hand in hand and the need for data forensics is increasing.

⇨ CYBER CRIME, CYBER WARFARE, etc

- In many countries we have today dedicated education for data forensic engineers.

Cyber crime – statistics

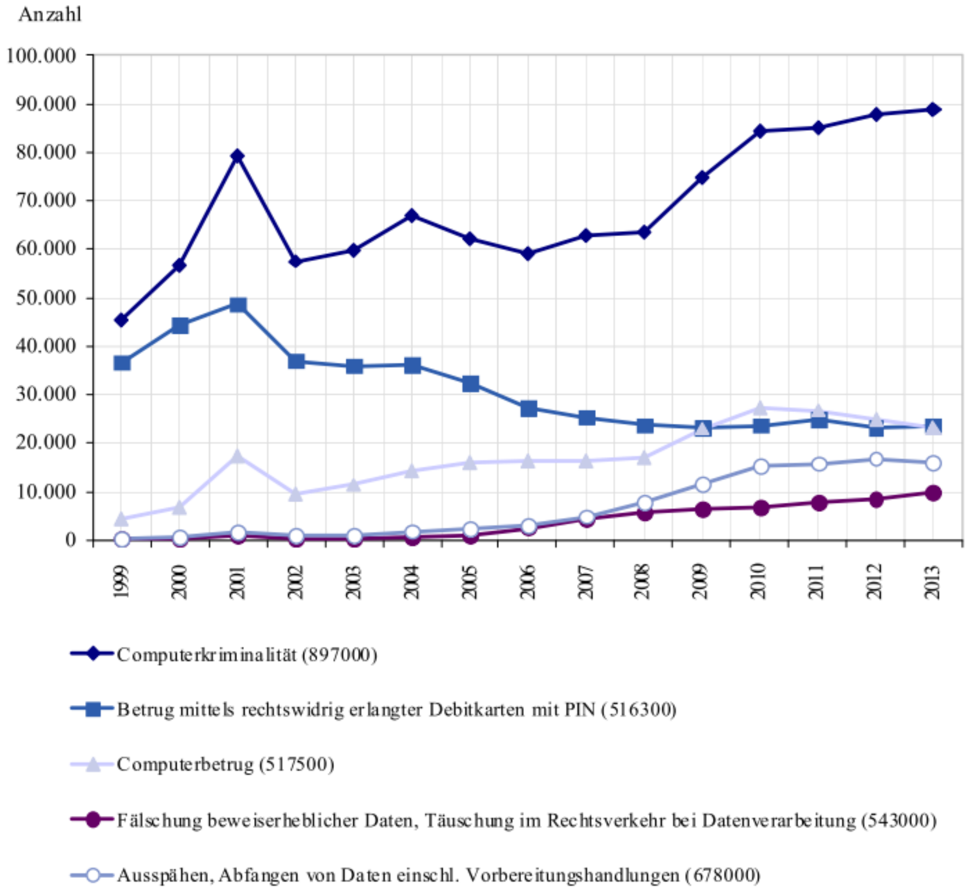
- USA: reports from http://www.ic3.gov
- Sweden: BRÅ responsible for statistics
- Germany: Polizeiliche Kriminalstatistik (PKS)



Statistik från Brottsförebyggande rådet. Siffrorna för 2013 är preliminära.

Statistics of cyber crime - Germany

Entwicklung ausgewählte Delikte der Computerkriminalität
8.5 – G02



Jahrbuch 2013: http://www.bka.de

Some terminology 1/4

- **Computer Emergency Response Team (CERT)**

  (you find also CIRT, I=incident) CERT origin: Carnegie Mellon, USA

- **Attribution**
  - Meta data and other logs can be used to attribute actions to an individual, e.g might identify who place a file in system
- **Alibis and statements**
  - Information provided by those involved can be cross checked with digital evidence
- **Intent**
  - As well as finding objective evidence of a crime being committed, investigations can also be used to prove the intent (known by the legal term mens rea)..

Some terminology 2/4

- **Evaluation of source**
  - File artifacts and meta-data can be used to identify the origin of a particular piece of data; for example, older versions of Microsoft Word embedded a Global Unique Identifier into files which identified the computer it had been created on. Proving whether a file was produced on the digital device being examined or obtained from elsewhere (e.g., the Internet) can be very important.[3]
- **Document authentication**
  - Related to "Evaluation of source," meta data associated with digital documents can be easily modified (for example, by changing the computer clock you can affect the creation date of a file). Document authentication relates to detecting and identifying falsification of such details.

Some terminology 3/4

- **Event**:
  - Unexpected behavior by a system that yields abnormal results or indicates unauthorized use or access, unexplained outages, denial of service, or presence of a virus

- **Incident:**
  - An attempt to exploit a computer network or system such that the actual or potential adverse effects may involve fraud, waste or abuse; compromise of information; loss or damage of property or information, or denial of service. Incidents may also include:
    - Penetration of a System
    - Exploitation / Attempted Exploitation
    - Malicious Mobile Code / Viruses
    - Violations of legal regulations

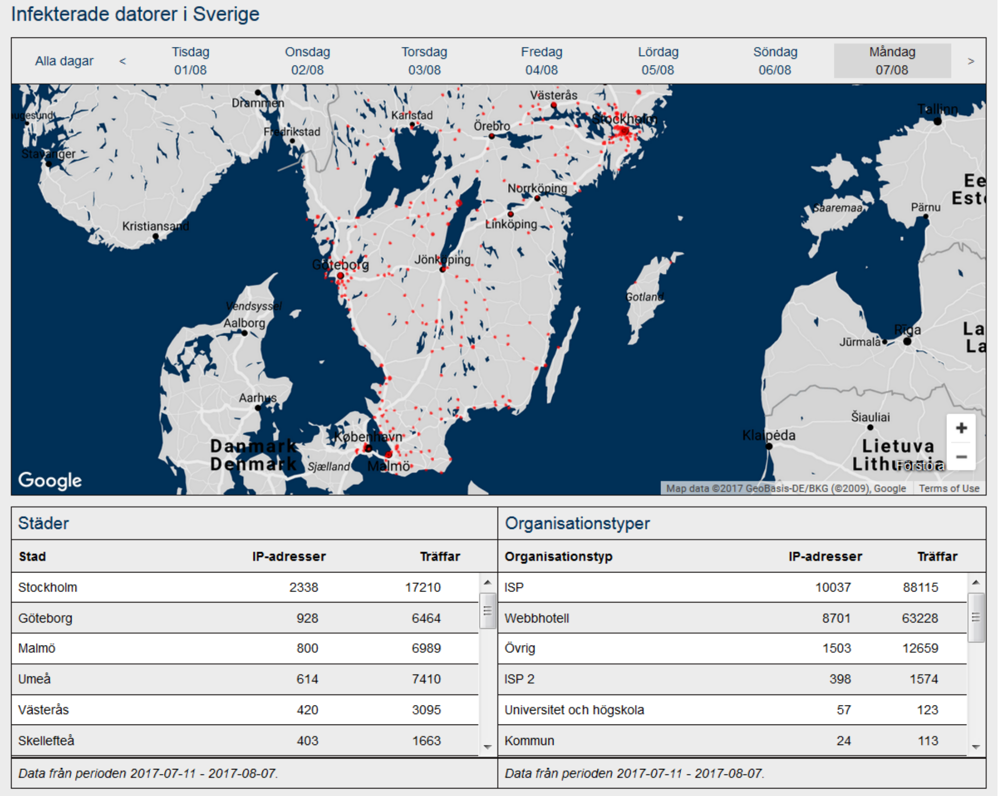
Some terminology 4/4

- **Technical Vulnerability**:
  - A hardware or software weakness or design deficiency that leaves a system open(vulnerable) to potential exploitation, either externally or internally, resulting in the risk of
    - compromise of information,
    - alteration of information or
    - denial of service

- **Administrative Vulnerability**:
  - A security weakness caused by incorrect or inadequate implementation and maintenance of a system's existing security features.

In Sweden: CERT.SE

- http://www.cert.se/



Infekterade datorer i Sverige

Alla dagar	<	Tisdag 01/08	Onsdag 02/08	Torsdag 03/08	Fredag 04/08	Lördag 05/08	Söndag 06/08	Måndag 07/08	>

Städer

Stad	IP-adresser	Träffar
Stockholm	2338	17210
Göteborg	928	6464
Malmö	800	6989
Umeå	614	7410
Västerås	420	3095
Skellefteå	403	1663

Data från perioden 2017-07-11 - 2017-08-07.

Organisationstyper

Organisationstyp	IP-adresser	Träffar
ISP	10037	88115
Webbhotell	8701	63228
Övrig	1503	12659
ISP 2	398	1574
Universitet och högskola	57	123
Kommun	24	113

Data från perioden 2017-07-11 - 2017-08-07.

- Cert.se is part of Myndigheten för samhällsskydd och beredskap, (MSB).

Digital forensics is about:

- Digital Forensic Science – “The use of scientifically derived and proven methods toward the preservation, collection, validation, identification, analysis, interpretation, documentation, and presentation of digital evidence derived from digital sources for the purpose of facilitation or furthering the reconstruction of events found to be criminal, or helping to anticipate unauthorized actions shown to be disruptive to planned operations.” (Palmer, 2001: 16)

- One uses also the term: Computer forensics

(the terms digital and computer forensics are not identical though: digital forensics has a broader scope, see also Altheide’s video)

Who uses digital evidence ?

Many types of criminal and civil proceedings can and do make use of evidence revealed by computer forensics specialists:

- Criminal justice agencies
- Prosecutor's Office/DA, Attorneys, and Judges
- Corporate Councils
- Company Legal resources
- Human Resources (HR=personal avdelningen)
- Auditors
- Individuals
- Crackers/Hackers – Caution !

Main steps in Digital Forensics

1. **Seizure**: steps to get access to(seize) resources where data of interest resides to ensure the preservation of evidence
2. **Acquisition:** Imaging (duplication) of data stored in electronic format
3. **Analysis:** Analyzing the image data
4. **Reporting** results in a factual manner: often to non-technical people.

Digital forensics is not:

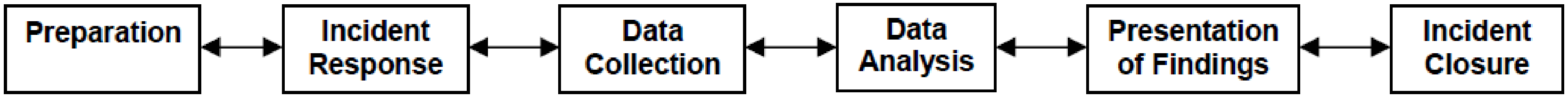
- Pro-active (security)
  - It is reactive to an event or request
- About finding the bad guy
  - It is about finding ***evidence***
- Something you do for fun
  - Proper forensic investigations require expertise
  - Legal limitations on seizing data
- Quick
  - Storage media in excess of 1TB are available
  - Data can be encrypted/hidden

Types of digital forensics:

- Device-level (computer vs mobile) investigation
- Network investigation
- Cloud system investigation
- Software investigation
- Steganographic investigation
  - Digital Imagery
  - Digital Sound
  - Digital Video
  - Encrypted or Embedded Content
  - Watermarking

Investigation frameworks

- Over the years different frameworks have been developed that specify a process for the investigation.
  - Standardize ways of working
  - Quality of investigation can be better relied upon
  - Toolset development



Preparation

Incident Response

Data Collection

Data Analysis

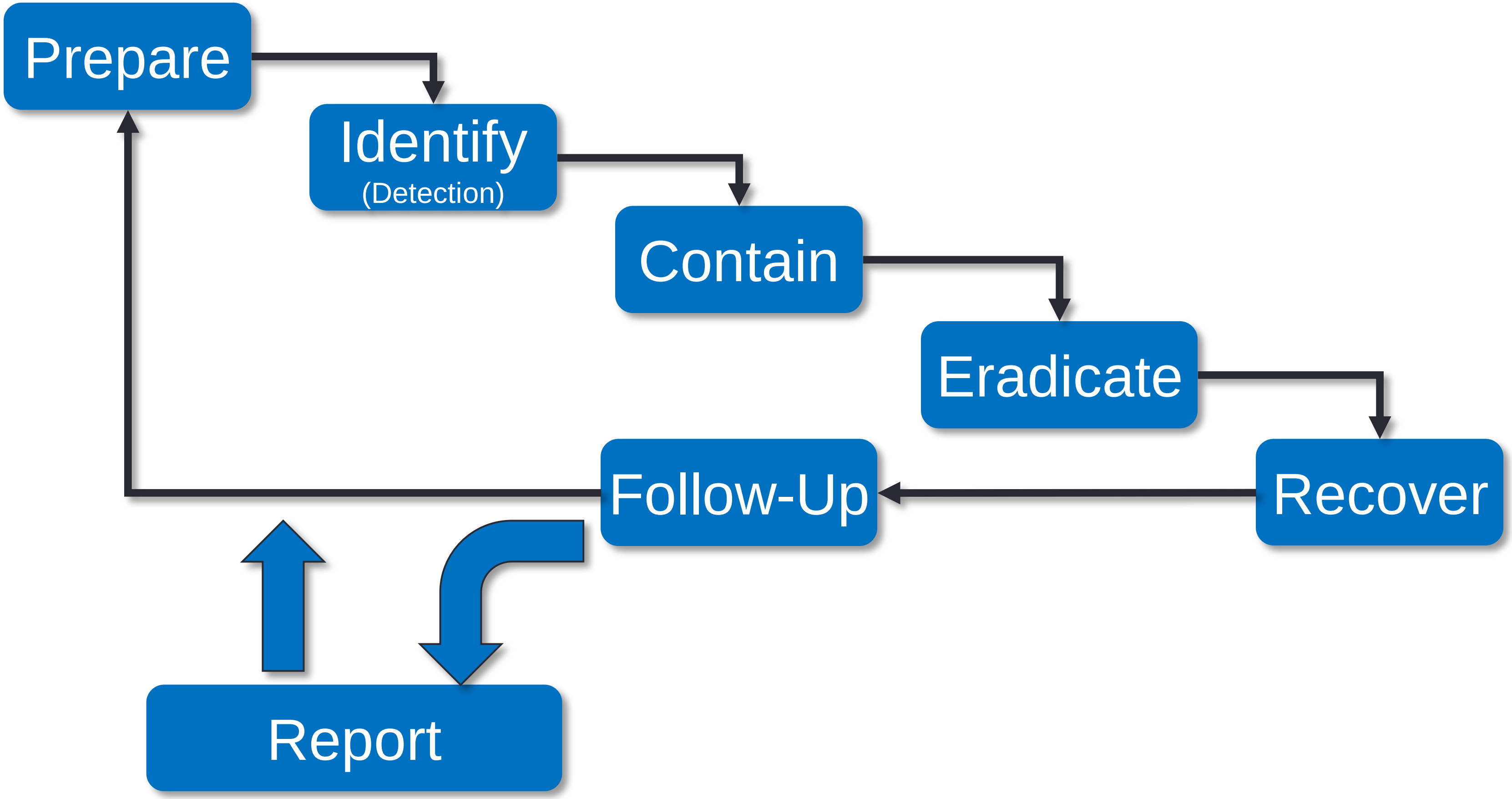
Presentation of Findings

Incident Closure

Figure 1.
Single Tier Digital Investigations Process Framework

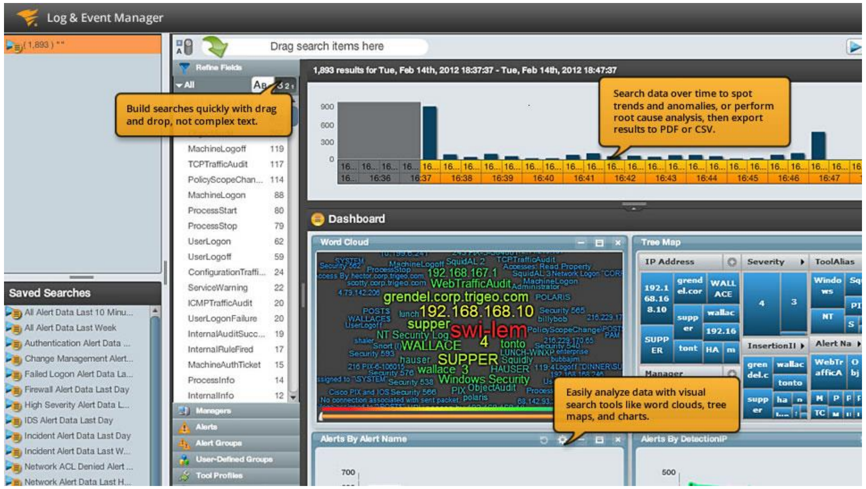
From: A Hierarchical, Objectives-Based Framework for the
Digital Investigations Process, Nicole Beebe, Jan Clark, DFRWS 2004

Incident Handling Basics – it is a process



Investigation: Event Correlation

1. Time-Based Correlation
2. Deviation from, say IETF RFC, Standards
3. Source IP Correlation
4. Target IP Correlation
5. Correlation with Other Data Sources (Data Fusion)



Build searches quickly with drag and drop, not complex text.

1,893 results for Tue, Feb 14th, 2012 18:37:37 - Tue, Feb 14th, 2012 18:47:37

Search data over time to spot trends and anomalies, or perform root cause analysis, then export results to PDF or CSV.

Easily analyze data with visual search tools like word clouds, tree maps, and charts.

MachineLogoff	119
TCPTrafficAudit	117
PolicyScopeChan...	114
MachineLogon	88
ProcessStart	80
ProcessStop	79
UserLogon	62
UserLogoff	59
ConfigurationTraffi...	24
ServiceWarning	22
ICMPTrafficAudit	20
UserLogonFailure	20
InternalAuditSucc...	19
InternalRuleFired	17
MachineAuthTicket	15
ProcessInfo	14
InternalInfo	12

Saved Searches

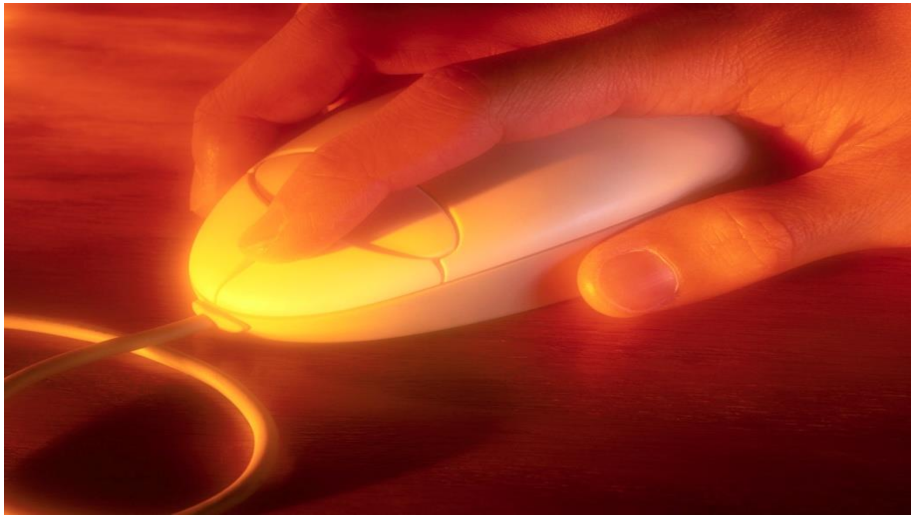
- All Alert Data Last 10 Minu...
- All Alert Data Last Week
- Authentication Alert Data ...
- Change Management Alert...
- Failed Logon Alert Data La...
- Firewall Alert Data Last Day
- High Severity Alert Data L...
- IDS Alert Data Last Day
- Incident Alert Data Last Day
- Incident Alert Data Last W...
- Network ACL Denied Alert ...
- Network Alert Data Last H...

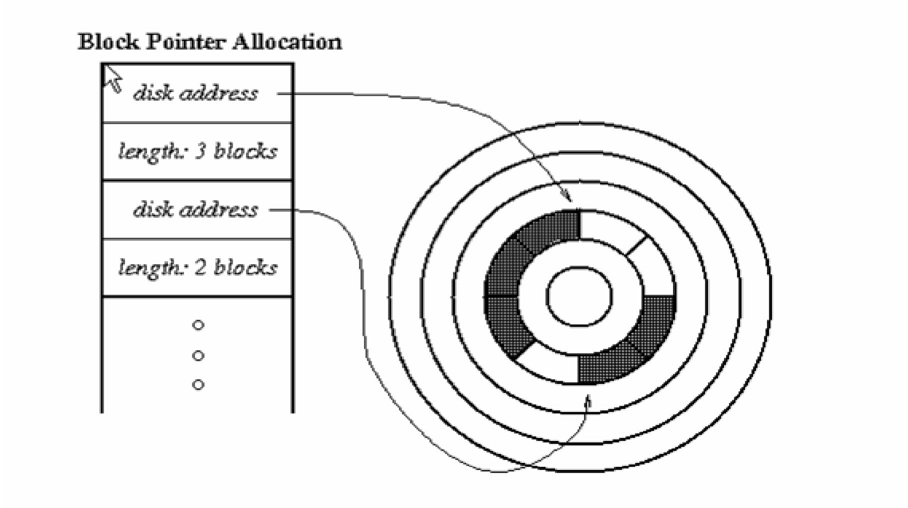
RFC=IETF Request For Comments series of documents

Desired analysis outcome:

Establish links:

- User ⇔ Platform
- Platform ⇔ O/S
- O/S ⇔ Logon
- Logon ⇔ Application
- Application ⇔ Data

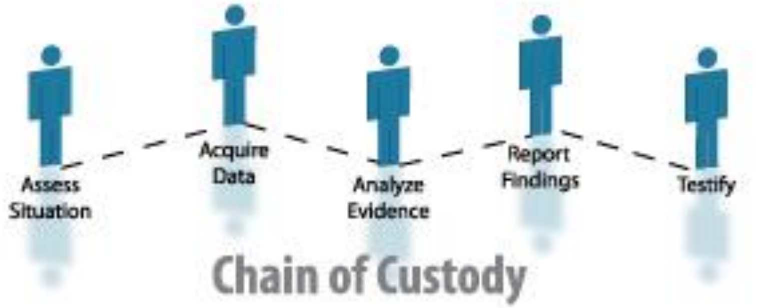




Chain of custody (CoC)

- The notion of CoC has a legal background where it refers to the chronological documentation, showing the seizure, custody, control, transfer, analysis, and disposition of physical or electronic evidence.

- Sv: **spårbarhetskedja**



Sweden

- In Sweden one has so-called ”fri bevisprövning” under which, in principle, unlawfully obtained evidence can be brought to the court.

- Still the quality of the evidence should not be questionable.

Analyzing a system

- When you want to analyze a system a good approach is to boot it from your own (say Linux) system that you know is clean. Most, e.g. BIOS, allow you to select where to boot from.

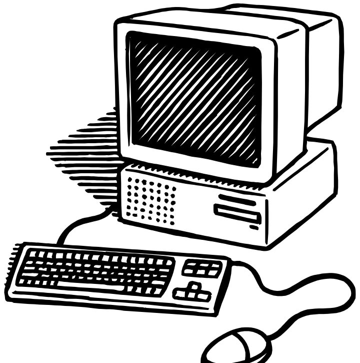
- But secure boot features may prevent this.

Digital forensic process – put it simple

- Observe and evaluate environment
  - “If it is on, leave it on – if it is off, leave it off.”
- Gather and safeguard evidence
- Maintain a clear chain of custody
- Perform an evidentiary evaluation
- Document findings

FORENSICS – HOW? TOOLS

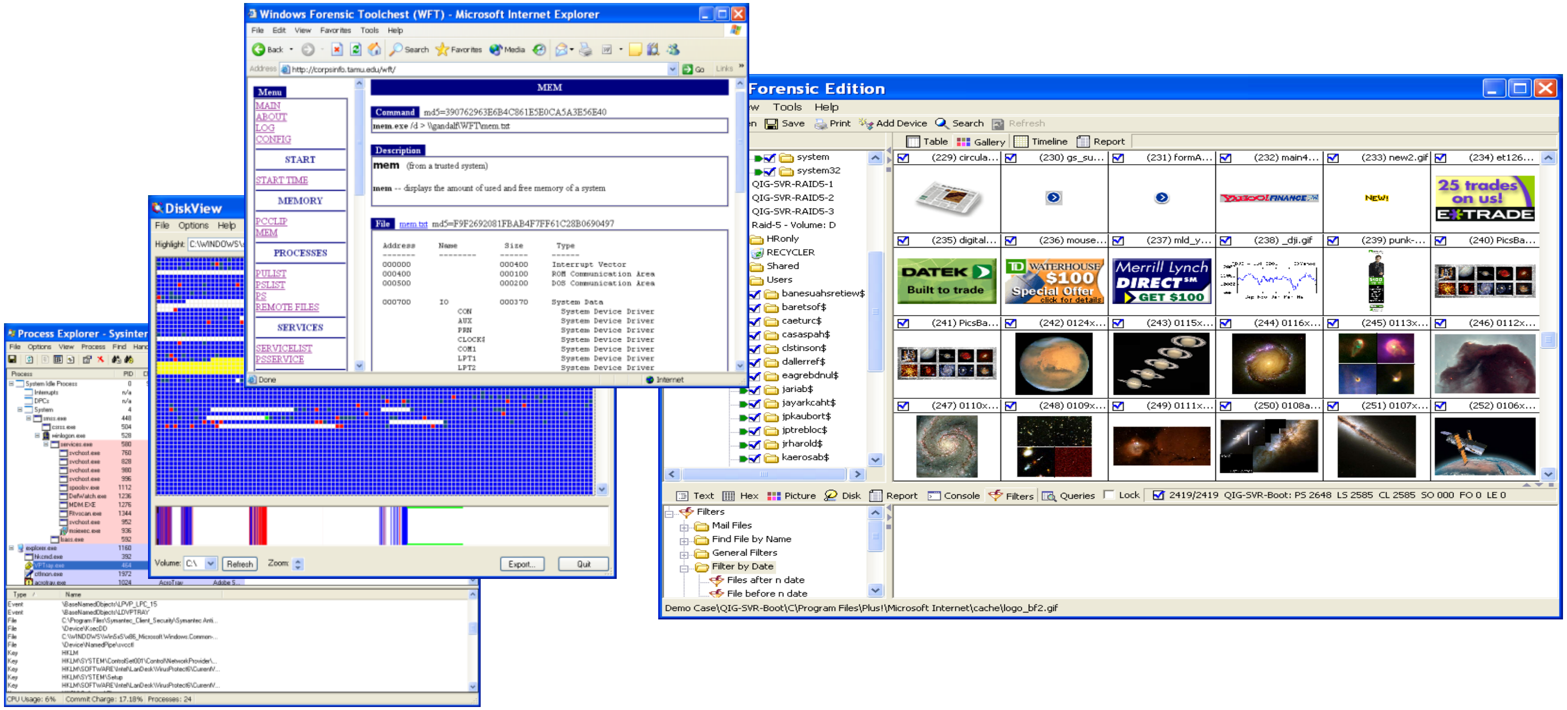
Sources of data:



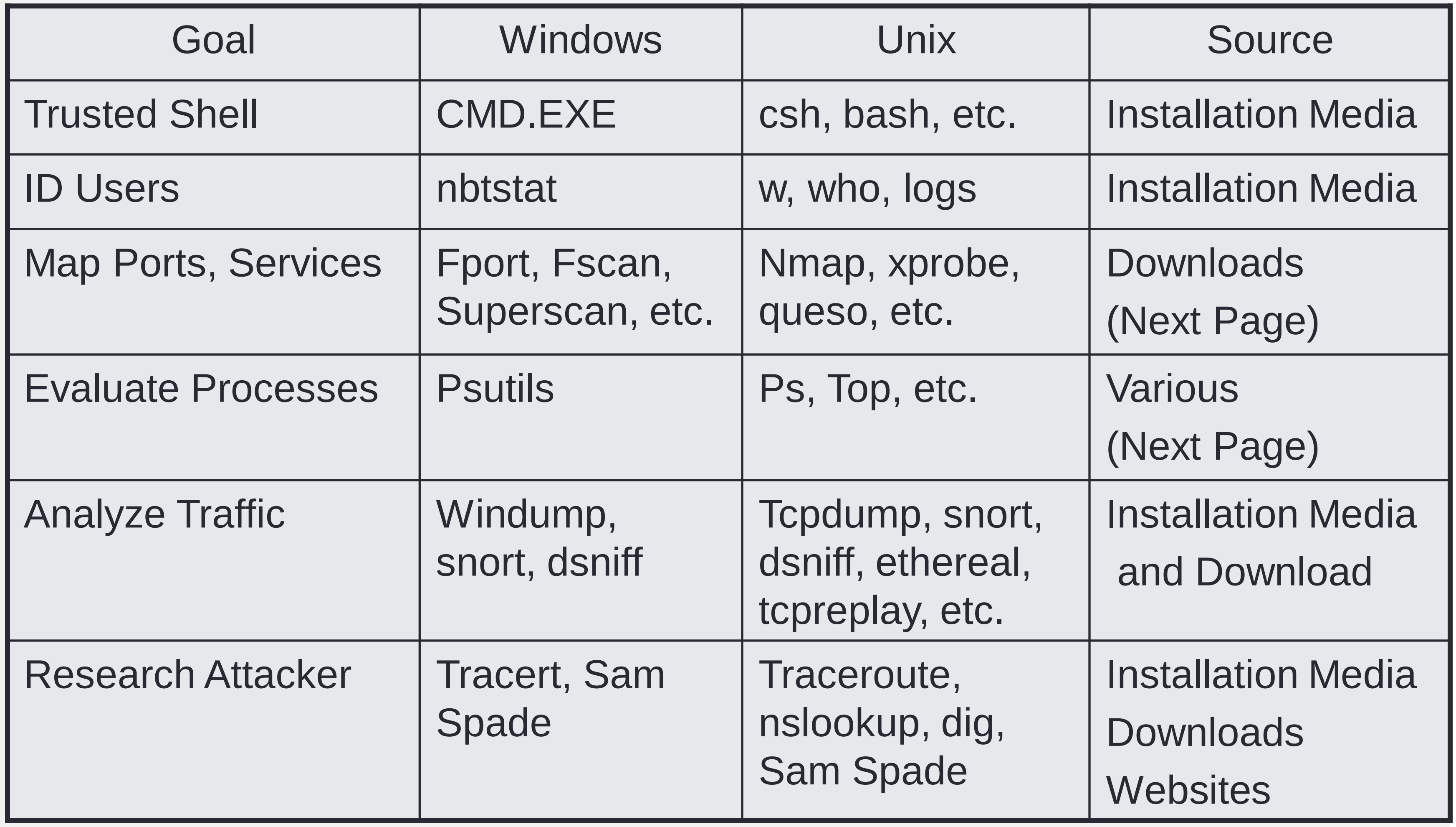
Digital forensic tools:

- Focused Products
  - HELIX
  - WinHex
  - Vision
  - S-Tools
- Log analyzer and SIEM

- Commercial Suites
  - EnCase
  - Forensic Toolkit
  - DMZ F.I.R.E.
  - Maresware



A 1st Responder's Toolkit

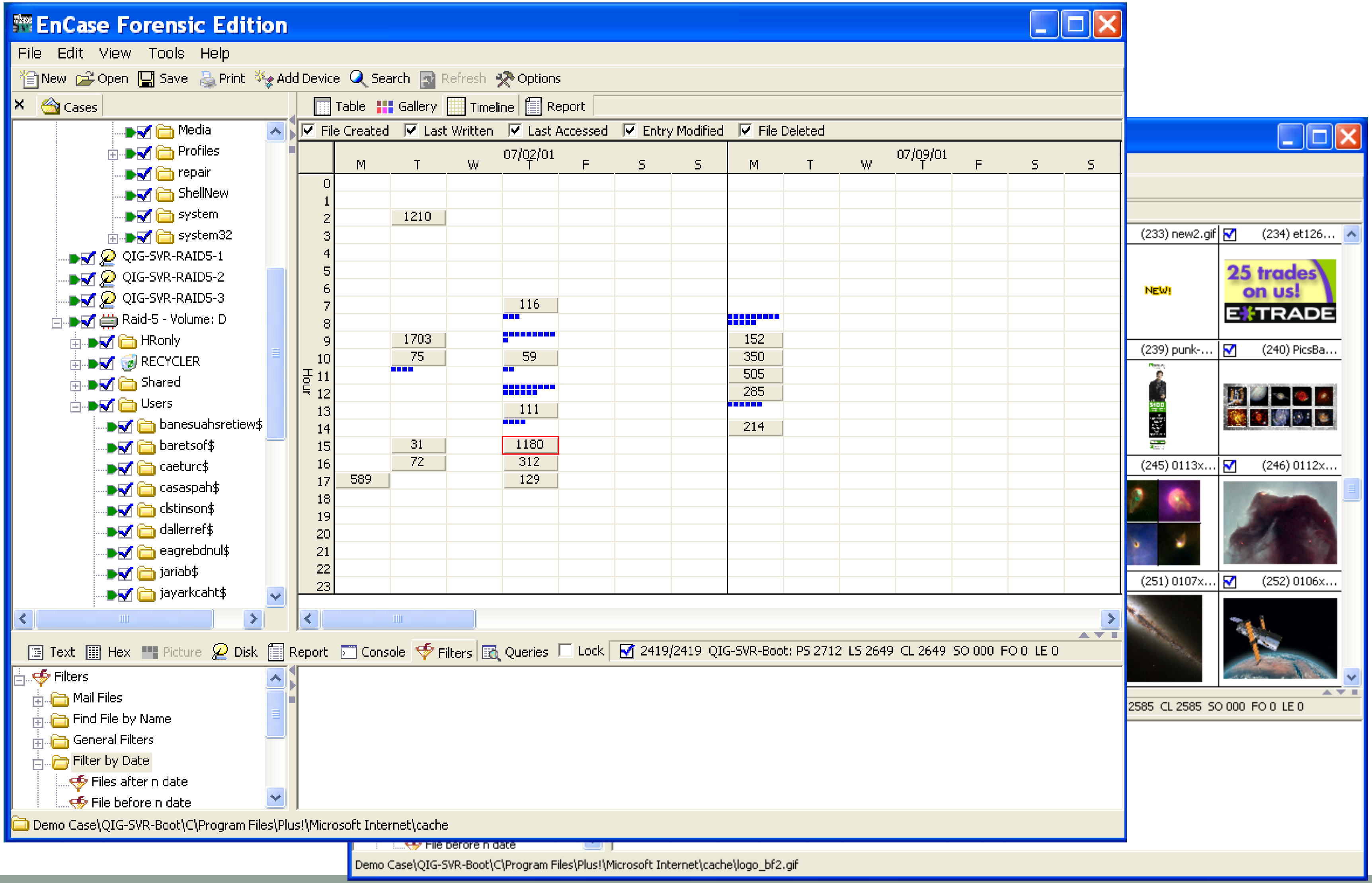


Goal	Windows	Unix	Source
Trusted Shell	CMD.EXE	csh, bash, etc.	Installation Media
ID Users	nbtstat	w, who, logs	Installation Media
Map Ports, Services	Fport, Fscan, Superscan, etc.	Nmap, xprobe, queso, etc.	Downloads (Next Page)
Evaluate Processes	Psutils	Ps, Top, etc.	Various (Next Page)
Analyze Traffic	Windump, snort, dsniff	Tcpdump, snort, dsniff, ethereal, tcpreplay, etc.	Installation Media and Download
Research Attacker	Tracert, Sam Spade	Traceroute, nslookup, dig, Sam Spade	Installation Media Downloads Websites

Digital forensic tools:

- (Drive) Blockers
- Drive Cloning
- Hot-operation Appliances

EnCase



☑ File Created ☑ Last Written ☑ Last Accessed ☑ Entry Modified ☑ File Deleted

				07/02/01							07/09/01			
Hour	M	T	W	T	F	S	S	M	T	W	T	F	S	S
0														
1														
2		1210												
3														
4														
5														
6														
7				116										
8														
9		1703						152						
10		75		59				350						
11								505						
12								285						
13				111										
14								214						
15		31		1180										
16		72		312										
17	589			129										
18														
19														
20														
21														
22														
23														

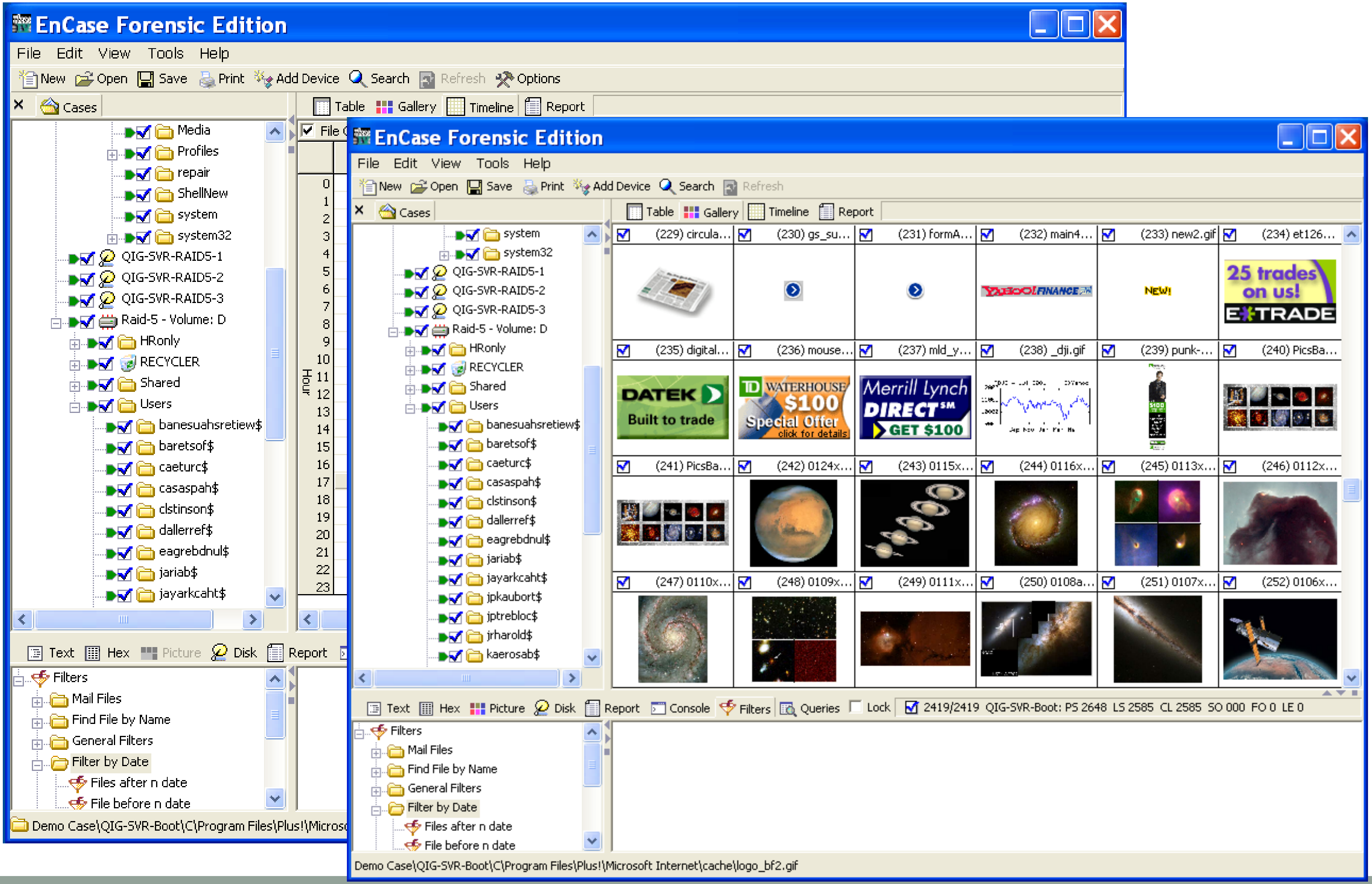
☑ 2419/2419 QIG-SVR-Boot: PS 2712 LS 2649 CL 2649 SO 000 FO 0 LE 0

Demo Case\QIG-SVR-Boot\C\Program Files\Plus!\Microsoft Internet\cache

2585 CL 2585 SO 000 FO 0 LE 0

Demo Case\QIG-SVR-Boot\C\Program Files\Plus!\Microsoft Internet\cache\logo_bf2.gif

EnCase


EnCase Forensic Edition
File Edit View Tools Help
New Open Save Print Add Device Search Refresh Options
Cases
Table Gallery Timeline Report
Media
Profiles
repair
ShellNew
system
system32
QIG-SVR-RAID5-1
QIG-SVR-RAID5-2
QIG-SVR-RAID5-3
Raid-5 - Volume: D
HRonly
RECYCLER
Shared
Users
banesuahsretiew$
baretsof$
caeturc$
casaspah$
clstinson$
dallerref$
eagrebdnul$
jariab$
jayarkcaht$
jpkaubort$
jptrebloc$
jrharold$
kaerosab$
Hour
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
Text Hex Picture Disk Report Console Filters Queries Lock
Filters
Mail Files
Find File by Name
General Filters
Filter by Date
Files after n date
File before n date
(229) circula… (230) gs_su… (231) formA… (232) main4… (233) new2.gif (234) et126…
(235) digital… (236) mouse… (237) mld_y… (238) _dji.gif (239) punk-… (240) PicsBa…
(241) PicsBa… (242) 0124x… (243) 0115x… (244) 0116x… (245) 0113x… (246) 0112x…
(247) 0110x… (248) 0109x… (249) 0111x… (250) 0108a… (251) 0107x… (252) 0106x…
YAHOO! FINANCE
NEW!
25 trades on us!
E*TRADE
DATEK Built to trade
TD WATERHOUSE $100 Special Offer click for details
Merrill Lynch DIRECT GET $100

2419/2419 QIG-SVR-Boot: PS 2648 LS 2585 CL 2585 SO 000 FO 0 LE 0

Demo Case\QIG-SVR-Boot\C\Program Files\Plus!\Micros

Demo Case\QIG-SVR-Boot\C\Program Files\Plus!\Microsoft Internet\cache\logo_bf2.gif

Open source forensics tools

- Look at (for example)
- https://digital-forensics.sans.org/blog/2012/10/06/digital-forensics-case-leads-open-source-forensics-edition
- Sleuthkit and Autopsy browser

- Live-CD, USB, VM.image
  - Kali or DEFT (used in project A)

Network Analysis

- Analyzing a network under attack or determine network usage is requires tools and experience.

- Logs and Tools to extract logs
- SNORT

- Network provides may be forced to support LI (legal intercept functions)

- End-2-end encryption may be a bottleneck. Note mobile networks, classical telephony networks do not have e2e encryption, but one may still gain knowledge such as which endpoints are involved and LI make give access anyhow

Analysis of Documents

- Documents may contain hidden (meta)data;
  - MS Office documents do
  - HTML/XML documents may

- Images may contain hidden data;
  - fingerprints
  - Steganographically hidden data

The Iraq document (1/2)

Back in February 2003, 10 Downing Street published a dossier on Iraq's security and intelligence organizations. The document was released as Word document and one could extract the following data from link below (local copy)

Source: (link no longer working)
http://www.computerbytesman.com/privacy/*blair*.*htm*

The Iraq document (2/2)

- Ten revisions in the log:
- Rev. #1: "cic22" edited file "C:\DOCUME~1\phamill\LOCALS~1\Temp\AutoRecovery save of Iraq - security.asd"
- Rev. #2: "cic22" edited file "C:\DOCUME~1\phamill\LOCALS~1\Temp\AutoRecovery save of Iraq - security.asd"
- Rev. #3: "cic22" edited file "C:\DOCUME~1\phamill\LOCALS~1\Temp\AutoRecovery save of Iraq - security.asd"
- Rev. #4: "JPratt" edited file "C:\TEMP\Iraq - security.doc"
- Rev. #5: "JPratt" edited file "A:\Iraq - security.doc"
- Rev. #6: "ablackshaw" edited file "C:\ABlackshaw\Iraq - security.doc"
- Rev. #7: "ablackshaw" edited file "C:\ABlackshaw\A;Iraq - security.doc"
- Rev. #8: "ablackshaw" edited file "A:\Iraq - security.doc"
- Rev. #9: "MKhan" edited file "C:\TEMP\Iraq - security.doc"
- Rev. #10: "MKhan" edited file "C:\WINNT\Profiles\mkhan\Desktop\Iraq.doc"

Note: Floppy discs were used in distributing the document

- This shows the authors:
  - P. Hamill = Paul Hamill - Foreign Office official
  - J. Pratt = John Pratt - Downing Street official
  - A. Blackshaw = Alison Blackshaw - The personal assistant of the Prime Minister's press secretary
  - M. Khan = Murtaza Khan - Junior press officer for the Prime Minister
  - "cic22" = "Communications Information Centre," a unit of the British Government.

FILE SYSTEM IMPLEMENTATION

Brief overview

Storage media recovery

- Hard disk
- Solid State memory: flash memory, SSDs
- Optical storage
- Cloud Storage

Hard disk: data recovery

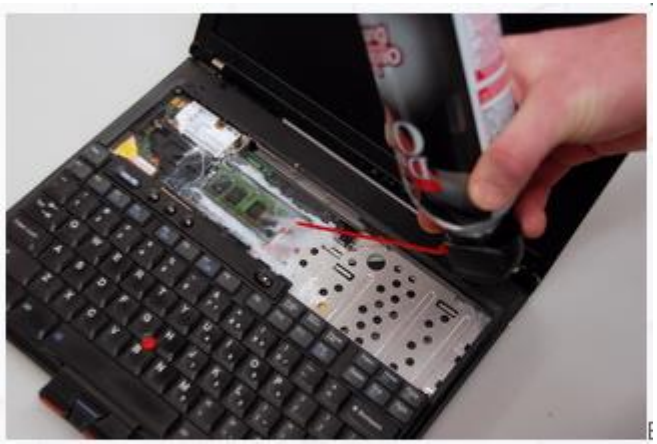
- allocated disk area:
  - used disk space for existing files: very easy
  - deleted/unused space: can be analyzed for data
- slack space: normally not accessible but forensic tools can analyze this type of space

- File Format:
  - Different OS have/support different file systems.
  - We look at a simple one: FAT. Other file systems have more features but use similar principle

Where can we find stored data?

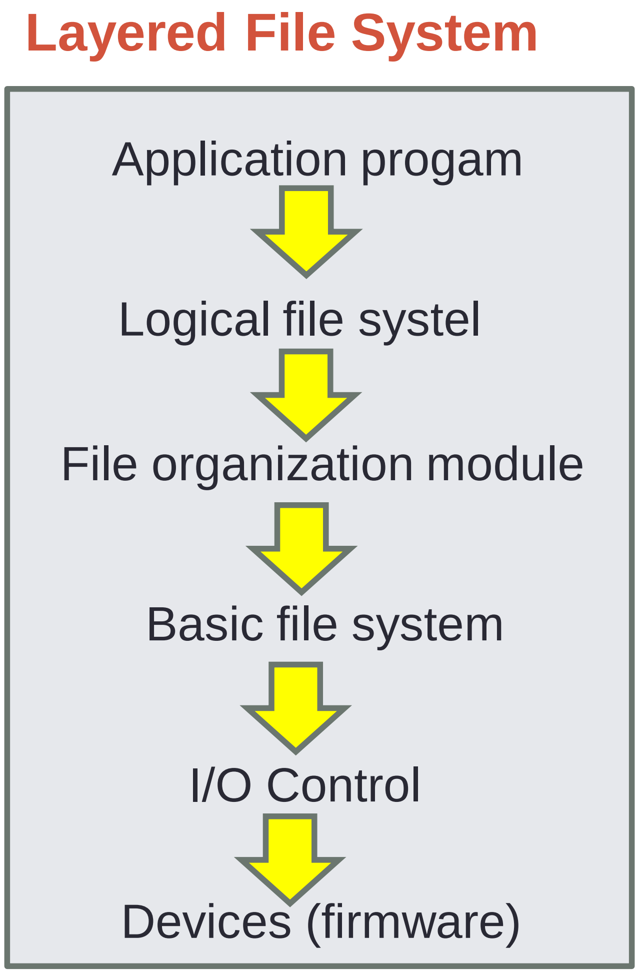
- Non-volatile storage

- Volatile storage



Reading: Lest We Remember: Cold Boot Attacks on Encryption Keys
JJ. Alex Halderman, Seth D. Schoen, Nadia Heninger, William Clarkson, William Paul, Joseph A. Calandrino, Ariel J. Feldman, Jacob Appelbaum, and Edward W. Felten.

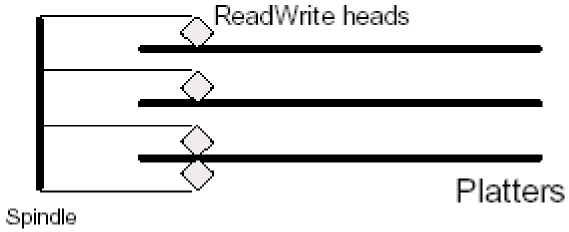
Layered File System

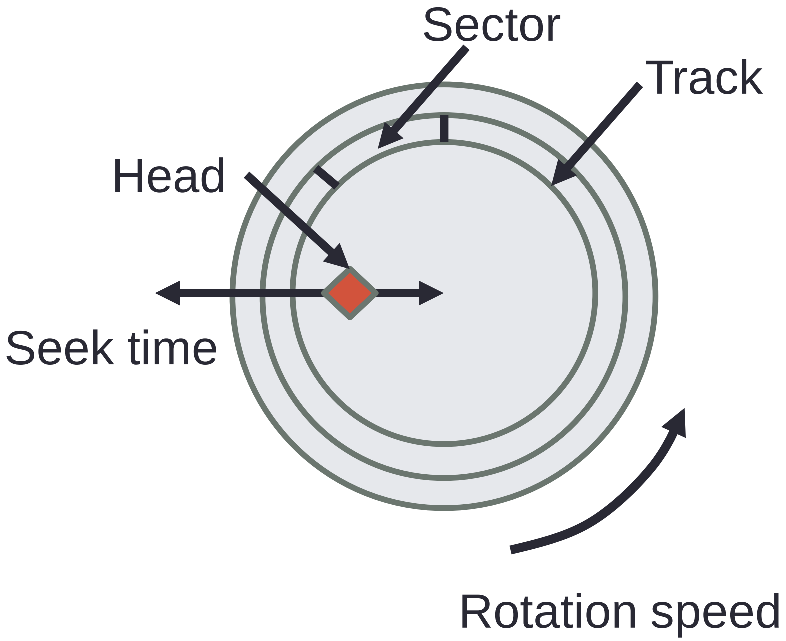


- Logical File System
  - Maintains file structure via file control block (FCB)
- File organization module
  - Translates logical block to physical block
- Basic File system
  - Converts physical block to disk parameters (drive 1, cylinder 73, track 2, sector 10 etc)
- I/O Control
  - Transfers data between memory and disk
- Devices
  - Firmware handles data transfer, storage and internal operations

Physical Disk Drive Structure

- Parameters to read from disk:
  - cylinder(=track) #
  - platter(=surface) #
  - sector #
  - transfer size
  - Read heads (single/double side)
  - Rotation speed (delay)
  - Seek time




Sector

Track

Head

Seek time

Rotation speed

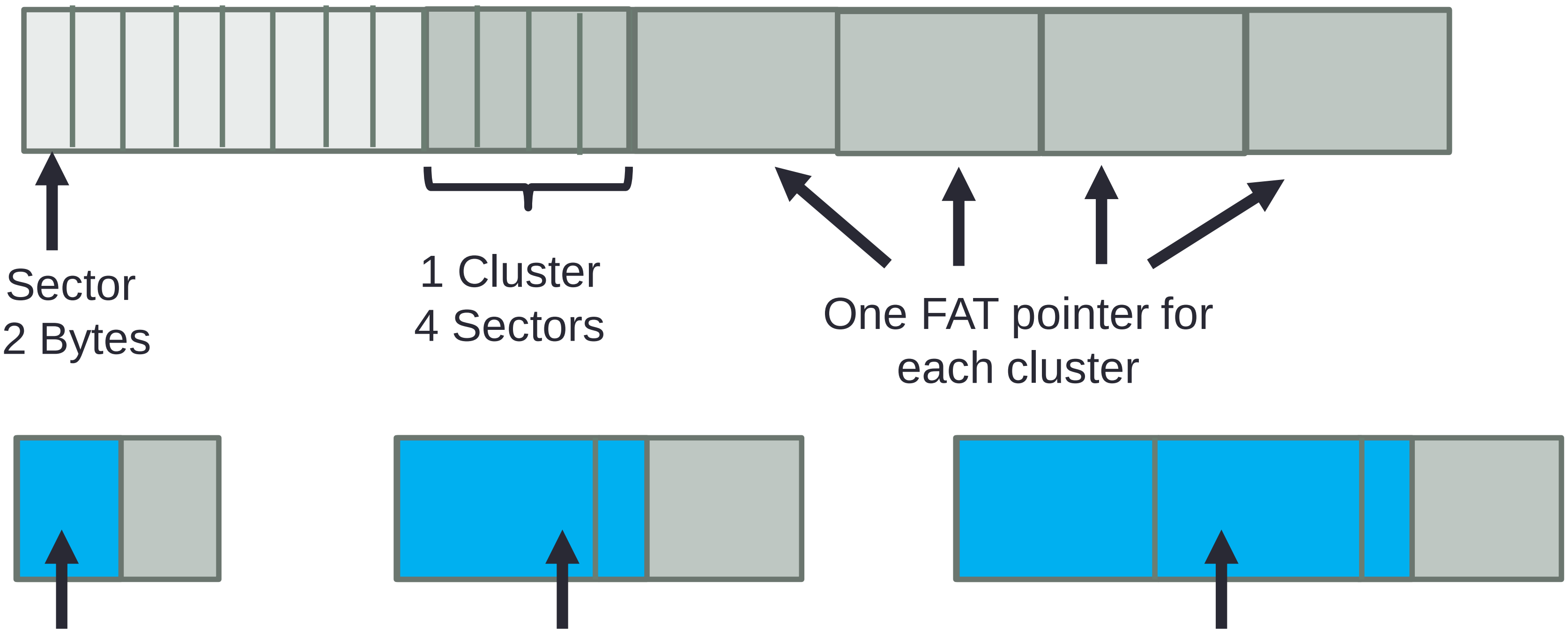
File system Units

- **Sector** – the smallest unit that can be accessed on a disk (typically 512 bytes)

- **Block(or Cluster)** – the smallest unit that can be allocated to construct a file

- What’s the actual size of 1 byte file on disk?
  - takes at least one cluster,
  - which may consist of 1~8 sectors,
  - thus 1byte file may require ~4KB disk space.

Sector → Cluster → File layout



1 Sector
512 Bytes

1 Cluster
4 Sectors

One FAT pointer for
each cluster

1024 Byte file
uses 1 cluster
50% of cluster space
is wasted

2458 Byte file
uses 2 clusters
1638 bytes (40%) of
cluster space is wasted

4506 Byte file
uses 3 clusters
1638 bytes (27%) of
cluster space is wasted

Partitions

- Disks are divided into one or more partitions.
- Each partition can have its own file system method (UFS, FAT, NTFS, …).

”unused” drive space – wasted space

- **Wasted cluster space**

  Space not used in a cluster for a file is wasted as it cannot be used by other file
- **Partition waste space** is the rest of the unused track which the boot sector is stored on – usually 10s, possibly 100s of sectors skipped
  - After the boot sector, the rest of the track is left empty

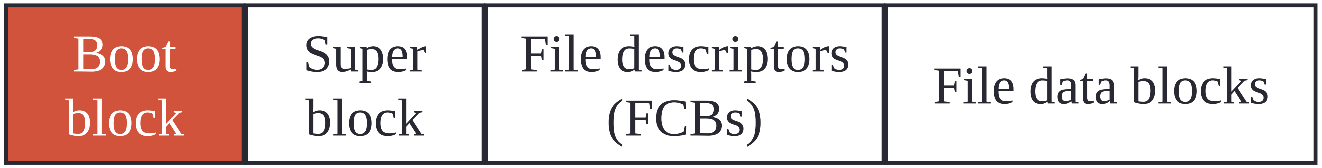
FCB – File Control Block

- Contains file attributes + block locations
  - Permissions
  - Dates (create, access, write)
  - Owner, group, ACL (Access Control List)
  - File size
  - Location of file contents
- FAT/FAT32 → part of FAT (File Alloc. Table)
- UNIX File System → I-node
  - There are many variants of unix file systens
- NTFS → part of MFT (Master File Table)

Remark

- Attributes like time of creation or change have to be interpreted with care.
  - The use may have changed this
  - Or are there guarantees in the system he/she cannot

Typical disk layout for a file system



Boot block

Super block

File descriptors (FCBs)

File data blocks

- Super block defines a file system
  - size of the file system
  - size of the file descriptor area
  - start of the list of free blocks
  - location of the FCB of the root directory
  - other meta-data such as permission and times
- Boot block is located at start, why?
  - Boot image location is indicated in boot block

Boot block

- Dual Boot
  - Multiple OS can be installed in one machine.
  - How system knows what/how to boot?

- Boot Loader
  - Understands different OS and file systems.
  - Reside in a particular location in disk.
  - Read Boot Block to find boot image.

Block Allocation

Different ways to organize memory blocks to a chain of storage space to store data/files

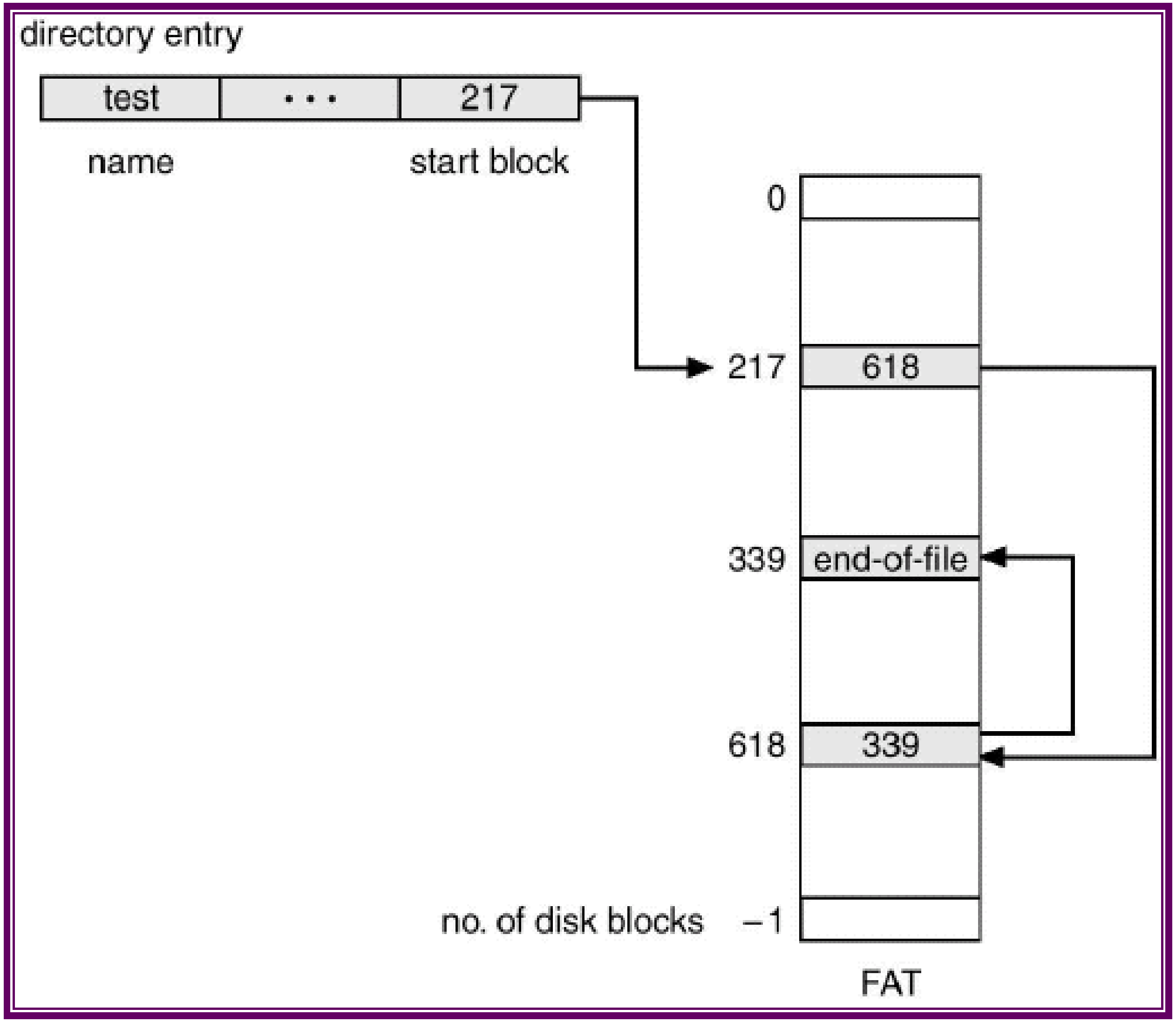
- Contiguous allocation
- Linked allocation
- Indexed allocation

FAT

- FAT == File Allocation Table
- FAT is located at the top of the volume.
  - two copies kept in case one becomes damaged.

- Cluster size is determined by the size of the volume.
  - Why?

FAT block indexing

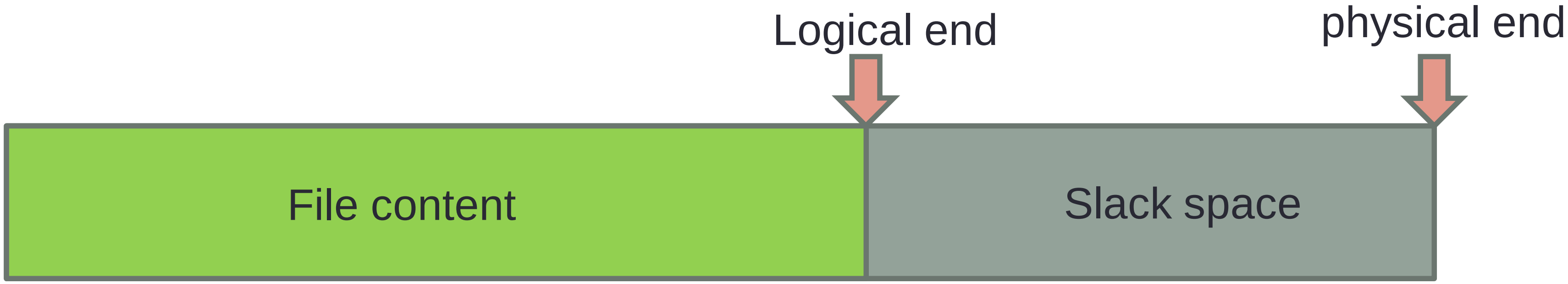


FAT12 Limitations

- limited to 2^{12} or approximately 4096 clusters. (In fact, the number is slightly less than this, since 000h and 001h are not used and FF0h to FFFh are reserved or used for other purposes, leaving 002h to FEFh as the range of possible clusters.)
- A cluster is made up of maximally 8 sectors but can be less (the actual value used is set)
- A sector is 512 bytes , hence a Cluster is 2048 bytes.
- So Fat-12 has a maximum of 4078 clusters * 2048 bytes/cluster = 8 Megabytes. But a floppy disk of 1,4MB needs only
- Two copies of FAT…
  - ➔ still susceptible to a single point of failure!

”unused” drive space – slack space

- **Slack Space** is the space between the logical end and the physical end of file and is called the file slack. The logical end of a file comes before the physical end of the cluster in which it is stored. The remaining bytes in the cluster are remnants of previous files or directories stored in that cluster.
  - Slack space can be accessed and written to directly using a hex editor.
  - This does not add any “used space” information to the drive



FAT(16) and FAT32

FAT or FAT16: Enhancements over FAT12

- More space
  - By having 16 bit entry

➔Thus at Partitions most 2^{16}=65,536 clusters accessible.
➔are limited in size to 2~4 GB
➔Wasted space in each cluster increases (> 200MB)

FAT32: Enhancements over FAT

- More efficient space usage
  - By smaller clusters.
  - Why is this possible? 32 bit entry…
- More robust and flexible
  - root folder became an ordinary cluster chain, thus it can be located anywhere on the drive.
  - back up copy of the file allocation table.
  - less susceptible to a single point of failure.

Volume size V.S. Cluster size

Drive Size	Cluster Size	Number of Sectors
512MB or less	512 bytes	1
513MB to 1024MB(1GB)	1024 bytes (1KB)	2
1025MB to 2048MB(2GB)	2048 bytes (2KB)	4
2049MB and larger	4096 bytes (4KB)	8

Magnetic vs Solid-state disks

- See Altheide’s video

- What do the differences mean for forensics analysis?

Extreme hard disk date recovery

Here the question is whether we can extract old data from an area on a magnetic disk that has been overwritten with data:

Gutmann (1996, Sixth USENIX Security Symposium Proceedings) : ->Yes (in principle)

In practice not easy if at all possible

For SSD: it depends but there could be data left.

Special secure delete functions may be available

Data recovery from flash memory

- Remember: Old data deletion -> may be left
- Relevant for SSD disks

- Interesting reading:
  - Guidelines for mobile phone forensics by NIST
  - Forensics of mobile phone memory: by Svein Y. Willassen. Norwegian University of Science and Technology
  - Forensic Data Recovery from Flash Memory, by M Breeuwsma, et al, Small scale digital device Forensics Journal, vol. 1, no. 1, June 2007

Hard disk and flash memory – architectural differences

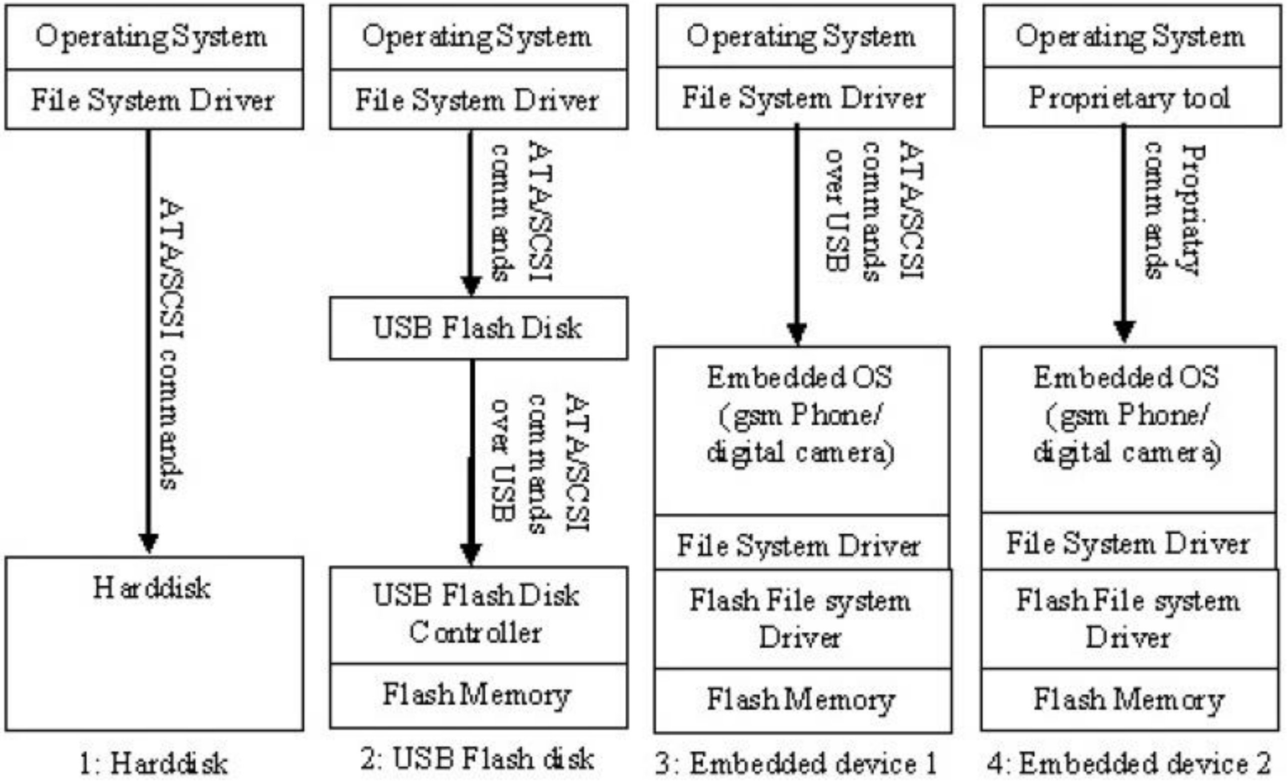


Fig. 3. Components involved in hard disk and flash memory access

From Breeuwsma et al, 2007

Analyzing a mobile phone

- A phone has an application subsystem and a modem subsystem
  - Application subsystem has its own processor(s), RAM and flash
  - Modem subsystem has its own processor and may have its own flash or could share it with the application processor
    - Modem uses a SIM card, holds keys for authentication and session keys for encryption
    - SIM card ID = IMSI
    - Modem ID = IMEI, interesting to track modem with prepaid SIM
  - Many phone hw use so-called JTAG interface for debugging and fault analysis. JTAG access is usually protected.
  - Also boundary scan of is often used which can be used for attack

Mandatory read: FORENSIC ANALYSIS OF MOBILE PHONE INTERNAL MEMORY
Svein Willassen (see literature list)

Forensics and Cloud Technology

- Cloud technology brings new challenges
  - Data can be distributed across multiple platforms that can geographically at different places and be in different countries
  - Cloud systems are elastic: they may grow and shrink on demand
  - Combined with web technologies services are not implemented as ”vertical silos” but as a distributed mesh of (sub)services. Logging use of services to be useful for forensics will be a complex task.

Forensics and Cloud Technology

C Altheide (Google)

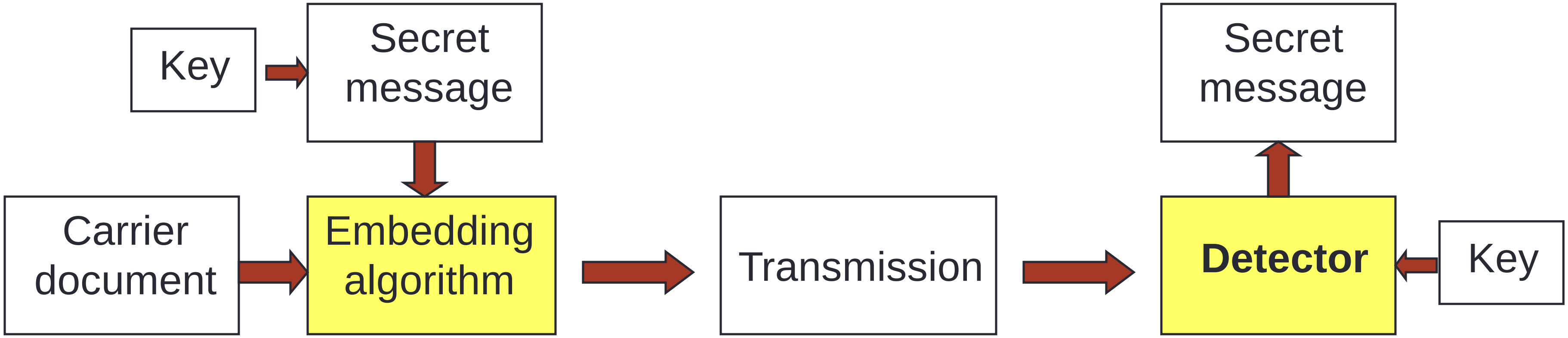
The death of computer forensics

Steganography

- Steganography means “to hide in plain sight” and is derived from the Greek term for *covered writing*. – Kruse & Heiser, 2004
- Automated steganographic tools exist for images, sound files, video, MP3s, documents, and other forms of transport.

1.4MB Source   400KB Message   1.4MB Composite

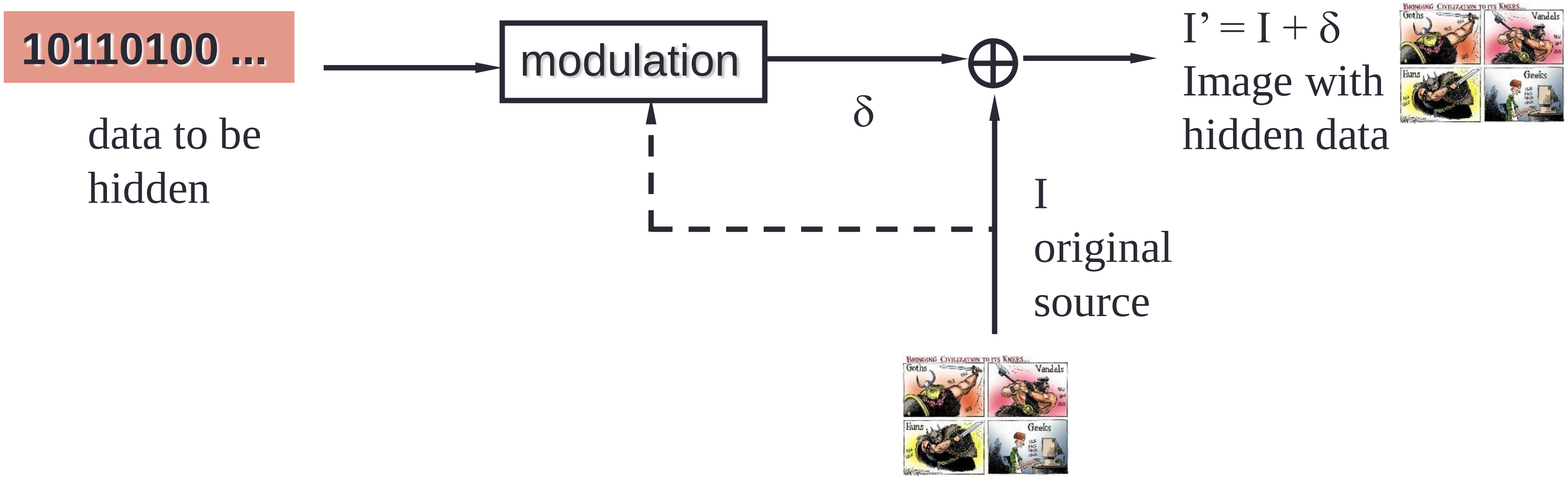
Data Hiding - Definition



- Relationship carrier - message
- Who extracts the message? (source versus destination coding)
- How many recipients are there?
- Type of key a public-key vs symmetric?
- Embedding / detection bundled with a key in a tamper-proof hardware?
- Is the speed of embedding / detection important?

Spread-spectrum Embedding

- Add a noise-like signal and detection via correlation
- Good tradeoff between security, imperceptibility & robustness
- Limited capacity: original signal often appears as major interferer
- Note: spread spectrum technology is traditionally used to combate noise from nature or jammers.



10110100 ...

data to be hidden

modulation

δ

$\oplus$

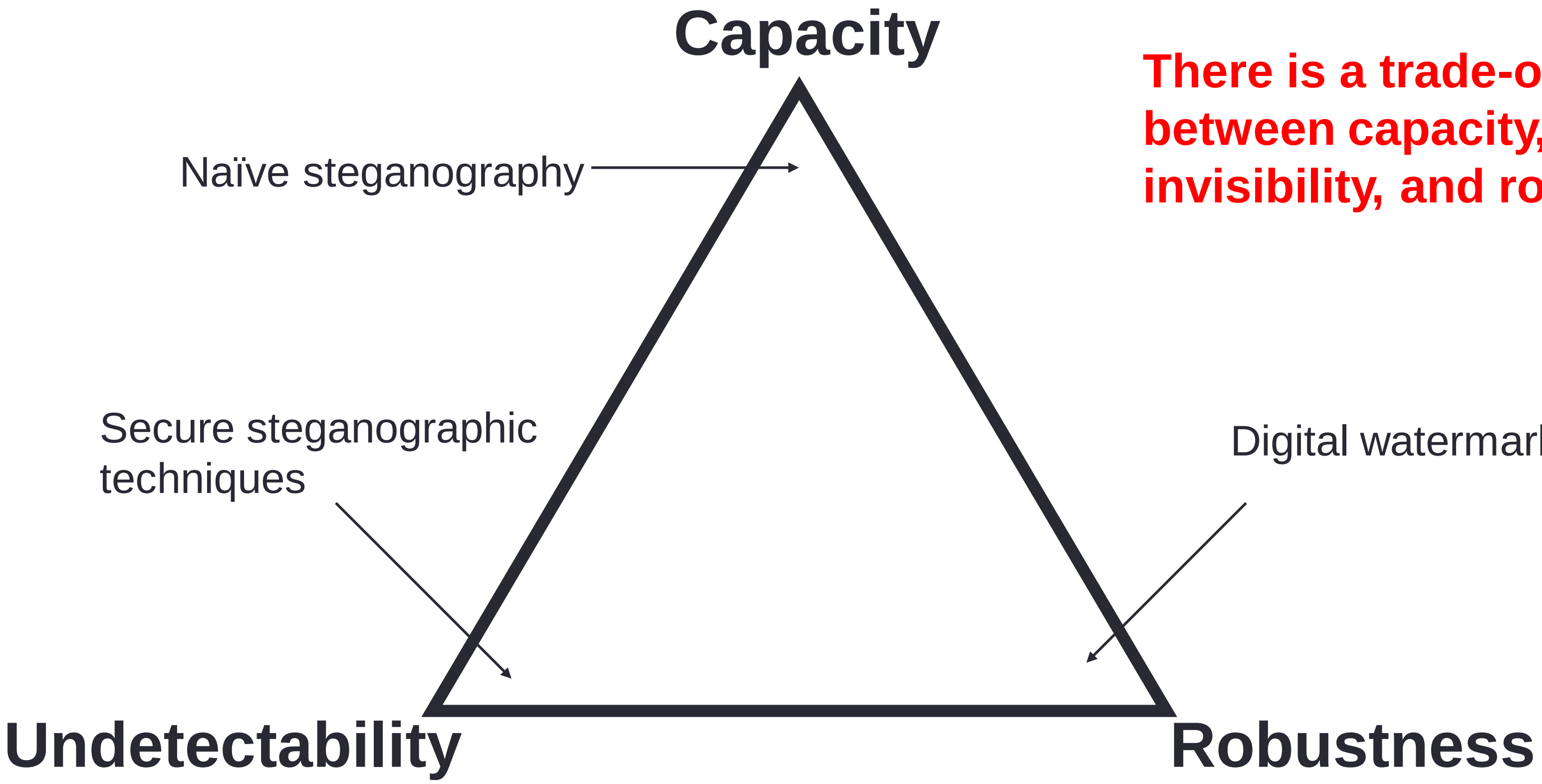
$I' = I + \delta$

Image with hidden data

I

original source

The “Magic” Triangle

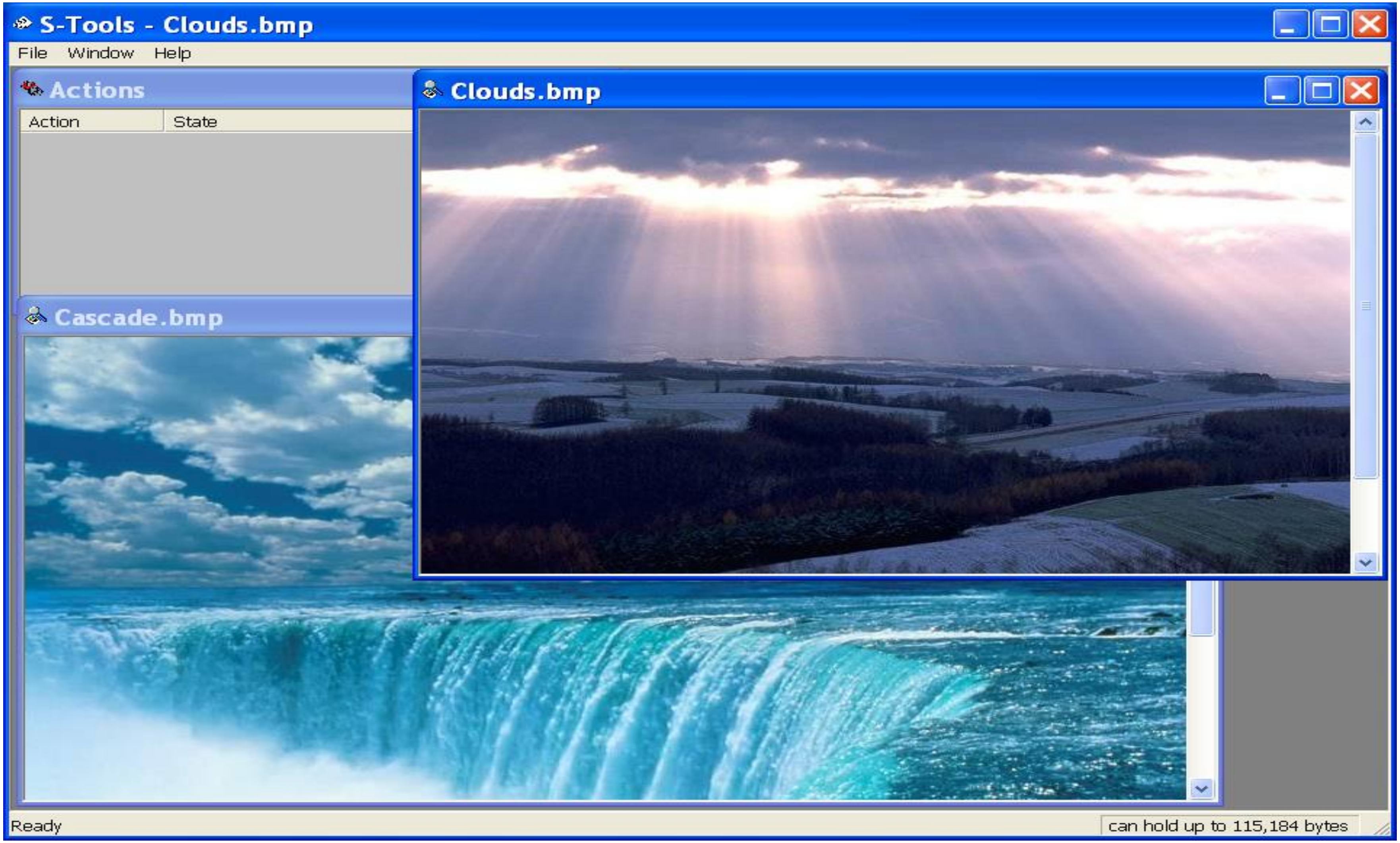


There is a trade-off between capacity, invisibility, and robustness

Additional factors:

- **Complexity of embedding / extraction**
- **Security**

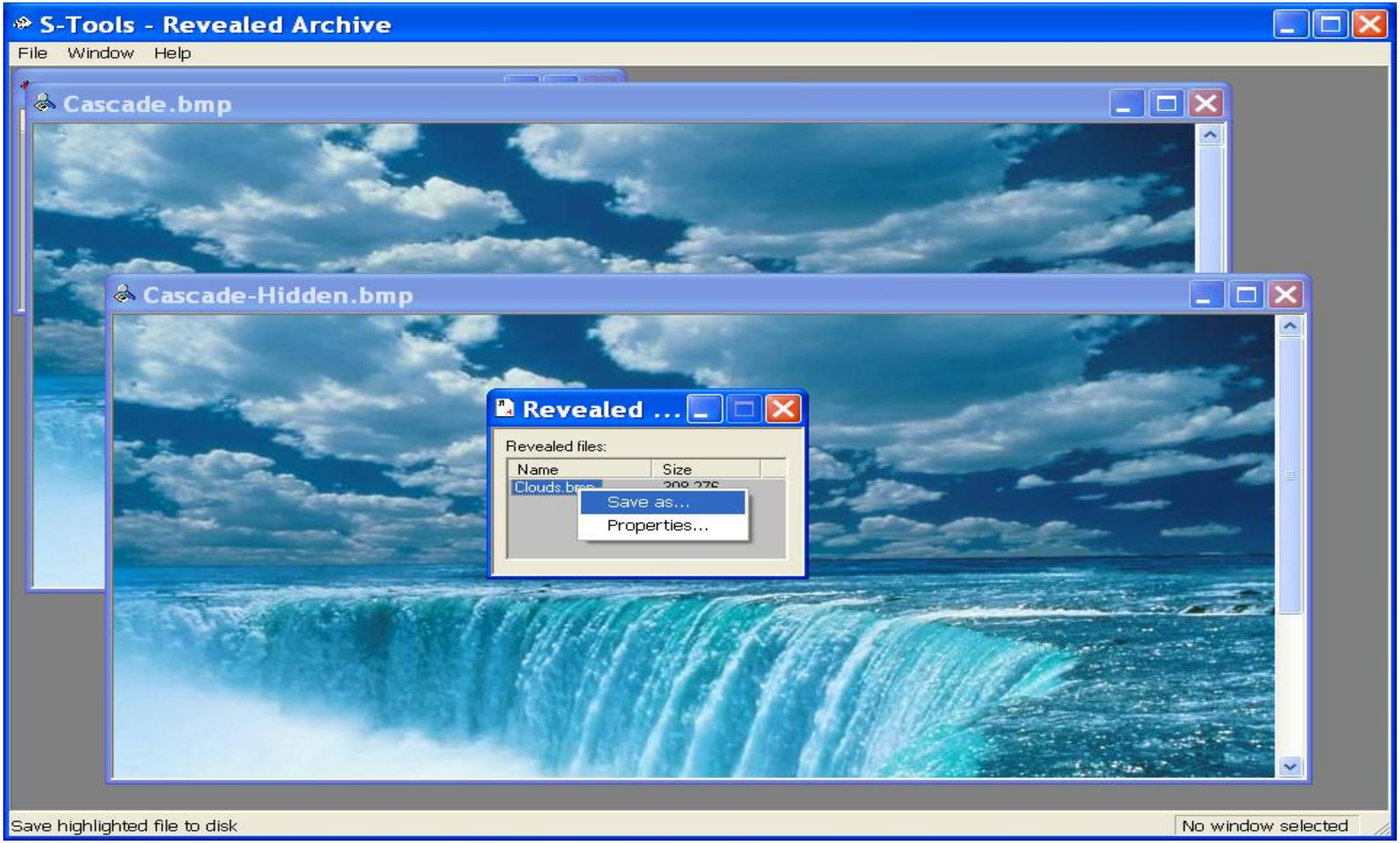
S-Tools



Ready

can hold up to 115,184 bytes

Revealing the hidden image



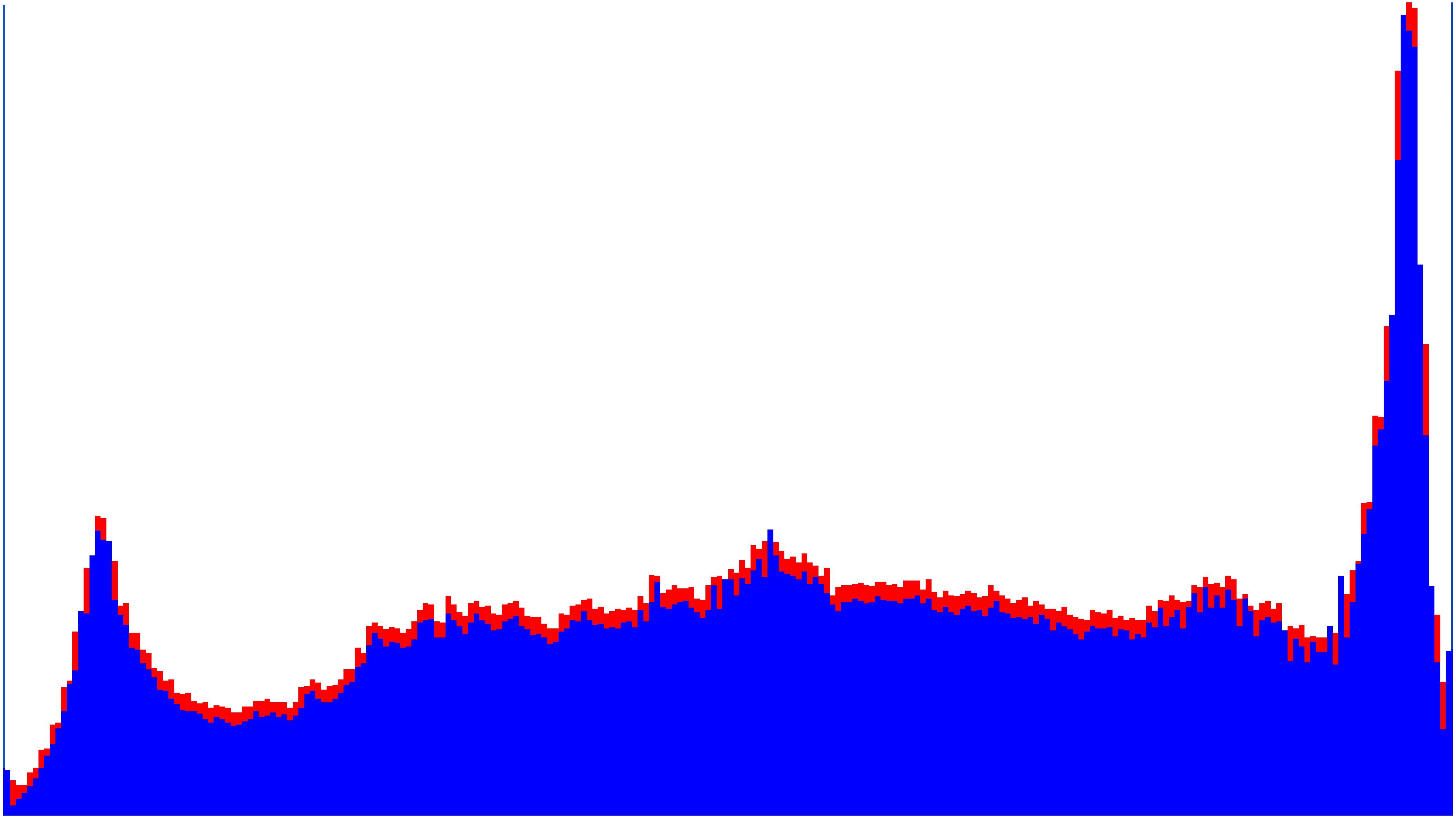
Revealed files:

Name	Size
Clouds.bmp	208 276

Save highlighted file to disk

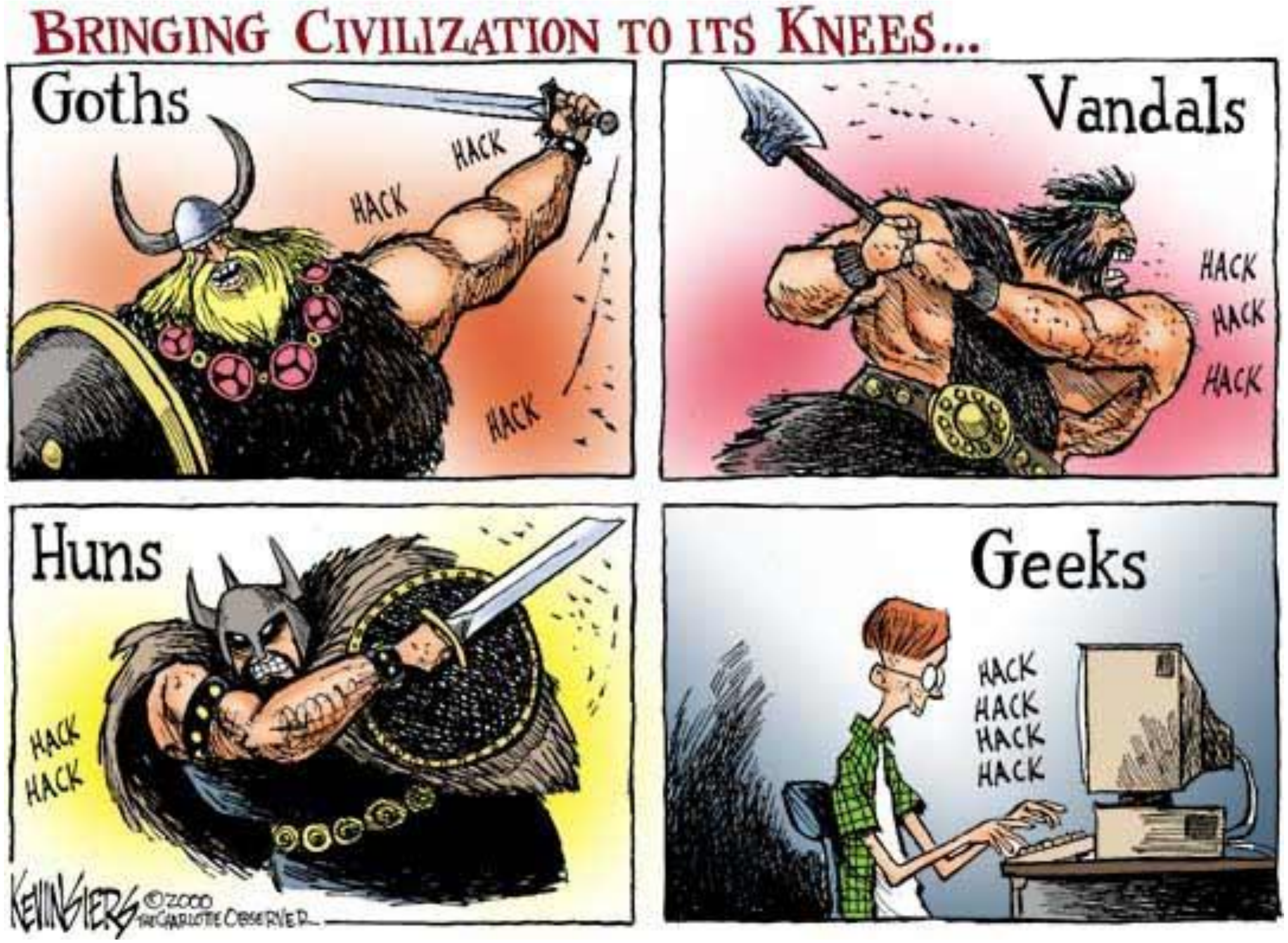
No window selected

Where did the data go?



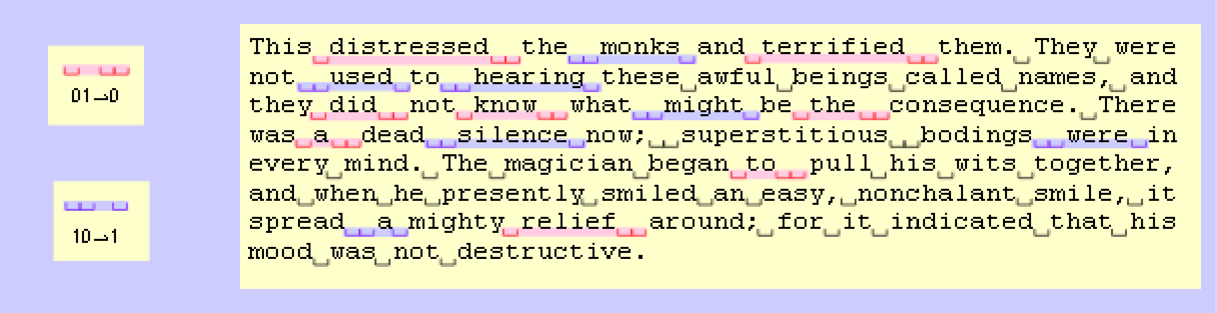
Example

contains what ?


BRINGING CIVILIZATION TO ITS KNEES...
Goths
HACK
HACK
HACK
Vandals
HACK
HACK
HACK
Huns
HACK
HACK
Geeks
HACK
HACK
HACK
HACK
KEVIN SIERS ©2000 THE CHARLOTTE OBSERVER

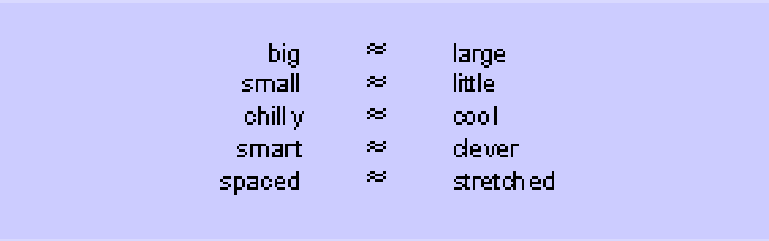
Hiding in Text

Example 1: justification:



This distressed the monks and terrified them. They were
not used to hearing these awful beings called names, and
they did not know what might be the consequence. There
was a dead silence now; superstitious bodings were in
every mind. The magician began to pull his wits together,
and when he presently smiled an easy, nonchalant smile, it
spread a mighty relief around; for it indicated that his
mood was not destructive.

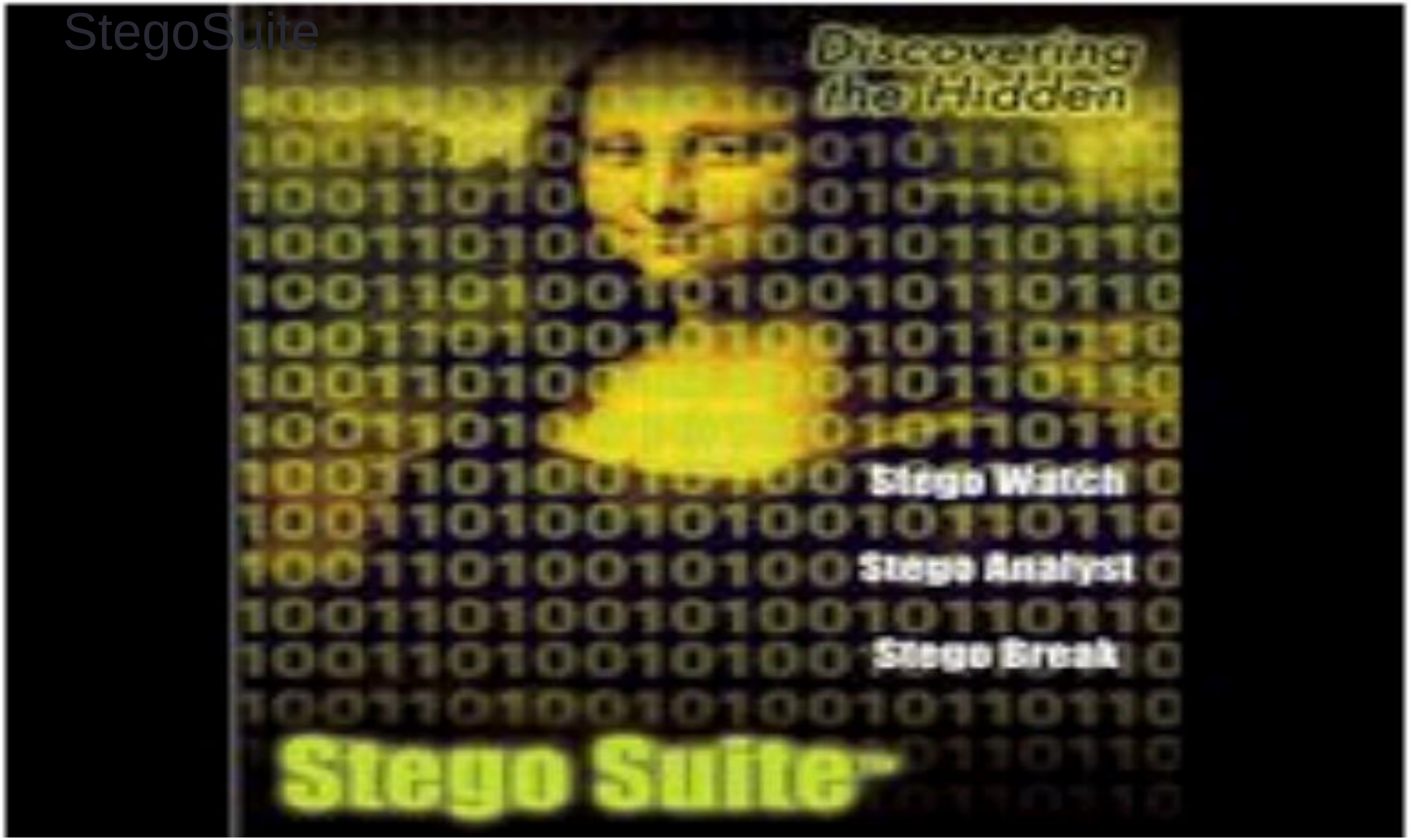
Example 2: synonymous pairs



big	≈	large
small	≈	little
chilly	≈	cool
smart	≈	clever
spaced	≈	stretched

See: Bender et.al. Techniques for data hiding, IBM SYSTEMS JOURNAL, VOL 35, NOS 3&4, 1996

Detecting Steganography - StegoSuite



Links for more information

- Guidance Software (EnCase) https://www.guidancesoftware.com/
- WinHex http://www.x-ways.net/winhex/forensics.html
- Foundstone (Forensic Toolkit) http://www.foundstone.com/
- E-Fense (HELIX) http://www.e-fense.com/helix/
- Computer Forensics, Cybercrime and Steganoraphy Resources http://www.forensics.nl/
- GIAC Certified Forensic Analyst Practical Papers Review http://www.giac.org/GCFA.php
- Video on steganography using images: construction and detection (https://youtu.be/TWEXCYQKyDc)
- tools: http://www.sysinternals.com/
- SANS Reading Room http://www.sans.org/rr/