Feedback of a Digital Forensics Expert - Part 1

Emmanuel Giguet

Feedback of a Digital Forensics Expert

Emmanuel Giguet
CNRS Researcher at GREYC Lab

emmanuel.giguet@cnrs.fr

Preliminaries

- Legal Expert for a Regional Court of Appeal
- During 10 years: From 2005 until 2015
- Average of 3 cases per year
- Independent expert
  - « Free » : no pressure, no advices from authorities
  - Reproducible Method : Forensically sound
  - Legally responsible (specific insurance)
  - Stay focus on the case and write facts

Different Kind of Cybercrimes

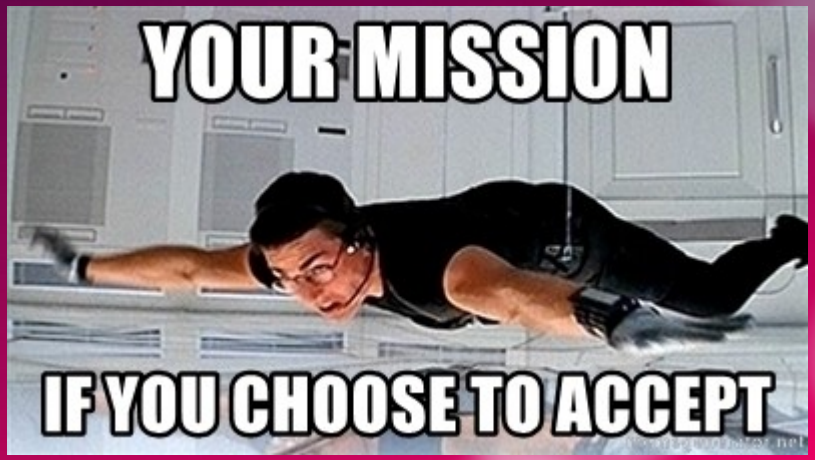
- Social fraud : forgery of official documents
- Trafficking in Intellectual Property
- Distribution and possession of child porn material
- Contacting children through the Internet
  - by sex predators/offenders or by "haters"
  - In order to get photos or videos
  - To meet the victim face to face
- Unconsented adult prostitution

Legal Framework (1/2)

Origin of the requisitions :

- Security forces (National Police / Gendarmerie => law enforcement services)
  - An Officer of the Judicial Police in charge of the investigation
- The regional public prosecutor
- A judge of the regional court

Accept or reject the mission ?



Legal Framework (2/2)

In case of preliminary investigations :

- A complaint against someone
  - The investigation has started
  - During the search warrant, the Force has seized devices
- A search warrant given by a Court judge
  - Help the Forces to locate and seize the material
  - Start the digital investigation on site
- A suspect is placed in custody by Forces
  - During 48 hours : suspicion of crime

About the mission (1/2)

- Locate and extract all information related to
  - Nominative information : names, pseudos, (filenames)
  - Media Type : (explicit/forged) content / categorized / counts
  - All information related to the possession and distribution of content
  - List the contacts, rebuild discussions, extract relevant excerpts
- Two main kind of targetted information :
  - Textual information
  - Images (photos) and videos
- To understand what happen :
  - Dive inside people’s life : deep inside...
  - Hope to rapidly find traces, evidences…

About the mission (1/2)

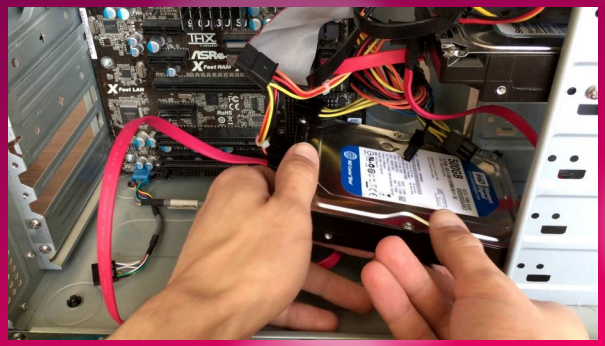
- Screen carefully the entire content :
  - The visible part
  - The invisible part
- Achieve Multi-source Analysis :
  - Computer Event Logs, Filetime Information, Web Browser Activity
  - Timestamped content => check/convert datetime
  - Media Content => Mail in Browser Cache, Attachment on FS
- Achieve Cross-device / Cross-computers Analysis :
  - Connected Devices : Run on a computer / Store in an external devices
  - Logs, Cache, Temp files / Content
  - Simultaneous usage of devices : Chat with a computer // Send a message by phone

The playground (1/2)

- In a single case, several materials to handle :
  - 1 or 2 tower computers (incl. 1 or 2 internal disks : 250-500Gb each),
  - 2 portable computers,
  - 1 or 2 external disks : 250-750Gb each
  - 2 or 3 USB Keys, 5 to 10 writable Cds or DVDs
- Analyze each device and Perform a Cross Analysis
  - Large quantity of information to handle : visible+recoverable
  - Time consuming process for a single person
  - Sources of errors and omissions are multiple

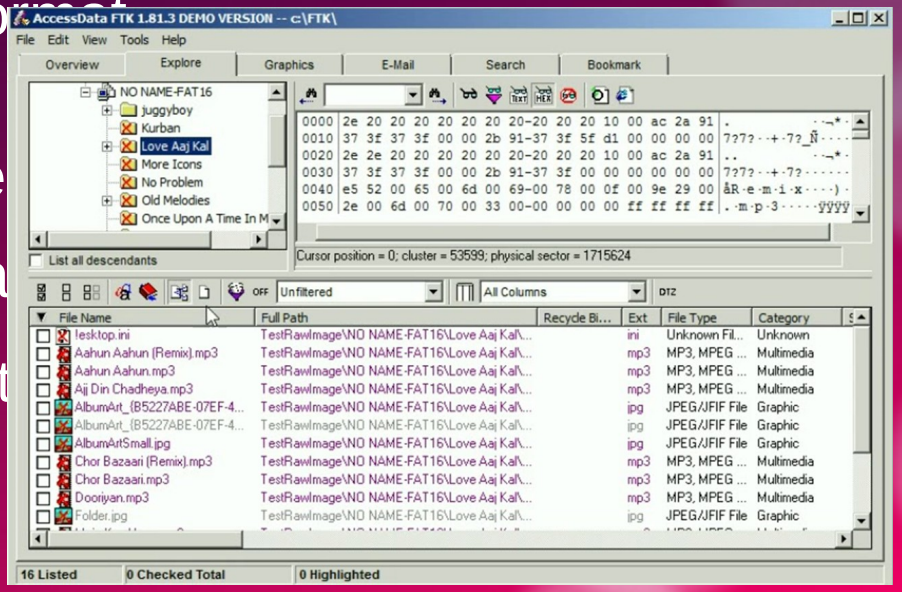
The playground (2/2)

- Concerning the Materials :
  - Mostly Tower Computers and Mobile Computers
  - Mostly Mechanical Disks
  - No disk encryption by default, or security by design
  - No mobile phones : analyzed by Forces with dedicated extraction devices (Cellebrite UFED)
- Concerning Internet Activities :
  - Search Activity
  - Access to online services : Mail, Bank, Travel & Hotel Booking, Maps
  - Social Networks and Instant Messaging : various adhoc tchat platforms
  - Peer to peer (file sharing) activities
  - No cloud services except for Mail : various mail hosting platforms



Technological Context

- Digital investigation platforms :
  - EnCase, FTK, Autopsy, X-Ways, Magnet IEF/Axiom
  - To extract traces and potential evidences
  - To manage the case
- Specific Forensic Tools :
  - To extract data from proprietary fo
    (Security Exploded)
  - To recover deleted file / To do File
- Not free, not open-source, not eva
  - Am I missing something important



Forensic Data Acquisition

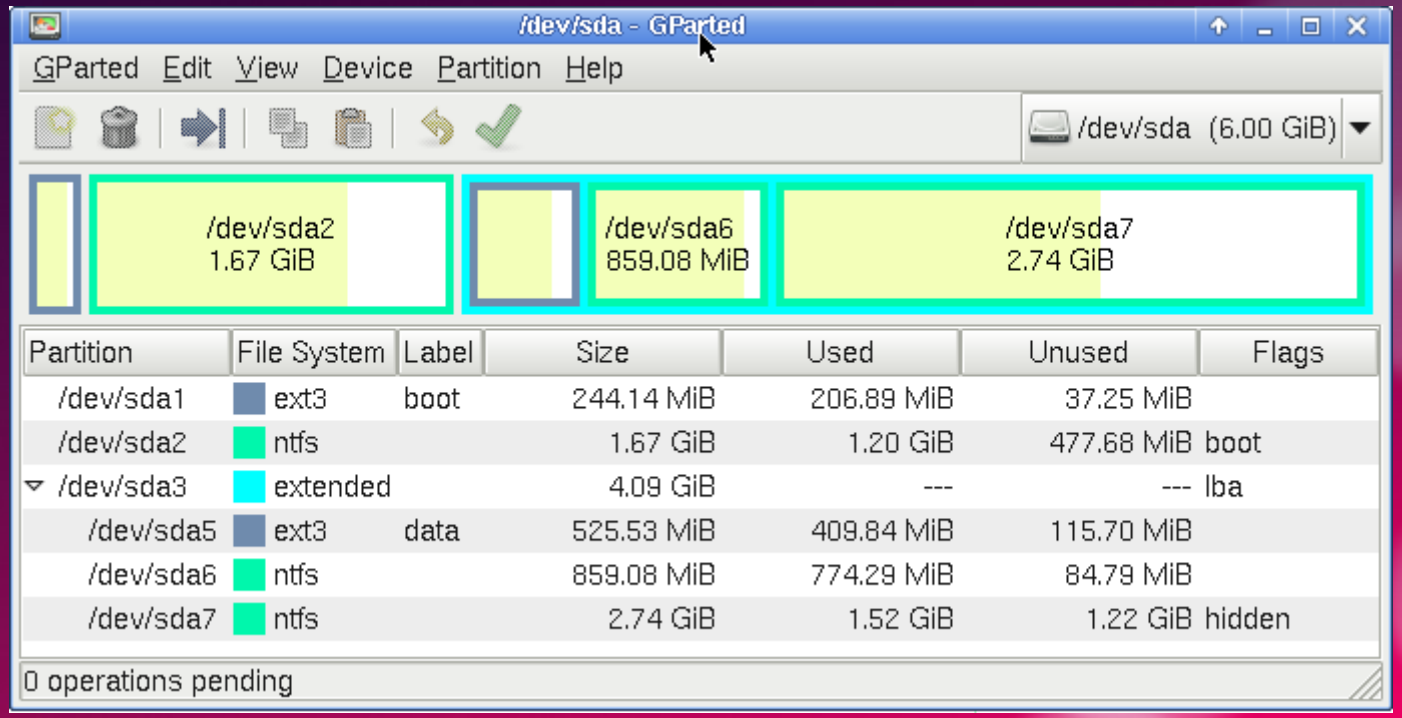
- Follow forensically sound procedures
  - Original data has not been altered
  - Ensure that evidences will be accepted by the Court
- Software vs Hardware write blockers
  - Work in a safe environment
  - Disable OS automount or autorun / mount -ro
  - Set Disk Jumpers
  - Use professional Write blockers & Disk imagers

Forensic Data Acquisition

- Forensically Sound Data Capture
  - Use Write blockers
  - Create disk image
  - Bit-by-bit copy with hashing on the fly (dcfldd, dc3dd)
- Limits
  - Disk stress
  - Time constraints

Starting the investigations

- Discover disk structure
  - Check Partitioning : os, data, recovery, swap, ...
  - Discover Unallocated Space



Starting the investigations

- Two strategies :
  - Examine via a terminal : sfdisk, mount
  - Boot the image in a virtual environment
- Mounting the image :
  - OS Name and version
  - Discover user accounts / Crack user passwords
  - Analyze with Forensic Tools & Forensic Platforms
- Boot the image :
  - Help diving in the case : Understand how the user works
  - Organization of the desktop, background image, shortcut icons
  - Help to rapidly get insights for easy cases

Digital Forensic Activities

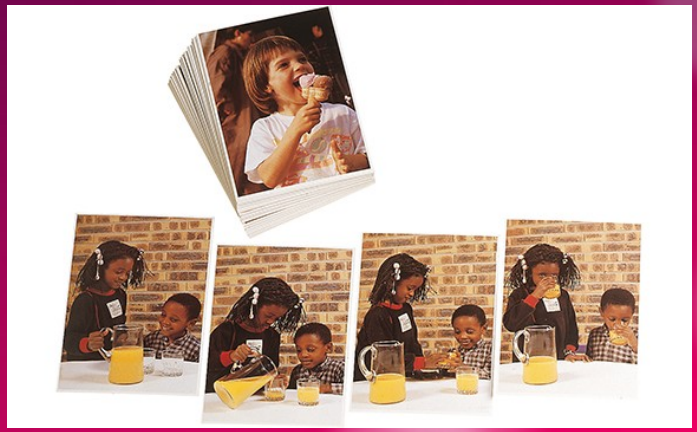
- Disk forensics
  - Including Web Browser Forensics
- Memory forensics
  - Passwords or hashes of passwords
  - Last activity
- Cloud forensics and Network forensics
  - Legal limit
- Encrypted content :
  - Search applications
  - Locate containers,
  - Ask credentials

Dealing with Identities

- Personal computers are shared
  - Often 1 shared account
  - Use of pseudos
- Linking Virtual and Real Identities
- Activity and Identity : Analysis of timeline
  - Reconstruction of timeline(s)
  - Detecting coherent/regular activity session
- « Written » Media Analysis (+OCR) :
  - Reports, Letters, Invoices, eMails

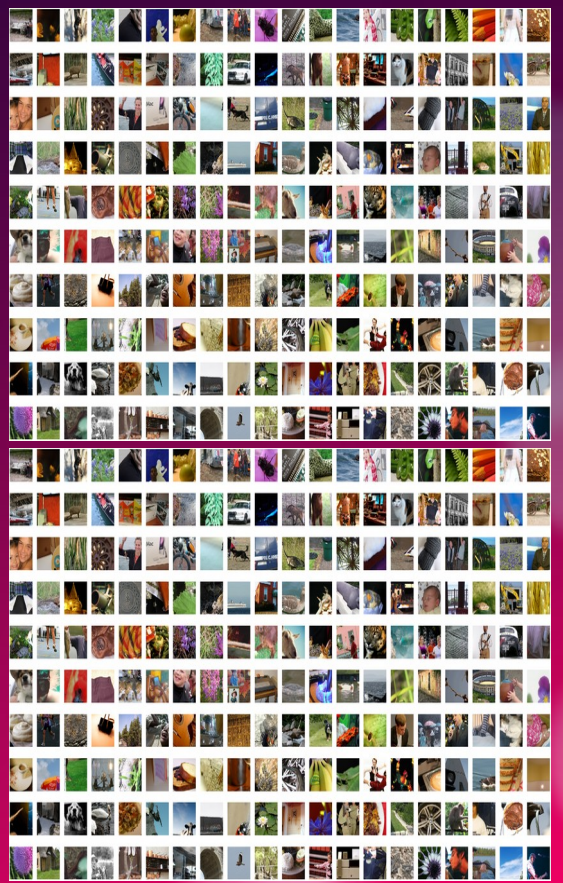
Dealing with images

- Dealing with tens of thousands of images per device
- The visible images (on the FS) :
  - Photos (personal/downloaded) : « burst mode »
  - Synthetic Images : Icons, logos, software images
- The photos and images embedded in documents or in archive
  - Office documents (doc, ppt, pdf, ...), : reports, presentation, mail…
  - Shared content : social networks
- The problem of near duplicates : PhotoDNA
  - Resized / cropped / multiple format
  - Annotated / altered / falsification



Dealing with images

- Thumbnail Databases :
  - To Accelerate Preview of Folder Content (OS based),
  - Digital Image Managers and Editors, Video Players, Media Players
- Software Cache :
  - Web Browser Forensics
- Lost images : File Carving Techniques
  - Deleted Files, Temporary files
  - Files in unallocated spaces
- The problem of duplicates :
  - Arise with File Carving & Cross device analysis



Dealing with images : the issues

- Filtering Techniques :
  - Databases of hashes of positive or negative
  - PhotoDNA
- The problem of duplicates or near duplicates :
  - Photos may be resized / cropped / edited / multiple format
- From exif metadata to content based analysis
  - Separate Photos from Synthetic Images
  - Authorship Attribution : Downloaded or Taken ? Group by camera model ?
  - Group by event : party, wedding, winter holidays,
  - Group by locations : indoor (in the same appartment), outdoor (garden, ...)
  - Face Gallery
  - Alteration/Falisication or Deep Fake Detection
  - Explicit content

Dealing with videos

- Time to analyse
  - Extract a frame every 10 seconds
- Issues :
  - Select Key Frames
  - Summarize

Dealing with text

- Often seen as Keyword search but…
  - Image of document => OCR
  - Keyword search => Encoding / Language
  - Propriatory format => Extractor
- Online discussions :
  - Protocol understanding
  - Locate received frames / sent frames
- Online conversation mining :
  - Tons of short messages with lots of contacts
  - specific langage (abbreviations, ...)
  - Need to combine with device event (camera / mic.)

Digital Forensics Issues

- Towards Forensics Labs ?
  - Initial Costs : Materials (computers, storage devices, write blockers, disk imagers, connectors)
  - Reccuring Costs : Software licences
- Improving Trust ?
  - Need to handle new technologies, usages, software, services, versions
  - Need for more independent evaluations, more open source tools
- Know-how & Sovereignty issue (Palantir-like corp.) ?
  - Default encryption / Hardware Security Module
  - Computer forensics companies
- Towards IA Tools ?
  - Text : Restricted to Keyword Search or Regexp Search ?
  - Image/Video : Towards Content-Based Search & Falsification Detection