



HAL
open science

Feedback of a Digital Forensics Expert - Part 1

Emmanuel Giguet

► **To cite this version:**

Emmanuel Giguet. Feedback of a Digital Forensics Expert - Part 1. Master. Master 2 Informatique Spécialité Cybersécurité - module Forensique, Université de Caen Normandie, France. 2023. hal-04454341

HAL Id: hal-04454341

<https://hal.science/hal-04454341>

Submitted on 13 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Feedback of a Digital Forensics Expert

Emmanuel Giguët
CNRS Researcher at GREYC Lab

emmanuel.giguët@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)

- 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

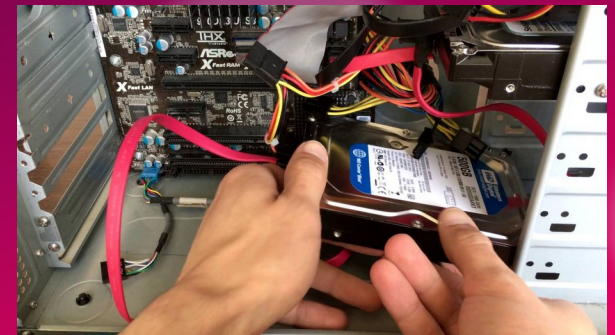
The playground (1/2)

- 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



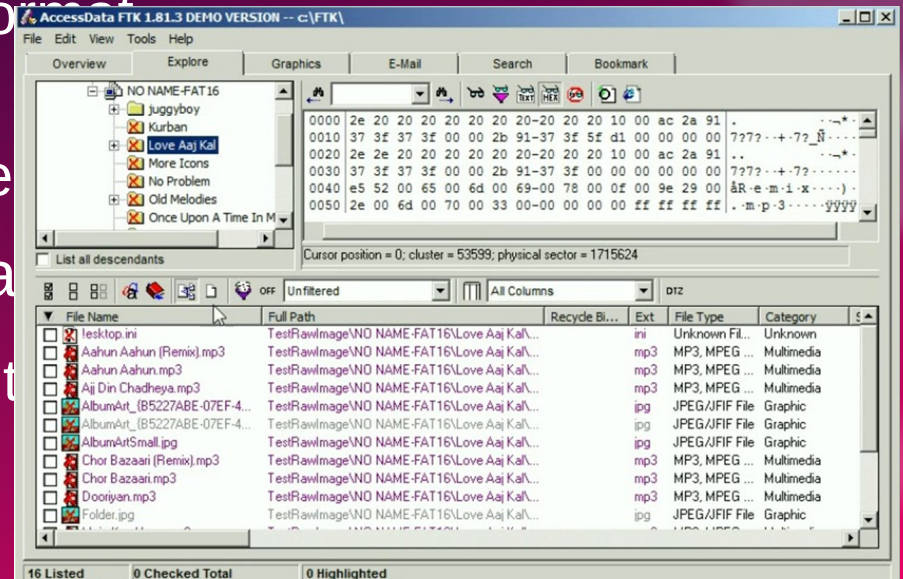
The playground (2/2)

- 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 formats (Security Exploded)
 - To recover deleted file / To do File Carving
- Not free, not open-source, not evaluated
 - 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



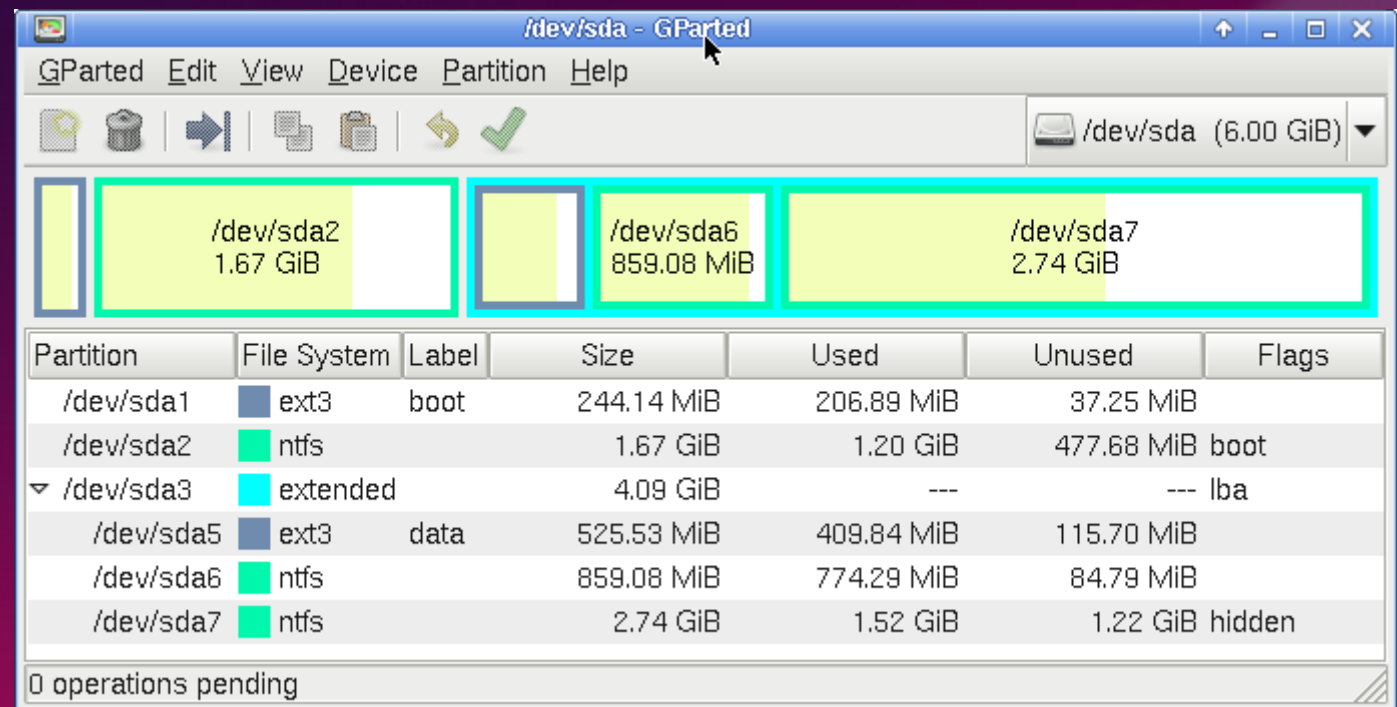
Forensic Data Acquisition

- 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



The screenshot shows the GParted application window titled "/dev/sda - GParted". The window displays a visual representation of the disk layout for /dev/sda (6.00 GiB). The partitions are shown as colored bars: /dev/sda1 (blue), /dev/sda2 (green), /dev/sda3 (cyan), /dev/sda5 (blue), /dev/sda6 (green), and /dev/sda7 (green). The /dev/sda2, /dev/sda6, and /dev/sda7 partitions are highlighted with a red border. Below the visual representation is a table with the following data:

Partition	File System	Label	Size	Used	Unused	Flags
/dev/sda1	ext3	boot	244.14 MiB	206.89 MiB	37.25 MiB	
/dev/sda2	ntfs		1.67 GiB	1.20 GiB	477.68 MiB	boot
▼ /dev/sda3	extended		4.09 GiB	---	---	lba
/dev/sda5	ext3	data	525.53 MiB	409.84 MiB	115.70 MiB	
/dev/sda6	ntfs		859.08 MiB	774.29 MiB	84.79 MiB	
/dev/sda7	ntfs		2.74 GiB	1.52 GiB	1.22 GiB	hidden

0 operations pending

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

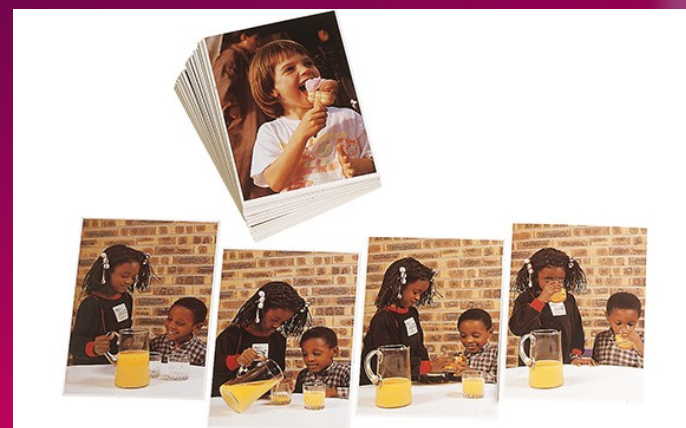
Dealing with Identities

- 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



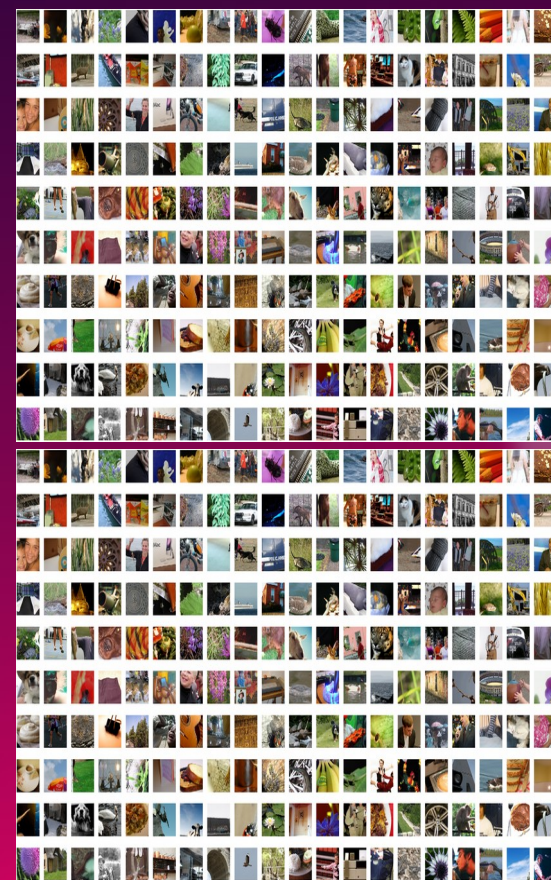
Dealing with images

- 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 apartment), outdoor (garden, ...)
 - Face Gallery
 - Alteration/Falsification or Deep Fake Detection
 - Explicit content

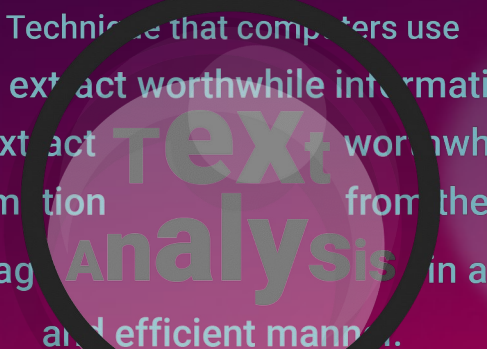
Dealing with videos

- 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
 - Proprietary format => Extractor
- Online discussions :
 - Protocol understanding
 - Locate received frames / sent frames
- Online conversation mining :
 - Tons of short messages with lots of contacts
 - specific language (abbreviations, ...)
 - Need to combine with device event (camera / mic.)



Techniques that computers use
to extract worthwhile information
to extract worthwhile information
information from the human
language in a smart
and efficient manner.

**Text
Analysis**

Digital Forensics Issues

- Towards Forensics Labs ?
 - Initial Costs : Materials (computers, storage devices, write blockers, disk imagers, connectors)
 - Reccurring 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



The main body of the page is a large, empty white space, suggesting a placeholder for content or a blank page.