# Computer Forensics and Cyber Crime Analysis

Alessandro Milani

October 9, 2024

# Contents

T	Le	egai	3									
1	Leg	al Introduction	4									
	1.1	GDPR aand alghoritm bias	4									
	1.2	Example - Lex Machina	4									
	1.3	Example - Compas	4									
2	Fou	Foundations of Digital Forensics 6										
	2.1	Intro and definitions	6									
		2.1.1 Digital and Electronic Evidence										
		2.1.2 What is digital Forensics	8									
		2.1.3 The "Big Five" of Digital Forensics (Council of Europe)	8									
	2.2	Digital Forensics Procedure										
		2.2.1 Identify the Suspect										
		2.2.2 Detecting and Seizing Digital Evidence										
		2.2.3 Validating Digital Evidence										
		2.2.4 Chain of Custody										
		2.2.5 Analysis of Digital Evidence										
		2.2.6 Presentation in Court										
	2.3	Privacy and Due Process Rights										
		2.3.1 Encryption										
		2.3.2 Case Law on Encryption										
		2.3.3 Mandatory Key Disclosure Laws										
		2.3.4 Remote Forensics										
		2.3.5 Cloud Computing										
II	$\mathbf{T}$	ech :	15									
3	Introduction 16											
9	3.1											
		<del>-</del>										
	3.2	Forensics History										
			16									
			16									
	9 P		16									
	3.3	Computer Forensics Definitions										
	3.4	1 1 ()	17									
		3.4.1 CF Q&A										
		3.4.2 CF Goals	17									

4			nology & revelant conceps	18
	4.1	Terms		18
		4.1.1	Digital evidence	18
		4.1.2	Chain of custody	18
		4.1.3	Data acuisition	18
		4.1.4	Hashing	19
		4.1.5	Write Blocker	19
		4.1.6	Forensic image	19
	4.2	Scenar	rios	19
	4.3	investi	igation phases	20
		4.3.1	Phases	20
		4.3.2	Identification	20
		4.3.3	Collection	21
		4.3.4	Acquisition	21
		4.3.5	Evaluation	21

Part I

Legal

# Chapter 1

# Legal Introduction

Beafore the technology was between us (classic telephon call), but with the advanced of the information technology, the technology is now about us (facial recognition, social media, etc).

Example: The IA act say taht is not possible utilize IA for real time facial recognition without an "important" reason.

This generation use technology also for be profiled by an alghoritm for varius scope and not only for comunicate.

The "technology was between us" was simpler and the only problem was to cehck if there is a conversetion and intercept it with a good quality. Now, we have the problem of quantity. If we need to analyze data from milions of people we can end to vaiolate fundamental rights and make mistaks (even with OSINT).

Survively theory: collective intelligence, the point is that if you have 10k persons say that the restuarant is good and only 10 that say it's a froud, the collective intelligence say that the majority have right. In Forensics this can not be applied because you need to be 100% secure of what you have. (if in a trial you have only the 1% that a person can't be guily you need to be in favor of him) [find better term]

In the technology in Us, and the advance of IA is important that the law split want is uman and what is not (like be transparent when a content is AI gen and when not)

"Tesla case" when there is an incident it need to understand in the percent of error that is from tesla and the percent from the partners.

# 1.1 GDPR aand alghoritm bias

Art.22 of GDPR, say that you always need to have human in the decision process

# 1.2 Example - Lex Machina

Tool that analyze all the legal case from a giurisdiction (like France) and classify all the case in different categorys. So if you have a case X in Paris with judge Y, you have 60% possibility to win. If the Attorny ins Z, the probability is 80%.

# 1.3 Example - Compas

alghoritm that help the jugde defice is the person can commit other crime or not and so decide it need to stay in jail or get a reduce in the sentence (sconto della pena)

- A False positive in digital forensics can change people's lives.

# Chapter 2

# Foundations of Digital Forensics

#### 2.1 Intro and definitions

#### 2.1.1 Digital and Electronic Evidence

By the Scientific Working Group on Digital Evidence (**SWGDE**) a definition of, **digital evidence** is "any information of evidential value whether memorized or sent in a digital format". It's used by the **Council of Europe** 

Another definition come from the **Eoghan Casey - 2004** that define a digital evidence or electronic evidence as "any probative informationstored or transmitted in digital form that a party to a court case may use at trial". It's more related to the juridical part.

A last definition of **Electronic evidence** is information generated, stored or transmitted using electronic devices that may be relied upon in court, defined by the **Council of Europe - 2013**.

#### For the exam, the first dfinition is the more important

So. in general, way we can say that a digital/electronic evidence need to be:

- invisible to the untrained eye
- Need to be **interpreted** by a specialist
- It may be altered or destroyed through normal use
- It can be **copied** without limits

#### Legal Requirements

The main characteristics that a Digital/Electronic Evidence need to have to be accepted in a trial are:

- Admissability: it need to be compliant with law and best practices.

  What can be seen is not what can be admissed in court (ex. if the italian police enter in a laptop, can only "wiretapping", by enabling mic and camera, and can't use other information like email or files in court )
- Authenticity: avoid any digital evidence tampering

- Reliability and believability: readily understandable for a judge. If a judge not understand the evidence can ignore it
- Proportionality: respect fundamental rights of parties affected by the measure

#### Find a digital evidence

A digital evicence can be hidden in different place and a criminal usully use some classical ways (not the cloud becasue the access is easy form a pocile force) like hidden folder, usb, extrnal memory etc.. can be hidden everywhere

#### Categories

Three main types of digital evidence

- Created by human: digital data result of an action taken by an human
  - Human to Human: Like an email
  - Human to PC: like a word document
- Created indipendently by the computer: data that are result of the processing of data by an algorithm and without human intervention
- Created by both human and the computer: somethings like a spreadsheet where the data are entered by the human, and the result is worked out by the computer

#### This case is not import for the exam

Julie Amero is a supply teacher at Kelly School in Norwich, Connecticut who was found guilty of showing pornography to children under the age of 16 for some popup that appear during a lesson

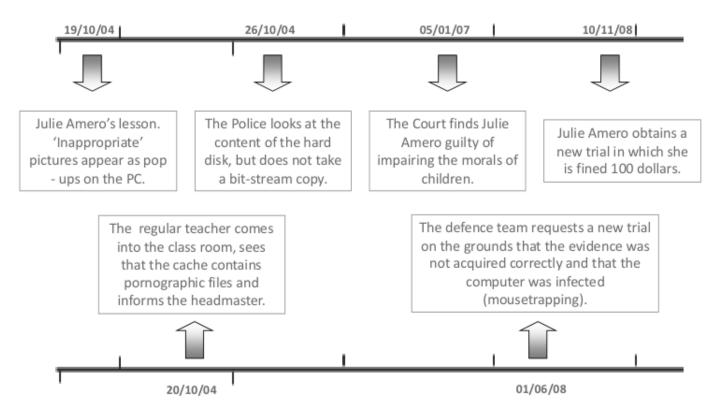


Figure 2.1: Timeline of the case

# 2.1.2 What is digital Forensics

**Digital Forensics** is get hold of evidence without modifing the IT system in which that evidence is found, ensure that the evicence acuired in another medium is identical to the original and analyse data without modifyin it.

# 2.1.3 The "Big Five" of Digital Forensics (Council of Europe)

- Data Integrity: No action taken should change electronic devices or media, which may subsequently be relied upon in court.
- Chain of custody: An *audit trail* of all actions taken when handling electronic evidence should be created and preserved
- Specialist Support: If investigations involving search and seizure of electronic evidence it may be necessary to consult *external specialists*.
- Appropriate Training: First responders must be appropriately trained to be able to search for and seize electronic evidence if no experts are available at the scene

• Legality: The person and agency in charge of the case are responsible for ensuring that the law and the above listed principles are adhered to.

# 2.2 Digital Forensics Procedure

Six phase of digital forensics procedure:

## 2.2.1 Identify the Suspect

There are 3 main phase for identify the suspect:

- Osint and Socmint: Very usefull for collect information reguarding criminal (even mafia ones), from social media, and other public sources.
- Data Retension Directive in EU: The investigator uses the Court System to compel the ISP to reveal a physical location that corresponds likely to the source of Network (IP Address)
- Multiple User ID or multiple Ips over time, open Wi-Fi, Proxy, Botnet: Under a warrant (depending from the Jurisdiction) the location is searched and any computer or other device is seized

#### **Data Retension**

With the Directive 2006/24/EC, the EU member states are required to store data for a period of **6 to 24 months** (but can change from state to state). The data stored are gerally call detail records (CDR) of telephony and internet traffic and location data (IPDRs).

So evert single country and ISP have different data retention policy, and this can be a problem for the investigator, but from a privacy point of view a short time or null data retention is better.

**Transparency report:** Every year the ISP need to publish a transparency report where they show the number of request of data retention and the number of request that they have accepted.

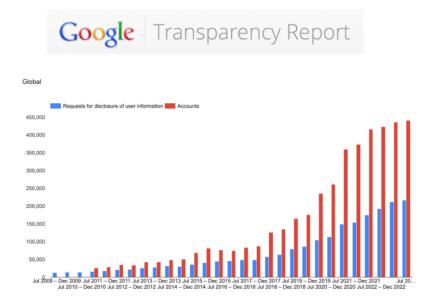


Figure 2.2: Timeline of Google transparency report

#### Identify the suspect

it's rigth use face recognision to identify a suspect? For the moment if the situation is critical is possible utiliza A live facial recognision. AI act also regulate the use of facial recognision.

# 2.2.2 Detecting and Seizing Digital Evidence

The seize of digital evicences has to respet two foundamental rules: **bit-Stram Copy** and **Hash Function**. (definitions are alredy knowed)

#### Where and how is the digital/electronic evidence hosted?

The digital evicence can be in the suspected PC or in a third party server.

In the first case, we need to mange the encrition of the data, and it's possible get a Key Mandatory Law.

In the case of evicende in a server, is needed a collaboration from the ISP/Telco/Banck, and so need there is Jurisdiction problem.

#### The role of third paries during digital investigation

A third party can give a lot of usefull information.

For example and **ISP** Could reveal from which place the email was sent, the **Mail Account Provider** could reveal from which places the email account was accessed and a **Credit Card Company** could reveal where the goods bought with a cloned credit card were delivered

#### 2.2.3 Validating Digital Evidence

There is some tool that help to valdiate online digital evidence.

These kind of tool are usefull during OSINT bacasue they permit to collect data (like a story, a reel), that are not sure remain online, in a proper way for a court.

#### 2.2.4 Chain of Custody

Digital storage media last less than analogue media and devices to read such media last even less. For example a LaserDisc last only 15 years, where there are books form thr 1086 (domesday book). So there at the moment, for trial, there are a lot of hard drive keeped in proprer way to avoid the data loss. It's a real mess.

#### 2.2.5 Analysis of Digital Evidence

Start after the sieze of suspect's device, and need to be performed besided a precise chain of custody. To perform the analysis are usually used some automatic tools, but in the recent time are used also some AI tools but only for post analysis and not for prediction policies (limitation imposed by the AI act) for example AI can not be used for kidnapping cases because the crime is in progress and not "finished".

- Text searches: aimed at scanning files, directories and even entire file systems for specific text terms (generative AI can be used for summarizing and analyzing documents, but it's not very precise, plus alucination)
- Image searches: aimed at identifying image files in various formats, and at generating still frames of digitally stored video footage. Mainly analysis of child pornography that can be lead also to false positive (like a video of a mum and child in a bath)
- Data recovery: aimed at recovering all files stored on mass memory units, including deleted or damaged data. Destroy data can also be a crime (even if there are some backups), based on the intention of the crime (like delete file to hide evidence)
- Data discovery: targeted at accessing hidden, encrypted or otherwise protected data
- Data carving: focused on reconstructing damaged files by retrieving portions of their content
- Metadata recovery and identification: this digital forensic tool is particularly useful for retracing the timeline of web accesses and file changes

Some other problem with the use of **AI** for prediction of crime are: the bias of the AI and possible consequences for privacy and **social control** by not very democratic government.

#### Two Italian issues

Repeatable or Unrepeatable forensics analysis: Repetable when you can do a bit-stream copy of the data and also give one copy to the defence to do the same analysis, and more in general i can repeat analysis again and again (when i want).

In a **non-repetable** analysis, we need do live forensics activity and ofter occure with mobile phone, where there isn't the possibility to do a bit-stream copy of the data. In the live forensics i also need the presence of the attorney or the defender when i do the analysis to make it admissable in court.

Open Source or Closed Source: Open source can be more transparent

#### 2.2.6 Presentation in Court

The presentation of digital evidence findings is a **crucial stage** for prosecutors, judges, and lawyers (the evidence need to be presented in a way that the judge can understand it otherwire he can ignor it). The outcome of the trial relies not only on the results of the investigation but also on the **clarity and comprehensibility** of the report provided.

#### **Operational Recommendations:**

- Presence of an index: The report should include a clear index for easy navigation through the document.
- Glossary and Reference Notes: If technical terms are used, a glossary and reference notes should be provided to ensure that all parties understand the terminology (judge and lawyers are not IT experts).
- Timeline Table and Flow Charts: A timeline table or flow charts should be included to visually represent the sequence of events and digital activities.
- Presentation Slides with Photos: Visual aids, such as presentation slides with photos, help in simplifying complex technical details.
- Video Recording: Where applicable, video recordings of the operations carried out during the investigation can provide further clarity.

Presentation in Court of the Digital Evidence Findings: Murtha Case

# 2.3 Privacy and Due Process Rights

## 2.3.1 Encryption

Encryption can be used to hide the fact that encrypted messages are exchanged and used by criminals can lead to difficulties collecting the necessary evidence

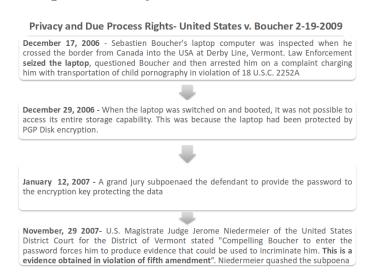


Figure 2.3: Case correlated with the use of encryption

#### 2.3.2 Case Law on Encryption

Anther the previus case, some state are starting to create law about "Mandatory Key Disclosure" that force the suspect to give the key/password to decrypt the data. (some are Australia, Belgium France etc...)

#### 2.3.3 Mandatory Key Disclosure Laws

These case of legislative instrument doesn't work fow two main reason:

- technical reason: An expert could always find a way yo hide a file
- Possible violation of European Convention on Human Rights: Article 6 Everyone charged with a criminal offence shall be presumed innocent until proved guilty according to law

#### 2.3.4 Remote Forensics

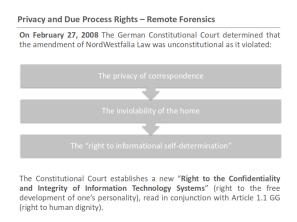


Figure 2.4: Case correlated to remote forensics

# 2.3.5 Cloud Computing

Cloud computing services face two key legal challenges: **Jurisdiction** and **Privacy**.

**Jurisdiction** The "loss of location" of digital evidence in the cloud introduces significant jurisdictional issues. In a cloud environment, the question arises: are the documents governed by the laws of the state in which they are physically located, the location of the company possessing them, or the laws of the state where the individual resides?

Over the last few years, various legal frameworks and approaches have been proposed to address this complex issue, but it remains an area of ongoing debate.

**Privacy** Cloud computing introduces several privacy concerns, including:

- Lack of Control: Cloud clients may no longer maintain exclusive control over their data, limiting their ability to implement necessary technical and organizational measures to comply with Data Protection Laws.
- **Absence of Transparency:** Cloud providers may not provide sufficient information regarding how data is processed, leading to significant risks in terms of compliance with data protection regulations.

#### **Privacy**

#### Jurisdiction

In addressing the "loss of location" issue within the realm of cloud computing, we have four possible legal principles that can be applied:

- Territorial Principle: The court in the jurisdiction where the data is physically located has authority. Jurisdiction is determined based on the geographical location of the data.
- Nationality Principle: The nationality of the perpetrator is used to establish criminal jurisdiction. The legal framework of the country of the individual committing the crime applies.
- Flag Principle: This principle applies to crimes committed on ships, aircraft, and spacecraft. They are subject to the jurisdiction of the country whose flag the vehicle flies under.
- Power of Disposal Approach: This approach focuses on who has control over the data. A regulation based on this would enable law enforcement to access a suspect's data in the cloud, regardless of its physical location, by considering the individual or entity with power over the data.

# Part II Tech

# Chapter 3

# Introduction

# 3.1 Topics

#### • Forensics Analysis

use of logic and meaningful knowledge and methodological approach to legal problems and criminal investigation.

#### • Computer Forensics

Collection, preservation and analysis of digital evidence (inside file system, email, cloud account etc...) to support investigation and legal proceedings

# 3.2 Forensics History

#### 3.2.1 Ancient Times

Forensic science dates back to **Babylon** (1900 BC) where fingerprints were used for identification, and China (1248 AC) with forensic pathology. In the UK (1835), bullet comparison solved a case, and by 1892, the first murder was solved using fingerprints.

#### 3.2.2 Modern Times

Forensic standards grew in police departments, with the first crime lab in 1923. DNA fingerprinting began in the 1950s, and DNA profiling was developed by 1985. AFIS systems emerged in the late 1980s. Today, AI, toxicology, and digital forensics are key areas of innovation.

## 3.2.3 Digital Field

#### Early Times

In 1989, Robert Morris was convicted under the Computer Fraud and Abuse Act, marking the first use of computer logs in forensics. That year, IACIS was founded, followed by IOCE in 1995 to share digital forensic practices.

#### Recent Times

- 1990: Forensic tools like EnCase emerged
- 2000: Digital forensics became widespread in law enforcement

- 2010: Growth of cloud and mobile forensics, automation, and machine learning
- 2020: Advances in crypto, blockchain, and AI improve digital forensics

# 3.3 Computer Forensics Definitions

US\_CERT: The discipline that combines elements of law and computer science to collect and analyze data from computer systems, networks, wireless communications, and storage devices in a way that is admissible as evidence in a court of law.

A. Ghirardini - Computer Forensics: The discipline whose goal is preservation, identification, analysis of information system to the aim of identification of evidences during investigation activities.

**NIST glossary:** The application of computer science and investigative procedures involving the examination of digital evidence - following proper search authority, chain of custody, validation with mathematics, use of validated tools, repeatability, reporting, and possibly expert testimony.

# 3.4 CF purpose(s)

#### 3.4.1 CF Q&A

During an investigation, digital forensics need to analysis data to answer some key questions:

• What happened?

• Where did it take place?

• Who was involved?

• Why did it take place?

• When did it take place?

• How did an incident occur?

The answers to these questions are essential for support legal proceedings and mitigate possibility of future incidents with a preventive approach.

#### **3.4.2** CF Goals

The goals of computer forensics (CF) are multifaceted and aim to provide a comprehensive understanding of digital incidents.

Firstly, CF seeks to retrieve what has been the input, such as what has been typed. It also aims to determine the actions performed, for example, what programs have been run and what peripherals have been connected.

Additionally, CF involves analyzing used files to understand what modifications have been done and when these modifications occurred (and the information from an OS are not enough, because are an abstraction managed by the file system  $\rightarrow$  needed bit analysis (like for ereased data)).

Another critical goal is to identify the damage done, such as what data have been erased. In essence, the overarching goal of CF is to **gain a technical comprehension of what happened during the incident** (from a technical point of view).

# Chapter 4

# CF Terminology & revelant conceps

#### 4.1 Terms

#### 4.1.1 Digital evidence

Technical definition of Digital evidence is very similar to the legal ones:

Data stored or transmitted in digital form that can be used in court.

The conrnerstones of digital forensics are the different levels of **abstraction**, requires **interpretation**, are **fragile**, may be **voluminous** and the difficulty to discover **connection** between data and reality (the connection need to be done beafore entering the court).

Digital evidence also requires a deep technical understanding of the possible types of data (files, emails, logs, metadata) and the legal requirements for each of one to collect and preserve it. (To make all of this effective, knowledge of file systems, network protocols and encryption are essential)

## 4.1.2 Chain of custody

Documented and **unbroken** process of handling evidence from the time it is collected until it is presented in court

This procedure is essential to ensure the integrity of the evidence and to avoid that them to be tampered or accessed by unauthorized people.

Keep the chain of custody requires knowledge about how to document evidence collection, storage, and access (logging procedures, secure storage, legal procedures).

If the chain of custody is **broken**, the evidence may be considered **inadmissible in court** (so is needed know the regulation of the state to decide how to manage it).

#### 4.1.3 Data acuisition

The process of collecting digital evidence from devices without altering or damaging the original data.

One of the biggest problem fo the management of digital evidence, because it's needed to be performed on hostile systems (that can be infected, compromised, have a malware or system to avoid copy like edited system call for make other program to fail). in a not coltrolled envorment like a crime scene. So are needed knowledge of disk imaging and live data capture in order to not alter what's going on on the suspect system. Are also required expertise in forensics acquisition, analysis tools (like FTK Imager, EnCase) and knowledge of file systems, write-blockers, and hashing (crucial for ensuring integrity).

#### 4.1.4 Hashing

The process of converting data into a fixed-length string of bits, which represents the data uniquely

It's used in the chain of custody for ensure the integrity of the digital evidence and so verity that a file has not been altered.

Require understanding of hashing algorithms (strengths and weaknesses, e.g. MD5 collision), formats (hex, base64 etc...) (if wrong formats are sued, the chain of custody is broken and each information gathered from that point is considered not valid) and expertise in hashing tools (sha256sum, hashdeep, FTK imager, Autopsy).

Have to be used any time an evidence is "managed" (copied, moved)

#### 4.1.5 Write Blocker

Hardware or software tool used to prevent any data from being written to a storage device during analysis, preserving the original data content

To be operated, require understanding of how write-blocking devices work and how they can be implemented in forensic procedures.

It's essential for the legally defensible acquisition.

# 4.1.6 Forensic image

A bit-by-bit copy of digital media, including deleted files and data in slack space, which is an exact replica of the original device

The goal of a forensic image is to preserve the original evidence and avoid the modification of the original data.

To be performed in a correct way, requires understanding of mechanisms to copy information in digital devices (file system knowledge and behavior) and familiarity with with bit-by-bit copy tools (DD, FTK Imager, EnCase, Guymager).

As the hashing, it's need to be used any time an evicene is "managed" (copied, moved)

# 4.2 Scenarios

There are some possible scenario that a computer forensic investigator can face:

- Internet abuse from employee
- computer-aided frauds
- Data unauthorized manupulation (theft or destruction)
- Computer/network manage assessment
- ... any other case that include digital evidence

# 4.3 investigation phases

A Computer forensics investigatos usually follow standard phases that guide him. There are different standards like: NIST family, ACPO guidelines (UK), ISO/IEC 27042, SWDGE.

#### 4.3.1 Phases

- identification: When the investigator come for the first time to the crime scene and need to identify potential source of relevant digital evicences.
- collection: The letteral pick up of the evidence (like a computer or a smartphone) or a remote taking possession of the evidence (like for a remote server) and its connection (e.g. network or physical, like USB disk).
  - It's splitted from acquisition because it's a critical phase where the evidence can be altered and lost utility for the investigation (es. data corruption, lost of metadata etc...)
- acquisition: Electronically retrieving data by running various CF tools and software suites
- evaluation: Evaluating the data recovered to determine if and how it could be used against the suspect (e.g. for prosecution in court)
- **presentation:** Presenting the evidence discovered in a manner which is suitable for lawyers, non-technical staff/management and the law (and internal rules)

#### 4.3.2 Identification

During the identification phase is important **recognize** all the **relevant data sources** before any acquisition, even if no physical present, like data in the cloud

A imple **list of example** are: hard drives (HDD/SSD), memory (RAM), mobile devices (smartphones, tablets), cloud storage, network traffic, removable media (USB drives, DVDs), IoT devices and embedded systems (like smart washing machines)

For identify these sourcer, the investigator can perform some actions, like:

- Perform an initial survey of the scene (physical or network environment)
- Identify key devices and data locations (local storage, remote servers, cloud services)
- Check for connected devices, including peripherals like printers, removable media, or networkattached devices
- Map all potential data sources using network topology diagrams or asset inventories

A particular aspect that need to be considered is the possible present of "ephemeral" storage or data, like cloud syncing, hidden secttor, tmp, dat in ram etc...

#### 4.3.3 Collection

During the collection, the focus is on gathering evidence from identified data sources while ensuring the preservation of its integrity. An important key point is the implementation of methods that **minimize** the risk of evidence tampering or data loss.

For enforce this key point, it's important **isolate devices** to prevent them from being tampered with remotely (e.g., disconnect them from the network), use devices to **block external communication** for mobile or wireless devices and use network isolation tools for virtual and cloud environments to prevent remote access (like use a virtual private cloud).

A particular note is for the managment of live systems where is needed ensure evidence integrity while maintaining system uptime (so not shoutting down the system for avoid the loss of volatile data).

Create a **detailed record** of the condition and state of the evidence

- take photographs of the devices in situ, including connected peripherals and the physical state
- record serial numbers, device models, and any other identifiable information
- document the scene, noting which devices were running, whether screens were active o locked, and any other visible indicators

hint: complete documentation is crucial to prevent legal challenges regarding the integrity of the evidence.

Beafore procede to the acquisition, is needed to ensuring no alteration will take place, so do somethings like enable write blockers for physical storage devices, disable connection and sinking. particulary complex is maintain integrity on live systems (e.g., using remote collection methods that minimize data alteration risks)

## 4.3.4 Acquisition

The act of performing a forensic copy (so a bit-by-bit copy) of the original data with the goal of ensure that the acquired data is a faithful replica of the source so to maintaining data integrity.

There 2 two main acquisition methods:

- Static: When the system is powered down, it's se most commod method for acquiring data from hard drives and external memory
- Live: The system is running and it's needed to deal with volatile data like RAM, network connections, or running processes.

#### 4.3.5 Evaluation