Understanding Digital Forensics

October 25, 2019; Wilmington, OH

John C. Ellis, Jr.

Presentation Outline

1. There are now over 7.2 billion mobile devices in the world, and in the United States, 92% of adults own a cell phone. These devices, in large part, have resulted in a significant increase in discovery in criminal cases. This presentation seeks to explain the process for acquiring data from devices, and how to integrate them into your case.

2. **Resource Materials**:
    a. Lee Reiber, *Mobile Forensic Investigations: A Guide to Evidence Collection, Analysis, and Presentation* (McGraw-Hill Education 2019);
    b. NIST Special Publication 800-101 (Revision 1), Guidelines on Mobile Device Forensics (May 2014);
    c. NIST Special Publication 800-124 (Revision 1), Guidelines for Managing the Security of Mobile Devices in the Enterprise (June 2013);
    d. Digital Forensics for Legal Professionals: Understanding Digital Evidence from the Warrant to the Courtroom 1st Edition, by Larry Daniel and Lars Daniel, ISBN-13: 978-1597496438.
    e. Criminal e-Discovery: A Pocket Guide for Judges – Federal Judicial Center, http://www.fjc.gov/public/pdf.nsf/lookup/Criminal-e-Discovery.pdf/$file/Criminal-e-Discovery.pdf.
    f. NIJ- Electronic Crime Scene Investigation: A Guide for First Responders, Second Edition, https://www.ncjrs.gov/pdffiles1/nij/219941.pdf.
    g. BJA- A Simplified Guide to Digital Evidence, http://www.crime-scene-investigator.net/SimplifiedGuideDigitalEvidence.pdf.

h. Digital Evidence Guide for First Responders Digital Evidence Guide for First Responders, http://www.iacpcybercenter.org/wp-content/uploads/2015/04/digitalevidence-booklet-051215.pdf.
i. Forensic Examination of Digital Evidence: A Guide for Law Enforcement, https://www.ncjrs.gov/pdffiles1/nij/199408.pdf.

3. **Digital Data—**Computers and Cellphones store data in the binary format.

   a. **Binary**—a number system that uses only 2 possible values: 0 and 1.

   b. **Decimal**—a number system that uses 10 unique symbols to represent a particular value: 0-9.

   c. **Hexadecimal**—a number system that uses 16 unique symbols to represent a particular value: 0-9 and A-F.

   d. **ASCII**—American Standard Code for Information Interchange is how digital devices translate binary, decimal and hexadecimal into a readable format on a digital device.

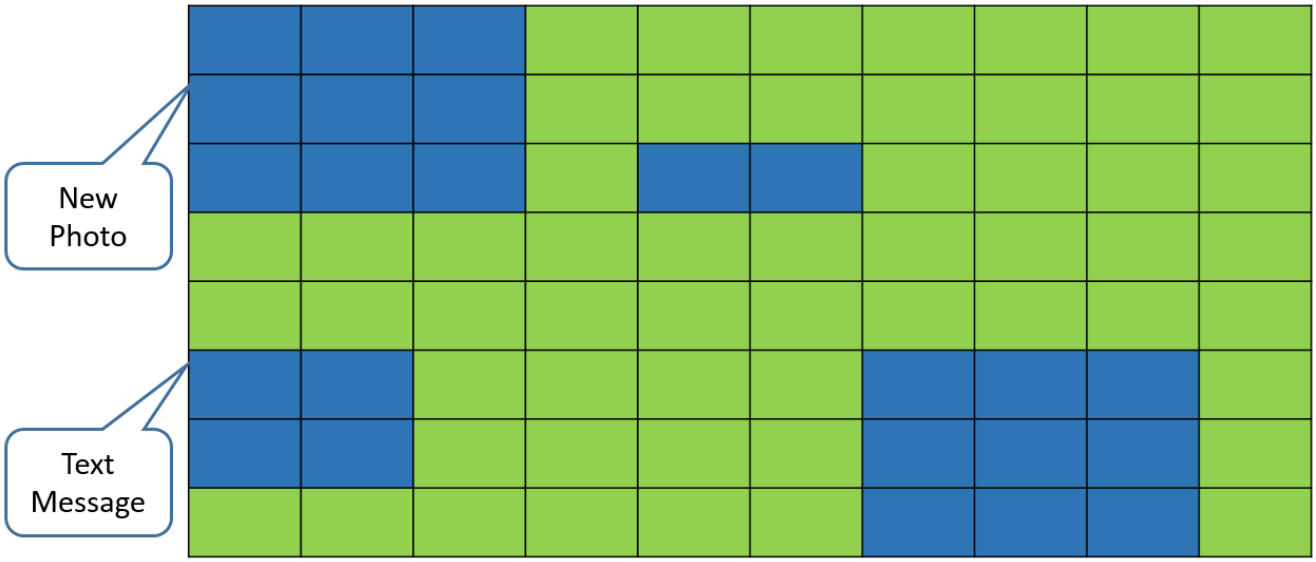
   e. **ASCII Table**—the following table is an example of how letters (a-g) and numbers (1-7) are represented by binary, decimal and hexadecimal.

ASCII	Binary	Decimal	Hexadecimal	ASCII	Binary	Decimal	Hexadecimal
A	01100001	97	61	1	00110001	49	31
B	01100010	98	62	2	00110010	50	32
C	01100011	99	63	3	00110011	51	33
D	01100100	100	64	4	00110100	52	34
E	01100101	101	65	5	00110101	53	35
F	01100110	102	66	6	00110110	54	36
G	01100111	103	67	7	00110111	55	37

f. **File Size Explained**: A bit is the smallest measurement: either a 1 or a 0. A byte contains 8 bits. In the following example, "Justice" has a file size of 9 bytes (8 letters and one space). The following chart shows how "Justice" is expressed in binary, decimal and hexadecimal.

ASCII	Binary	Decimal	Hexadecimal	File Size
J	01001010	74	4A	
u	01110101	117	75	
s	01110011	115	73	
t	01110100	116	74	7 bytes
i	01101001	105	69	
c	01100011	99	63	
e	01100101	101	65	

4. **Data Storage**—Many digital devices store data in packets. When data is stored on a device, it is in **allocated space** (in other words, that space is being used). When space is available to save new data, it is referred to as **unallocated space.** Unallocated space includes spaces where data has been deleted. In the following example, the green represents unallocated space and the blue allocated space.



5. **Metadata**— Information contained within digital data describing information about other data, including:
    a. The time and date a photograph was created;
    b. The date an application was downloaded;
    c. The last time a file was accessed; and
    d. The type of device used to capture the data (*e.g.,* iPhone 6s, computer, etc.).
6. **Mobile Devices (or Cell Phones)—**Are now "such a pervasive and insistent part of daily life that the proverbial visitor from Mars might conclude they were an important feature of human anatomy." *Riley v. California*, 134 S. Ct. 2473, 2484 (2014). But, as NIST explains in Special Publication 800-124, it is not always easy to define the term "mobile device" because the features are constantly changing. The following hardware and software characteristics collectively define the baseline of mobile devices:
    a. A small form factor (*i.e.,* is portable);
    b. At least one wireless network interface for network access (data communications);
    c. The wireless network interface uses Wi-Fi, cellular networking, or other technologies that connect the mobile device to network infrastructures with connectivity to the Internet or other data networks;
    d. Local built-in (non-removable) data storage;
    e. An operating system that is not a full-fledged desktop or laptop operating system;
    f. Applications available through multiple methods (provided with the mobile device, accessed through web browser, acquired and installed from third parties).

7. **Mobile Forensics**—Is the process of recovering, analyzing, and interpreting data related to a mobile device. It has four primary steps: (1) collection and extraction; (2) analysis; (3) reporting; and (4) verification.
    a. **Collection and Extraction**—includes the seizure of the device and the extraction (or copying) of the data in it.
        i. **Collection**—the condition of the device at the time of seizure can impact information later retrieved (*e.g.,* new data can be saved deleting old data; all the data on the device can be deleted remotely, etc). As such, the following questions should be considered:

1. Was the device on or off at the time of seizure;
2. If the device was on, was it isolated from all networks; and
3. If the device was isolated, what techniques were used (*e.g.,* placed in a Faraday bag, powered off, etc.).

ii. **Extraction**—the process of copying data from a mobile device. The amount of data received from the device depends on the type of extraction that is performed. Generally, there are three types of extractions: logical, file system, and physical. It is not always possible to perform all three extractions on every device. When possible, however, every extraction that is possible should be performed.

1. **Logical Extraction**— acquires data where it logically resides on a device using the original application (*i.e.,* iTunes for an iPhone). One benefit of a logical extraction is that the types of data being extracted can be limited (*e.g.,* contacts, SMS messages, calendar, etc.). This means that law enforcement can, when a logical extraction is available, limit the type of data it will seize from the device.
2. **File System Extraction**—acquires data by relying on software to access the device's memory including unallocated space. Data is only extracted from the device's file system—which controls the flow of data on mobile devices.
3. **Physical Extraction**—acquires data from both allocated and unallocated space, including logical data, deleted data, and hidden data.

4. **Cellebrite Extractions**—according to Cellebrite, the following information can generally be acquired during each extraction:

LOGICAL	FILE SYSTEM	PHYSICAL
SMS	SMS	SMS
Contacts	Contacts	Contacts
Call Logs	Call Logs	Call logs
Media	Media	Media
Audio	Audio	Audio
	Files	Files
	Hidden Files	Hidden Files
		Deleted Files

b. **Analysis**—the process of reviewing and understanding data with the goal of discovering useful information.

c. **Reporting**—the process of taking the useful data and reducing it to a readable report (e.g., a .PDF document).

d. **Verification**—the process of ensuring that the data that appears after the forensic process is accurate.

8. **Types of Information Contained in Cellular Phones**—Cellular phones contain a significant amount of information, including text messages, photographs, internet searches, and location data.
   a. **Text Messages**—generally refer to both SMS (Short Message Service) and apps such as Facebook Messenger, WhatsApp, and iMessage. In 2014, 561 billion SMS messages were sent worldwide (or approximately 18.7 billion texts per day). As for apps, by one estimate, 60 billion messages are sent per day from Facebook Messenger and WhatsApp, and Apple claims that, at times, 200,000 iMessages are sent per second. Much of this information can be found on the device.

b. **Photographs**—it is estimated that 1.3 trillion photographs will be taken worldwide this year. Of those, nearly 80% will come from a cellular phone. And these photographs only account for a portion of the photos stored on mobile phones. In addition, cellular phones include photographs saved from the internet (both intentionally and unintentionally). Cellular phones, like computers, often store photographs and content from websites—referred to as "cache data."

c. **Internet Searches**—people use their cell phones to search the internet. Often these searches, as well as websites that are visited on mobile devices, can be extracted during the mobile forensic process.

d. **Location Data**—cellular phones store a significant amount of location data. This information comes from four primary areas: (1) GPS; (2) WiFi hotspots; (3) cell sites; and (4) metadata in media files (*i.e.,* information hidden within photographs and videos). Some of the data remains stored in the device and can be obtained during the mobile forensic process.
    i. **GPS**—mobile devices use GPS for a variety of reasons, including providing the device's location and directions. Generally, there are two types of GPS data stored: (1) GPS fixes—where the device was at a given time; and (2) journeys—multiple GPS locations from a trip.

    ii. **WiFi**—when WiFi is enabled on a mobile device, the device is constantly looking for WiFi networks and storing this information. Each WiFi router has a unique number assigned to it—called a MAC Address. When the physical address of the WiFi router is known, you can use it to determine the historical location of a cellular phone.

    iii. **Cell Site**—mobile devices are designed to connect to cellular networks. Even when a device is not in use, it is constantly looking to determine nearby cellular networks. It does this by looking for signals from nearby cell sites. The sites within range are stored, recorded, and sometimes can be found during the forensic process.

iv. **Media**—photographs and videos captured with a mobile device may include the location data (called geotagging). Additionally, there is usually data about the time and date the phone or video was captured. Taken together, this information can show when and where a person was at a given time.

9. **Hash Function and Value**—A **hash function** is the process (or algorithm) that produces a numeric value that uniquely identifies digital data. The product of this function is called a **hash value**, (also sometimes called the digital fingerprint of the data). Hash values can be compared to two sets of data to confirm authenticity. There are several hash functions, the most popular are MD5 and SHA.

   The following example highlights how hash values operate. For illustration purposes, this example uses the CRC32 (which uses 32 bits whereas SHA-256 uses 256 bits). Here, the hash for the following these three sentences is very different, even when only small changes are made.

Input	Hash Function	Hash Value
John	→	914846a8
John was a Federal Defender	→	9862989c
John was a federal Defender	→	b6c1aa66

   With images (both photographs and videos), the Hash Function looks to the image itself, not the file name. In this example, if two images are the same they will have the same Hash Value, even if the file name is different.

Image	File Name	Hash Function	Hash Value
	happyface.jpg	→	76d6f21d
	sadface.jpg	→	76d6f21d

10. Tools for the Digital World

- **a.** FTK Imager
- **b.** Cellebrite Reader
- **c.** Phone Report Searching
  - **i.** Windows search
  - **ii.** dtSearch
  - **iii.** Excel
  - **iv.** CaseMap

11. Social Media

a. Common applications
   i. Facebook
   ii. Instagram
   iii. YouTube
   iv. Twitter
   v. LinkedIn

b. Other applications and tools to be aware of
   i. SnapChat
   ii. Tumblr
   iii. Periscope
   iv. Reddit
   v. Tinder

c. Messaging Apps
   i. Kik
   ii. Facebook Messenger
   iii. Line
   iv. Viber
   v. WhatsApp
   vi. TikTok (formerly Musical.ly)
   vii. Telegram

d. How to Capture Information
   i. Tools for web page preservation
      1. Adobe Acrobat Pro DC
      2. CamStudio
      3. Snagit
      4. Faststone Capture
      5. Camtasia Studio

6. The Wayback Machine
7. X1 Social Discovery

e. Considerations for capturing social media
   i. Document your process
   ii. Potential information to document about process.
      1. Case Information
      2. Who performed the capture
      3. Who requested the capture
      4. Credentials used (or whether any required)
      5. Equipment used (software versions)
      6. URLs & Date/Time Stamps