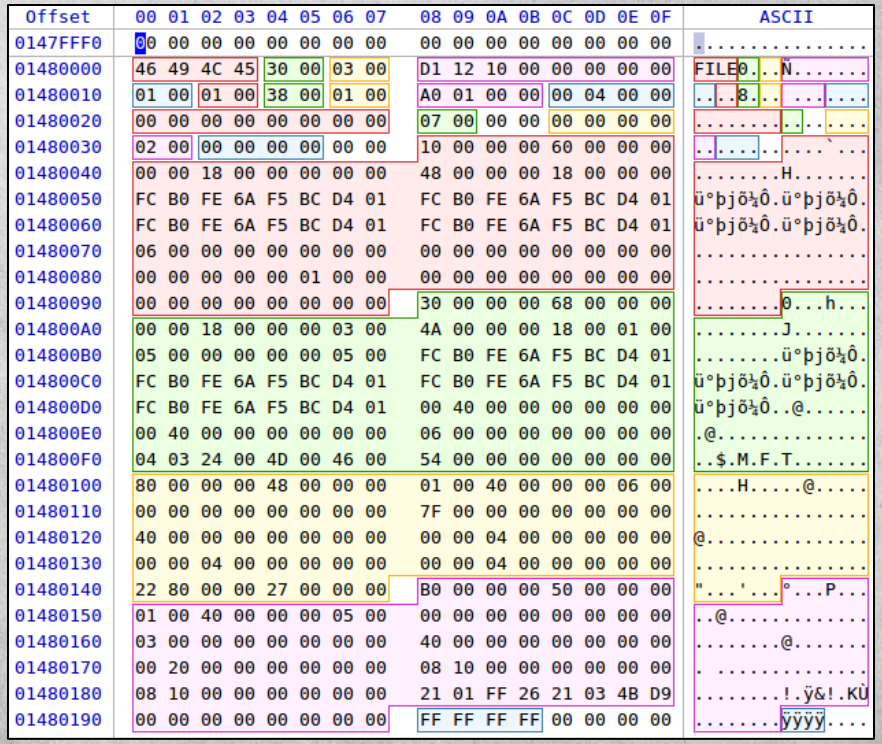
COMP 5350 / 6350
Digital Forensics

File Analysis

Which of the following file attributes are associated with the file below? (Select all that apply)

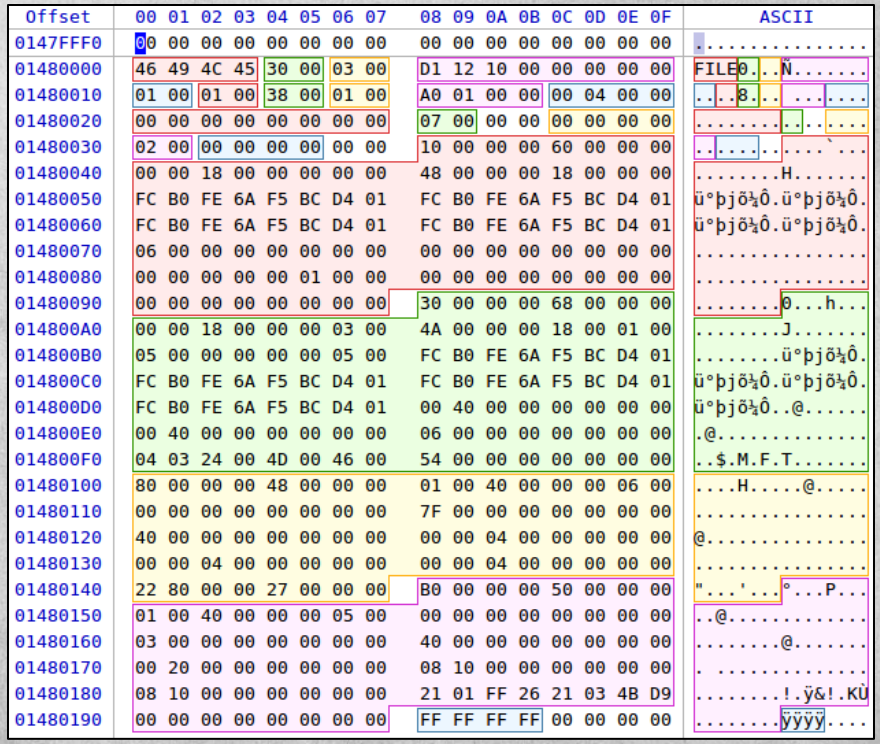
a) $STANDARD_INFORMATION
b) $FILE_ENTRY
c) $COMPRESSION
d) $DATA



Offset	00 01 02 03 04 05 06 07	08 09 0A 0B 0C 0D 0E 0F	ASCII
0147FFF0	00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00	....
01480000	46 49 4C 45 30 00 03 00	D1 12 10 00 00 00 00 00	FILE0...Ñ.......
01480010	01 00 01 00 38 00 01 00	A0 01 00 00 00 04 00 00	....8... .....
01480020	00 00 00 00 00 00 00 00	07 00 00 00 00 00 00 00	....
01480030	02 00 00 00 00 00 00 00	10 00 00 00 60 00 00 00	............`...
01480040	00 00 18 00 00 00 00 00	48 00 00 00 18 00 00 00	........H.......
01480050	FC B0 FE 6A F5 BC D4 01	FC B0 FE 6A F5 BC D4 01	ü°þjõ¼Ô.ü°þjõ¼Ô.
01480060	FC B0 FE 6A F5 BC D4 01	FC B0 FE 6A F5 BC D4 01	ü°þjõ¼Ô.ü°þjõ¼Ô.
01480070	06 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00	....
01480080	00 00 00 00 00 01 00 00	00 00 00 00 00 00 00 00	....
01480090	00 00 00 00 00 00 00 00	30 00 00 00 68 00 00 00	........0...h...
014800A0	00 00 18 00 00 00 03 00	4A 00 00 00 18 00 01 00	........J.......
014800B0	05 00 00 00 00 00 05 00	FC B0 FE 6A F5 BC D4 01	........ü°þjõ¼Ô.
014800C0	FC B0 FE 6A F5 BC D4 01	FC B0 FE 6A F5 BC D4 01	ü°þjõ¼Ô.ü°þjõ¼Ô.
014800D0	FC B0 FE 6A F5 BC D4 01	00 40 00 00 00 00 00 00	ü°þjõ¼Ô..@......
014800E0	00 40 00 00 00 00 00 00	06 00 00 00 00 00 00 00	.@..............
014800F0	04 03 24 00 4D 00 46 00	54 00 00 00 00 00 00 00	..$.M.F.T.......
01480100	80 00 00 00 48 00 00 00	01 00 40 00 00 00 06 00	....H.....@.....
01480110	00 00 00 00 00 00 00 00	7F 00 00 00 00 00 00 00	....
01480120	40 00 00 00 00 00 00 00	00 00 04 00 00 00 00 00	@...............
01480130	00 00 04 00 00 00 00 00	00 00 04 00 00 00 00 00	....
01480140	22 80 00 00 27 00 00 00	B0 00 00 00 50 00 00 00	"...'...°...P...
01480150	01 00 40 00 00 00 05 00	00 00 00 00 00 00 00 00	..@.............
01480160	03 00 00 00 00 00 00 00	40 00 00 00 00 00 00 00	........@.......
01480170	00 20 00 00 00 00 00 00	08 10 00 00 00 00 00 00	. ...
01480180	08 10 00 00 00 00 00 00	21 01 FF 26 21 03 4B D9	........!.ÿ&!.KÙ
01480190	00 00 00 00 00 00 00 00	FF FF FF FF 00 00 00 00	........ÿÿÿÿ....

Which of the following file attributes are associated with the file below? (Select all that apply)

a) $STANDARD_INFORMATION
b) $FILE_ENTRY
c) $COMPRESSION
d) $DATA



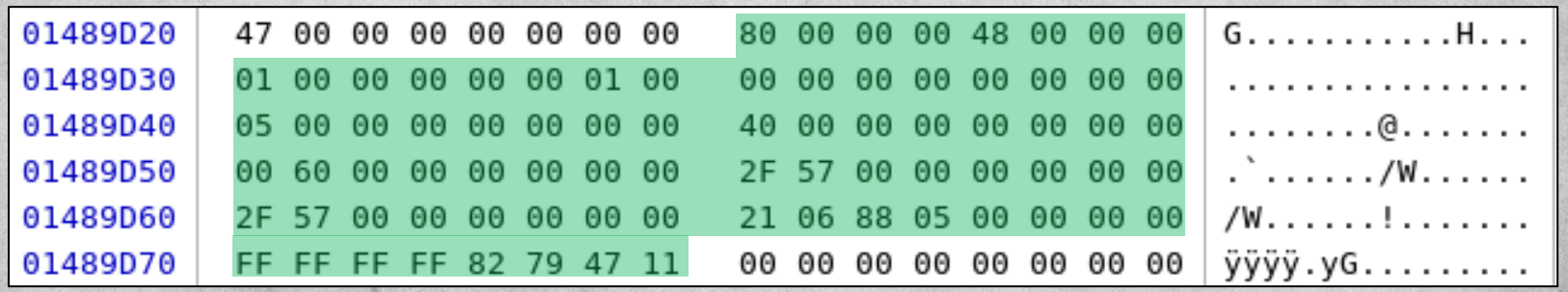
Offset	00 01 02 03 04 05 06 07	08 09 0A 0B 0C 0D 0E 0F	ASCII
0147FFF0	00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00	....
01480000	46 49 4C 45 30 00 03 00	D1 12 10 00 00 00 00 00	FILE0...Ñ.......
01480010	01 00 01 00 38 00 01 00	A0 01 00 00 00 04 00 00	....8... .....
01480020	00 00 00 00 00 00 00 00	07 00 00 00 00 00 00 00	....
01480030	02 00 00 00 00 00 00 00	10 00 00 00 60 00 00 00	............`...
01480040	00 00 18 00 00 00 00 00	48 00 00 00 18 00 00 00	........H.......
01480050	FC B0 FE 6A F5 BC D4 01	FC B0 FE 6A F5 BC D4 01	ü°þjõ¼Ô.ü°þjõ¼Ô.
01480060	FC B0 FE 6A F5 BC D4 01	FC B0 FE 6A F5 BC D4 01	ü°þjõ¼Ô.ü°þjõ¼Ô.
01480070	06 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00	....
01480080	00 00 00 00 00 01 00 00	00 00 00 00 00 00 00 00	....
01480090	00 00 00 00 00 00 00 00	30 00 00 00 68 00 00 00	........0...h...
014800A0	00 00 18 00 00 00 03 00	4A 00 00 00 18 00 01 00	........J.......
014800B0	05 00 00 00 00 00 05 00	FC B0 FE 6A F5 BC D4 01	........ü°þjõ¼Ô.
014800C0	FC B0 FE 6A F5 BC D4 01	FC B0 FE 6A F5 BC D4 01	ü°þjõ¼Ô.ü°þjõ¼Ô.
014800D0	FC B0 FE 6A F5 BC D4 01	00 40 00 00 00 00 00 00	ü°þjõ¼Ô..@......
014800E0	00 40 00 00 00 00 00 00	06 00 00 00 00 00 00 00	.@..............
014800F0	04 03 24 00 4D 00 46 00	54 00 00 00 00 00 00 00	..$.M.F.T.......
01480100	80 00 00 00 48 00 00 00	01 00 40 00 00 00 06 00	....H.....@.....
01480110	00 00 00 00 00 00 00 00	7F 00 00 00 00 00 00 00	....
01480120	40 00 00 00 00 00 00 00	00 00 04 00 00 00 00 00	@...............
01480130	00 00 04 00 00 00 00 00	00 00 04 00 00 00 00 00	....
01480140	22 80 00 00 27 00 00 00	B0 00 00 00 50 00 00 00	"...'...°...P...
01480150	01 00 40 00 00 00 05 00	00 00 00 00 00 00 00 00	..@.............
01480160	03 00 00 00 00 00 00 00	40 00 00 00 00 00 00 00	........@.......
01480170	00 20 00 00 00 00 00 00	08 10 00 00 00 00 00 00	. ...
01480180	08 10 00 00 00 00 00 00	21 01 FF 26 21 03 4B D9	........!.ÿ&!.KÙ
01480190	00 00 00 00 00 00 00 00	FF FF FF FF 00 00 00 00	........ÿÿÿÿ....

$SI

$DATA

For the highlighted $DATA attribute for the NTFS file below, what is the current status of the file?

a) Active
b) Archive
c) Deleted
d) Hidden

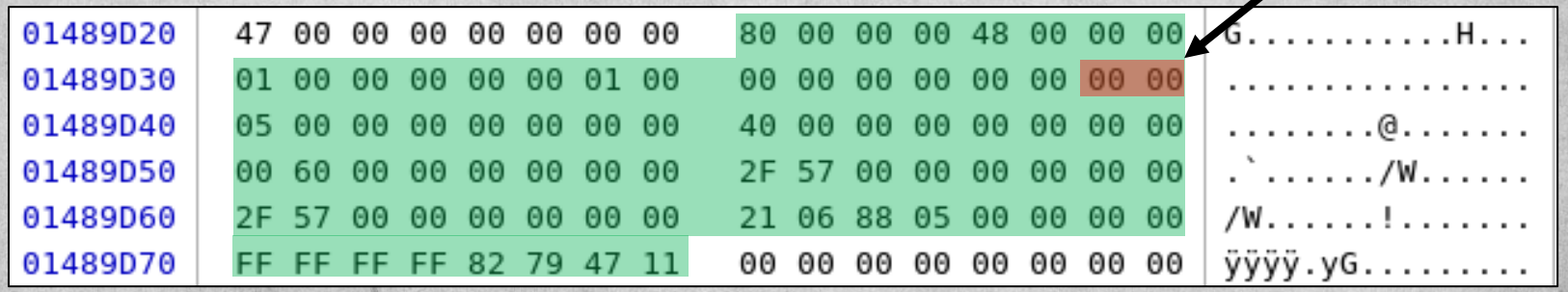


01489D20	47 00 00 00 00 00 00 00	80 00 00 00 48 00 00 00	G.....H...
01489D30	01 00 00 00 00 00 01 00	00 00 00 00 00 00 00 00	.....
01489D40	05 00 00 00 00 00 00 00	40 00 00 00 00 00 00 00	.....@.....
01489D50	00 60 00 00 00 00 00 00	2F 57 00 00 00 00 00 00	.`...../W.....
01489D60	2F 57 00 00 00 00 00 00	21 06 88 05 00 00 00 00	/W.....!.....
01489D70	FF FF FF FF 82 79 47 11	00 00 00 00 00 00 00 00	ÿÿÿÿ.yG.....

For the highlighted $DATA attribute for the NTFS file below, what is the current status of the file?

a) Active
b) Archive
c) Deleted
d) Hidden

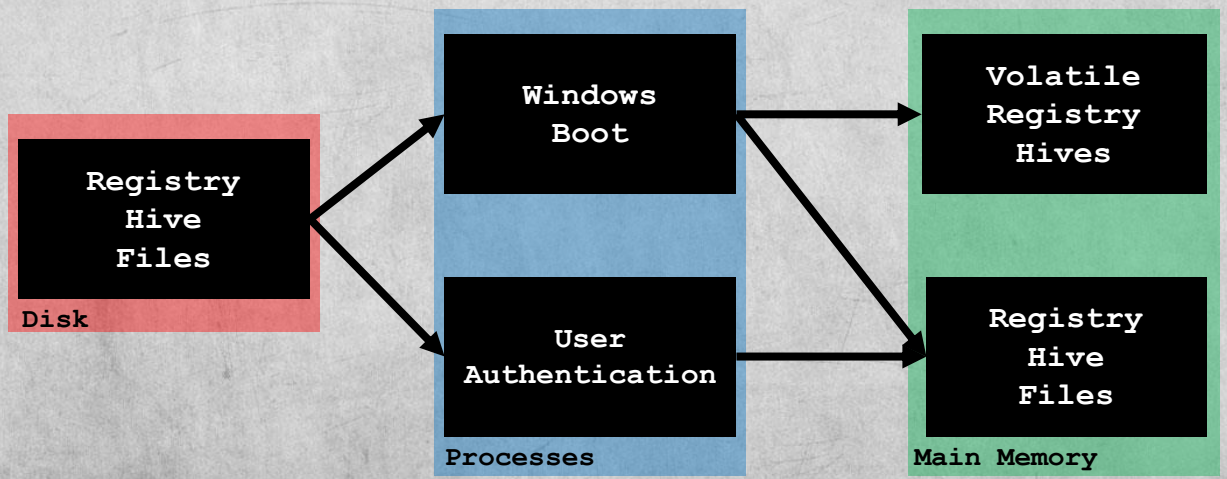
Byte 23 and 24



01489D20	47 00 00 00 00 00 00 00	80 00 00 00 48 00 00 00	G.....H...
01489D30	01 00 00 00 00 00 01 00	00 00 00 00 00 00 00 00	.....
01489D40	05 00 00 00 00 00 00 00	40 00 00 00 00 00 00 00	.....@.....
01489D50	00 60 00 00 00 00 00 00	2F 57 00 00 00 00 00 00	.`....../W.....
01489D60	2F 57 00 00 00 00 00 00	21 06 88 05 00 00 00 00	/W......!.....
01489D70	FF FF FF FF 82 79 47 11	00 00 00 00 00 00 00 00	ÿÿÿÿ.yG.....

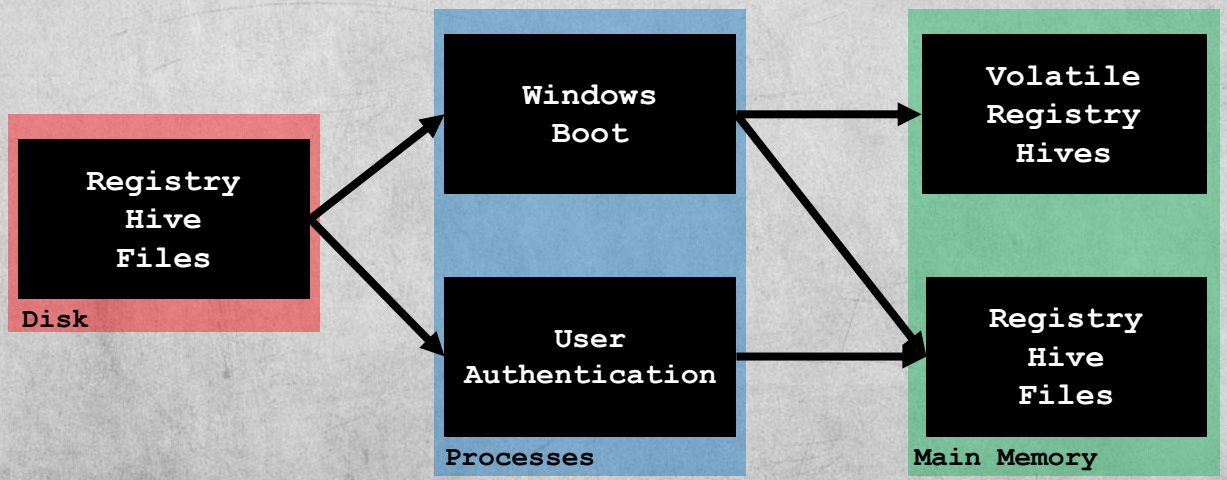
In order to capture a complete Windows Registry from main memory, what two activities must occur first?

a) System Boot
b) UEFI Access
c) User Authentication
d) LDAP Configuration



In order to capture a complete Windows Registry from main memory, what two activities must occur first?

a) System Boot
b) UEFI Access
c) User Authentication
d) LDAP Configuration



How many bytes are in each NTFS MFT entry?

a) 512 bytes
b) 1024 bytes
c) 2048 bytes
d) 4096 bytes

How many bytes are in each NTFS MFT entry?

a) 512 bytes
b) 1024 bytes
c) 2048 bytes
d) 4096 bytes

Forensic Analysis Methods

Analysis Methods

- **There are various methods of forensic analysis that can be used in the recovery and examination of data**
  - ✓ **File System Analysis***
    - **Analysis of disk partitions to extract files**
  - ✓ **Media Analysis**
    - **Analysis of bulk storage devices without a file system**
  - ✓ **Media Management Analysis**
    - **Analysis of RAID arrays and FLASH memory**
  - ✓ **Application Analysis**
    - **Analysis of data inside a file using application specific file format information**

Application Analysis Categories

- **Application analysis can be divided into the following categories:**
  - ✓ **Operating System Analysis**
    - **Investigation of operating system parameters including system settings, network settings, installed software, and authorization**
  - ✓ **Program Analysis**
    - **Analysis of application data and system logs**
  - ✓ **Multimedia Analysis**
    - **Analysis of files that use specific multimedia**

Application Analysis Steps

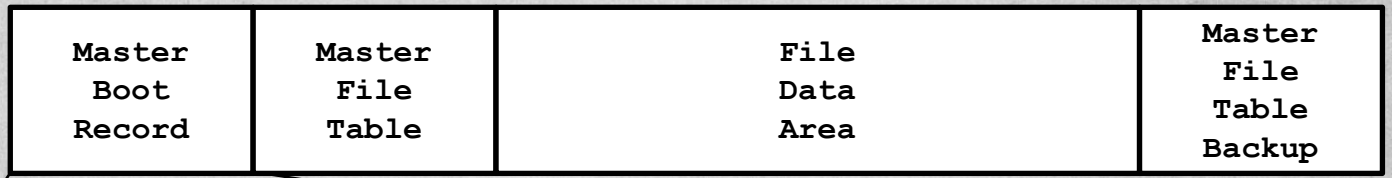
- When conducting file analysis at the application level, there are several key steps that should be applied:
  - ✓ File Signature Analysis (i.e. “Magic Numbers”)
    - File Header and Footer Detection
  - ✓ Hash Verification
    - Integrity Checking and Comparison
  - ✓ Keyword Analysis
    - Keyword Search and Assessment
  - ✓ Statistical Analysis
    - Mathematical-Based Analysis
  - ✓ Content Analysis
    - Pattern Analysis Based on Known File Structure

Cluster-Based File Analysis

Cluster-Based File Carving

- **Our attention up to this point has focused on specific file systems to recover digital artifacts**
- **This is a stepwise process which requires moving through the structures of each file system**
- **Based on the structure of a given file system we can search for files along cluster boundaries**
  - ✓ **Cluster-Based File Carving (CBFC)**
- **There will be times when digital evidence will be stored without the aid of a file system**

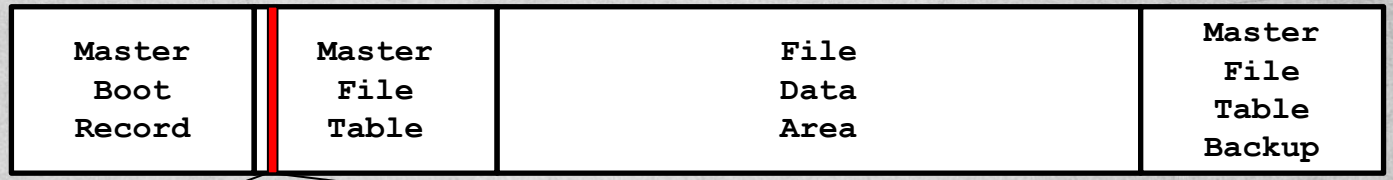
Cluster-Based File Carving - MBR



Offset	00 01 02 03 04 05 06 07	08 09 0A 0B 0C 0D 0E 0F	ASCII
00100000	EB 52 90 4E 54 46 53 20	20 20 20 00 02 08 00 00	ëR.NTFS ...
00100010	00 00 00 00 00 F8 00 00	3F 00 FF 00 00 08 00 00	.....ø..?.ÿ.....
00100020	00 00 00 00 80 00 00 00	FF AF 04 00 00 00 00 00	....€...ÿ¯......
00100030	00 32 00 00 00 00 00 00	02 00 00 00 00 00 00 00	.2..... .....
00100040	F6 00 00 00 01 00 00 00	C7 2D C5 78 72 C5 78 BA	ö.....Ç-ÅxrÅxº
00100050	00 00 00 00 FA 33 C0 8E	D0 BC 00 7C FB 68 C0 07	....ú3À.Ð¼.|ûhÀ.

NTFS Master Boot Record provides information for Master File Table…

Cluster-Based File Carving - MFT



Offset	00	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F	ASCII
09714000	46	49	4C	45	30	00	03	00	00	00	00	00	00	00	00	00	FILE0.....
09714010	02	00	00	00	38	00	00	00	A8	01	00	00	00	04	00	00	....8...¨.......
09714020	00	00	00	00	00	00	00	00	04	00	00	00	40	00	00	00	........@...
09714030	16	00	00	00	00	00	00	00	10	00	00	00	48	00	00	00	....H...
09714040	00	00	00	00	00	00	00	00	30	00	00	00	18	00	00	00	........0.......
09714050	74	DB	59	7B	EF	91	D6	01	3A	38	8D	2F	EE	91	D6	01	tÛY{ï.Ö.:8./î.Ö.
09714060	82	4E	5A	7B	EF	91	D6	01	80	C2	7F	7B	EF	91	D6	01	.NZ{ï.Ö..Â.{ï.Ö.
09714070	20	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	....
09714080	30	00	00	00	70	00	00	00	00	00	00	00	00	00	03	00	0...p...........
09714090	52	00	00	00	18	00	01	00	05	00	00	00	00	00	05	00	R...............
097140A0	74	DB	59	7B	EF	91	D6	01	74	DB	59	7B	EF	91	D6	01	tÛY{ï.Ö.tÛY{ï.Ö.
097140B0	74	DB	59	7B	EF	91	D6	01	74	DB	59	7B	EF	91	D6	01	tÛY{ï.Ö.tÛY{ï.Ö.
097140C0	00	80	00	00	00	00	00	00	00	00	00	00	00	00	00	00	....
097140D0	20	00	00	00	00	00	00	00	08	00	4D	00	65	00	6F	00	.........M.e.o.
097140E0	77	00	2E	00	4A	00	50	00	47	00	00	00	18	00	00	00	w...J.P.G.......
097140F0	50	00	00	00	68	00	00	00	00	00	00	00	00	00	01	00	P...h...........
09714100	50	00	00	00	18	00	00	00	01	00	04	80	14	00	00	00	P...............
09714110	24	00	00	00	00	00	00	00	34	00	00	00	01	02	00	00	$.......4.......
09714120	00	00	00	05	20	00	00	00	20	02	00	00	01	02	00	00	....
09714130	00	00	00	05	20	00	00	00	20	02	00	00	02	00	1C	00	....
09714140	01	00	00	00	00	03	14	00	FF	01	1F	00	01	01	00	00	........ÿ.......
09714150	00	00	00	01	00	00	00	00	80	00	00	00	48	00	00	00	............H...
09714160	01	00	40	00	00	00	02	00	00	00	00	00	00	00	00	00	..@.............
09714170	07	00	00	00	00	00	00	00	40	00	00	00	00	00	00	00	........@.......
09714180	00	80	00	00	00	00	00	00	96	7C	00	00	00	00	00	00	.........|......
09714190	96	7C	00	00	00	00	00	00	21	08	00	1B	00	00	00	00	.|......!.......
097141A0	FF	FF	FF	FF	00	00	00	00	FF	FF	FF	FF	00	00	00	00	ÿÿÿÿ....ÿÿÿÿ....

NTFS Master File Table entry provides information about file metadata and content…

Cluster-Based File Carving – File Recovery

Offset	00 01 02 03 04 05 06 07	08 09 0A 0B 0C 0D 0E 0F	ASCII
09714150	00 00 00 01 00 00 00 00	80 00 00 00 48 00 00 00	....H...
09714160	01 00 40 00 00 00 02 00	00 00 00 00 00 00 00 00	..@..... ....
09714170	07 00 00 00 00 00 00 00	40 00 00 00 00 00 00 00	.... @.......
09714180	00 80 00 00 00 00 00 00	96 7C 00 00 00 00 00 00	....|......
09714190	96 7C 00 00 00 00 00 00	21 08 00 1B 00 00 00 00	.|...... !.......
097141A0	FF FF FF FF 00 00 00 00	FF FF FF FF 00 00 00 00	ÿÿÿÿ.... ÿÿÿÿ....

1) Determine file offsets
2) Find file name
3) Locate data contents
4) Recover file

dd if=ntfs.dd of=Recovered.jpg bs=1 skip=657827 count=10252

File Signature Analysis

File Applications

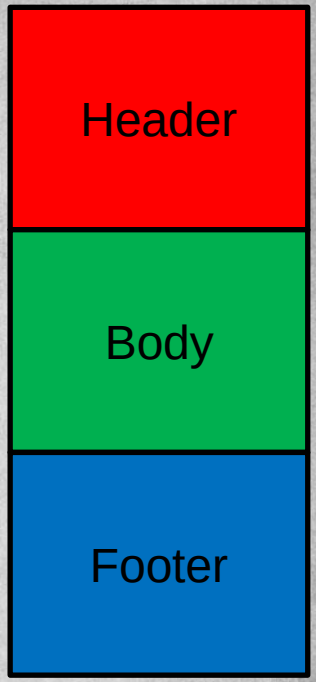
- **Common file application types that we will have to contend with during a forensic investigation include:**
  - ✓ **Images**
  - ✓ **Audio**
  - ✓ **Video**
  - ✓ **Archive**
  - ✓ **Documents**

- **It is necessary to understand how application data is configured before an analyst can properly evaluate raw data for recovery**

- **We will now introduce application data architecture and file signature analysis**

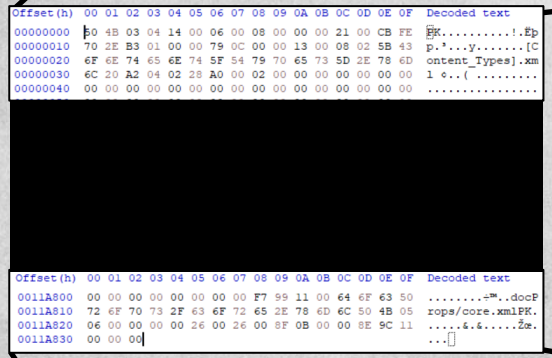
Application Data Architecture

- **For non-flat files, application data generally contains a header, body, and footer or some combination of each**
- **Forensic examiners attempting to recover data may be able to build parts of files due to corruption or active manipulation**
- **Application headers store file format information immediately after the signature and application metadata can indicate key information such as file size, data format, software version**



Header & Footer Example

Offset(h)	00	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F	Decoded text
00000000	50	4B	03	04	14	00	06	00	08	00	00	00	21	00	CB	FE	PK.....!.Ëþ
00000010	70	2E	B3	01	00	00	79	0C	00	00	13	00	08	02	5B	43	p.³...y.....[C
00000020	6F	6E	74	65	6E	74	5F	54	79	70	65	73	5D	2E	78	6D	ontent_Types].xm
00000030	6C	20	A2	04	02	28	A0	00	02	00	00	00	00	00	00	00	l ¢..(.....



Offset(h)	00	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F	Decoded text
0011A800	00	00	00	00	00	00	00	00	F7	99	11	00	64	6F	63	50	.....÷™..docP
0011A810	72	6F	70	73	2F	63	6F	72	65	2E	78	6D	6C	50	4B	05	rops/core.xmlPK.
0011A820	06	00	00	00	00	26	00	26	00	8F	0B	00	00	8E	9C	11	.....&.&.....Žœ.
0011A830	00	00	00														...

Image Applications

Image Data

- Files that contain data to be rendered as graphical output
- Many image applications contain metadata with three main image metadata types being used:
  - ✓ EXIF
    - Exchangeable Image File Format
    - Embeds data relative to capture devices and uses tags and values including date-time and geolocation
  - ✓ IPTC
    - Information Interchange Model
    - A legacy metadata model that embeds information about images used by newspapers and news agencies
  - ✓ XMP
    - eXtensible Metadata Platform
    - Developed by Adobe in 2001 to provide open metadata methods

Image File Types – JPEG

- **Joint Photographic Experts Group**
  - ✓ **Developed in 1992 as a method to compress images while maintaining image content**
  - ✓ **Headers**
    - **0xFFD8FFE0 – Standard JPEG/JFIF File**
    - **0xFFD8FFE1 – Standard JPEG File with EXIF Metadata**
    - **0xFFD8FFE2 – Canon Camera Image File Format (CIFF) JPEG**
    - **0xFFD8FFE8 – Still Picture Interchange File Format (SPIFF)**

```
File Types $ hexdump Mickey.jpg -C -s $(( 0*512 )) -n $(( 1*512 ))
00000000  ff d8 ff e0 00 10 4a 46  49 46 00 01 01 00 00 01  |......JFIF......|
00000010  00 01 00 00 ff db 00 84  00 03 02 02 08 08 08 08  |................|
00000020  08 08 08 08 08 08 08 08  08 08 08 08 08 08 08 08  |................|
```

Image File Types – GIF

- **Graphic Interchange Format**
  - ✓ **Initially designed for icons and simple graphics and supports transparency and animations**
  - ✓ **GIF do not generally contain extensive metadata**
  - ✓ **Headers:**
    - **0x474946383761 – GIF87a**
    - **0x474946383961 – GIF89a**
  - ✓ **Footer:**
    - **0x003B – ;|**

```
File Types $ hexdump Minions.gif -C -s $(( 0*512 )) -n $(( 1*512 ))
00000000  47 49 46 38 39 61 c8 00  c8 00 f7 f6 00 00 18 3a  |GIF89a.........:|
00000010  00 67 98 00 73 a7 00 88  b9 00 b8 de 00 cb df 01  |.g..s...........|
00000020  21 34 01 29 53 01 57 88  01 73 9b 01 7b af 01 dc  |!4.)S.W..s..{...|
```

Header

```
000ed680  1d 48 c5 ba ab cd 12 ed  c4 1e 2d c0 18 2c bd 42  |.H........-..,.B|
000ed690  e9 b5 2a ed ce 4a ed a5  ce ec b9 0e 2d d1 6a ed  |..*..J.......-.j.|
000ed6a0  5b 06 04 00 3b                                    |[...;|
```

Footer

Image File Types – PNG

- **Portable Network Graphics**
  - ✓ **An open and free replacement for GIF images**
  - ✓ **PNG images do not generally contain extensive metadata**
  - ✓ **Headers:**
    - **0x89504E470D0A1A0A – .PNG….**
  - ✓ **Footer:**
    - **0x49454E44AE426082 – IEND.B`.|**

```
File Types $ hexdump Mario.png -C -s $(( 0*512 )) -n $(( 1*512 ))
00000000  89 50 4e 47 0d 0a 1a 0a  00 00 00 0d 49 48 44 52  |.PNG........IHDR|
00000010  00 00 05 c1 00 00 07 3e  08 06 00 00 00 26 76 ed  |.......>.....&v.|
00000020  9c 00 00 00 09 70 48 59  73 00 00 35 d4 00 00 35  |.....pHYs..5...5|
```

Header

```
001aee00  af d6 25 4e 08 21 84 10  42 08 21 84 10 42 6e 01  |..%N.!..B.!..Bn.|
001aee10  fe ff 01 00 9b 8d f7 1e  1b d4 77 1a 00 00 00 00  |..........w.....|
001aee20  49 45 4e 44 ae 42 60 82                          |IEND.B`.|
```

Footer

Image File Types – TIFF

- **Tagged Image File Format**
  - ✓ **Image file format for publishing and graphic design and was originally created for fax and scanning applications**
  - ✓ **Default file format for OS X image applications**
  - ✓ **Headers:**
    - **0x4D4D002A – MM.* – Motorola**
    - **0x4D4D002B – MM.+ – Files > 4 GB**

```
File Types $ hexdump f14.tif -C -s $(( 0*512 )) -n $(( 1*512 ))
00000000  4d 4d 00 2a 00 04 b0 08  a4 a4 a2 a4 a6 a7 a7 a6  |MM.*............|
00000010  a4 a8 a9 a8 a3 a3 a4 a5  a5 a3 a8 a9 a4 a7 a3 a3  |................|
00000020  a3 a3 a5 a7 a2 a6 a3 a4  a4 a4 a7 a5 a3 a5 a4 a5  |................|
```

Header

Image File Types – BMP / DIB

- **Bitmap / Device Independent Bitmap**
  - ✓ **A legacy format created by Microsoft for use in Windows image applications**
  - ✓ **Files are stored as lists of Red-Green-Blue (RGB) values**
  - ✓ **Headers:**
    - **0x42 4D XX XX XX XX**

```
File Types $ hexdump -C Marbles.BMP -s $(( 0*512 )) -n $(( 1*512 ))
00000000  42 4d 7c 11 41 00 00 00  00 00 36 00 00 00 28 00  |BM|.A.....6...(.|
00000010  00 00 8b 05 00 00 e9 03  00 00 01 00 18 00 00 00  |................|
00000020  00 00 00 00 00 00 20 2e  00 00 20 2e 00 00 00 00  |...... ... .....|
00000030  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00  |................|
```

Image Standards

- **Using known information about file signatures and formats, it will be possible to recover files without cluster carving**

Reserved Area		
Description	Offset	Bytes
Signature - "BM"	0000h	2
File Size*	0002h	4
Reserved - "00 00"	0006h	2
Reserved - "00 00"	0008h	2
Data Offset	000Ah	4
Header Size - "40"	000Eh	4
Image Width	0012h	4
Image Height	0016h	4
Image Planes	001Ah	2
Bits / Pixel	001Ch	2
Compression Type	001Eh	4
Image Data Size	0022h	4
Horizontal Resolution	0026h	4
Vertical Resolution	002Ah	4
Number of Colors	002Eh	4
Number of Important Colors	0032h	4

```
File Types $ hexdump -C Marbles.BMP -s $(( 0*512 )) -n $(( 1*512 ))
00000000  42 4d 7c 11 41 00 00 00  00 00 36 00 00 00 28 00  |BM|.A.....6...(.|
00000010  00 00 8b 05 00 00 e9 03  00 00 01 00 18 00 00 00  |................|
00000020  00 00 00 00 00 00 20 2e  00 00 20 2e 00 00 00 00  |...... ... .....|
00000030  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00  |................|
```

Image Analysis Tools

- **Analysis of image files can be achieved with standard Linux commands and applications**
  - ✓ **file**
    - **Determines the file type of a digital artifact**
    - **Performs filesystem tests, magic number analysis, and language assessment**
  - ✓ **extract**
    - **Command-line tool extracts document metadata**
  - ✓ **exiftool**
    - **Command-line application and Perl library for reading and writing EXIF, GPS, IPTC, XMP data**
  - ✓ **imagemagick, identify**
    - **Software suite to create, edit, and compose bitmap images**
    - **It can also read, convert, and write images in multiple formats**

Audio Applications

Audio Data

- **Files that contain data for audio applications including:**
  - ✓ **Music**
  - ✓ **Voice Mail Messages**
  - ✓ **RF Audio Applications**

- **Some of the image-based metadata types also apply to audio applications**
  - ✓ **EXIF**
  - ✓ **XMP**

- **Due to intellectual property considerations for audio applications, metadata can provide significant information about ownership**

Audio File Types – WAV

- **Waveform Audio File Format**
  - ✓ **An audio standard developed by IBM and Microsoft in 1991 for storing audio bitstream on desktop PC's**
  - ✓ **WAV files are composed of tagged data chunks inside of a Resource Interchange File Format (RIFF) container**
  - ✓ **Headers**
    - **0x52494646xxxxxxxx57415645666D7420 – RIFF….WAVEfmt**
      - **Middle 4 bytes indicate file size**

```
File Types $ hexdump Moderato.wav -C -s $(( 0*512 )) -n $(( 1*512 ))
00000000  52 49 46 46 3a 60 10 00  57 41 56 45 66 6d 74 20  |RIFF:`..WAVEfmt |
00000010  10 00 00 00 01 00 02 00  40 1f 00 00 00 7d 00 00  |........@....}..|
00000020  04 00 10 00 64 61 74 61  34 5f 10 00 8e ff 17 00  |....data4_......|
```

0x10603a -> 1073210 bytes

Audio File Types – MPEG

- **Moving Picture Experts Group**
  - ✓ **Published by MPEG in 1993 and has become the most commonly used format for digital music**
  - ✓ **There are numerous MPEG versions including with MP3 (MPEG-1 Audio Layer 3) and M4A (Apple Lossless Audio Codec) being the most prevalent**
  - ✓ **MP3 was used extensively in Peer-to-Peer file sharing applications such as Napster, Kazaa, and Gnutella**
  - ✓ **Headers**
    - **0x494433 – ID3**
    - **0x667479704D344120 – ftypM4A**
      - **4-byte offset**

```
File Types $ hexdump MPEG.mp3 -C -s $(( 0*512 )) -n $(( 1*512 ))
00000000  49 44 33 03 00 00 00 00  00 66 54 43 4f 4e 00 00  |ID3......fTCON..|
00000010  00 0a 00 00 00 43 69 6e  65 6d 61 74 69 63 54 41  |.....CinematicTA|
00000020  4c 42 00 00 00 16 00 00  00 59 6f 75 54 75 62 65  |LB.......YouTube|
```

Audio File Types – ASF / WMA(V)

- **Advanced Systems Format / Windows Media Audio (Video)**
  - ✓ **ASF is a Microsoft designed container format for streaming media and is used to store WMA(V) data**
  - ✓ **Although this format is proprietary, it has been reverse engineered and can be played with open source codecs including ffmpeg**
  - ✓ **Metadata can be extracted from ASF**
  - ✓ **Headers**
    - **0x3026B2758E66CF11A6D900AA0062CE6C**

```
File Types $ hexdump ASFExample.asf -C -s $(( 0*512 )) -n $(( 1*512 ))
00000000  30 26 b2 75 8e 66 cf 11  a6 d9 00 aa 00 62 ce 6c  |0&.u.f.......b.l|
00000010  7f 02 00 00 00 00 00 00  05 00 00 00 01 02 a1 dc  |................|
00000020  ab 8c 47 a9 cf 11 8e e4  00 c0 0c 20 53 65 68 00  |..G........ Seh.|
```

Audio Analysis Tools

- **Analysis of audio files can be achieved with standard Linux commands and applications**
  - ✓ **file**
    - **Determines the file type of a digital artifact**
    - **Performs filesystem tests, magic number analysis, and language assessment**
  - ✓ **extract**
    - **Command-line tool extracts audio metadata**
  - ✓ **exiftool**
    - **Command-line application and Perl library for reading and writing EXIF, GPS, IPTC, XMP data**
  - ✓ **AtomicParsley**
    - **Extracts metadata from audio files**
    - **Writes metadata onto MPEG files**

Video Applications

Video Data

- Due to both video and audio components of a video file, there is usually a container that encompasses both video and audio streams
- The containers used to compress and encode data streams is known as a “codec” and is necessary to properly playback video and audio content
- Video files generally contain useable metadata and the examples defined next will illustrate their use

Video File Types – MPEG

- Moving Picture Experts Group
  - ✓ As described in the audio version of MPEG, this standard includes video and audio compression and was used extensively in video CD applications
  - ✓ The common versions of MPEG video include
    - MPEG-1 – CD MPEG1
      - Header: 0x52494646
    - MPEG-2 – DVD MPEG2
      - Header: 0x000001BA
      - Footer: 0x000001B9
    - MP4 – MPEG-4
      - 0x667479706D703432
        - 4-byte offset

```
File Types $ hexdump Earth.mp4 -C -s $(( 0*512 )) -n $(( 1*512 ))
00000000  00 00 00 20 66 74 79 70  6d 70 34 32 00 00 00 00  |... ftypmp42....|
00000010  6d 70 34 32 6d 70 34 31  69 73 6f 6d 61 76 63 31  |mp42mp41isomavc1|
00000020  00 00 2e 30 6d 6f 6f 76  00 00 00 6c 6d 76 68 64  |...0moov...lmvhd|
```

Video File Types – AVI

- **Audio Video Interleave**
  - ✓ **A video container format developed by Microsoft in 1992 which uses the RIFF contains discussed earlier**
  - ✓ **AVI can use multiple codecs to compress and decode video and audio streams**
  - ✓ **The contents of AVI files contain a sequence of values known as “FourCC” codes that specify which codecs are necessary for replay**
  - ✓ **Headers**
    - **0x52494646xxxxxxxx415649204C495354 – RIFF….AVI LIST**
      - **4-byte File Size**

```
File Types $ hexdump AVIExample.avi -C -s $(( 0*512 )) -n $(( 1*512 ))
00000000  52 49 46 46 46 54 0b 00  41 56 49 20 4c 49 53 54  |RIFFFT..AVI LIST|
00000010  be 22 00 00 68 64 72 6c  61 76 69 68 38 00 00 00  |."..hdrlavih8...|
00000020  35 82 00 00 2e 44 00 00  00 00 00 00 10 09 00 00  |5....D..........|
```

0xb5446 -> 742470 bytes

Video File Types – MOV

- QuickTime File Format
  - ✓ A legacy video format developed by Apple in 1991 that has been integrated into the MPEG-4 standard
  - ✓ Headers
    - 0x66747970 71742020
      - 4-byte offset
    - 0x6D6F6F76
      - 4-byte offset

```
File Types $ hexdump MOVExample.mov -C -s $(( 0*512 )) -n $(( 1*512 ))
00000000  00 00 00 14 66 74 79 70  71 74 20 20 00 00 02 00  |....ftypqt  ....|
00000010  71 74 20 20 00 00 00 08  77 69 64 65 00 0a 4d 7f  |qt  ....wide..M.|
00000020  6d 64 61 74 00 00 02 9f  06 05 ff ff 9b dc 45 e9  |mdat..........E.|
```

Video Analysis Tools

- **Analysis of video files can be achieved with standard Linux commands and applications**
  - ✓ **file**
    - **Determines the file type of a digital artifact**
    - **Performs filesystem tests, magic number analysis, and language assessment**
  - ✓ **extract**
    - **Command-line tool extracts video metadata**
  - ✓ **exiftool**
    - **Command-line application and Perl library for reading and writing EXIF, GPS, IPTC, XMP data**
  - ✓ **qtinfo**
    - ✓ **QuickTime utilities package on Ubuntu**
  - ✓ **AtomicParsley**
    - **Extracts metadata from video files**

Archive Applications

Archive Data

- An archive is a file container that holds the contents of other files and provides compression and encryption capabilities
- Archives can provide metadata relative to archive creation and in many cases provide information about file contents when unencrypted

Archive File Types – ZIP

- **ZIP Archive**
  - ✓ **The original compression and archive file format created by Phil Katz (PK) in 1989**
  - ✓ **ZIP archives provide compression and encryption**
  - ✓ **There are numerous ZIP file signatures:**
    - **PKZIP Archive File**
      - **Header: 0x504B0304 – PK..**
      - **Footer: Filename xx(?) 50 4B xx(17) 00 00 00**
    - **WinZIP Compressed Archive**
      - **Header: 0x57696E5A6970**
        - **29,152-byte offset**

```
File Types $ hexdump Mario.zip -C -s $(( 0*512 )) -n $(( 1*512 ))
00000000  50 4b 03 04 14 00 00 00  08 00 8c 7b 3e 51 f2 43  |PK.........{>Q.C|
00000010  c0 6f 8f da 1a 00 28 ee  1a 00 09 00 1c 00 4d 61  |.o....(.......Ma|
00000020  72 69 6f 2e 70 6e 67 55  54 09 00 03 17 a4 74 5f  |rio.pngUT.....t_|
```

```
001adb00  4d 61 72 69 6f 2e 70 6e  67 55 54 05 00 03 17 a4  |Mario.pngUT.....|
001adb10  74 5f 75 78 0b 00 01 04  e8 03 00 00 04 e8 03 00  |t_ux............|
001adb20  00 50 4b 05 06 00 00 00  00 01 00 01 00 4f 00 00  |.PK..........O..|
001adb30  00 d2 da 1a 00 00 00                              |.......|
```

Archive File Types – RAR

- The Roshal Archive
  - ✓ Proprietary archive format designed by Eugene Roshal in 1993 that provides compression, archive repair, archive splitting, and encryption
  - ✓ Compression multiple files into one archive
    - rar a Images Mickey.jpg Minions.gif Mario.png
  - ✓ Headers
    - RAR Version 4
      - 0x526172211A0700
    - RAR Version 5
      - 0x526172211A070100

```
File Types $ hexdump Images.rar -C -s $(( 0*512 )) -n $(( 1*512 ))
00000000  52 61 72 21 1a 07 01 00  09 83 a4 5e 0c 01 05 08  |Rar!.......^....|
00000010  00 07 01 01 a2 e9 a9 81  00 2a f5 0d e9 2a 02 03  |.........*...*..|
00000020  0b e1 f9 02 04 e2 8e 03  b4 83 02 30 19 44 5b 80  |...........0.D[.|
```

Archive File Types – 7z

- 7-zip
  - ✓ Open archive and compression format that provides similar capability to ZIP and RAR archives
  - ✓ 7-zip compression process
    - p7zip Mario.png
  - ✓ Header
    - 0x377ABCAF271C

```
File Types $ hexdump Mario.png.7z -C -s $(( 0*512 )) -n $(( 1*512 ))
00000000  37 7a bc af 27 1c 00 04  6c 72 1f 46 3c e6 1a 00  |7z..'...lr.F<...|
00000010  00 00 00 00 5a 00 00 00  00 00 00 00 2d dc 8d 72  |....Z.......-..r|
00000020  e0 e3 89 e0 03 5d 00 44  94 05 c4 7a 27 f6 f7 ee  |.....].D...z'...|
```

Archive Analysis Tools

- **Analysis of archive files can be achieved with standard Linux commands and applications**
  - ✓ **file**
    - **Determines the file type of a digital artifact**
    - **Performs filesystem tests, magic number analysis, and language assessment**
  - ✓ **extract**
    - **Command-line tool extracts archive metadata**
  - ✓ **exiftool**
    - **Command-line application and Perl library for reading and writing EXIF, GPS, IPTC, XMP data**
  - ✓ **zip / unzip**
  - ✓ **rar / unrar**
  - ✓ **7za**

Document Applications

Document Data

- Document data contains underlying text and image data and can also contain application specific data
- From a forensic standpoint, documents will be the largest category of data that an analyst will have to analyze and recover
- Documents provide an extensive list of metadata such as MAC time stamps, author information, and document revision history
- Documents also provide a threat vector for attacks due to numerous malware attacks against document applications

Document File Types – OLE

- **Object Linking and Embedding Compound File**
  - ✓ **Documents created with Microsoft Office 1997-2003 formats are OLE compound files**
  - ✓ **From a technical perspective OLE files are portable file systems that provide the same structure as standard file systems like FAT and NTFS**
  - ✓ **OLE files provide two mechanisms to store data:**
    - **Storage Objects**
    - **Stream Objects**
  - ✓ **Header**
    - **0xD0 CF 11 E0 A1 B1 1A E1**

"DOCFILE"

```
File Types$ hexdump -C Auburn.doc -s $(( 0*512 )) -n $(( 1*512 ))
00000000  d0 cf 11 e0 a1 b1 1a e1  00 00 00 00 00 00 00 00  |................|
00000010  00 00 00 00 00 00 00 00  3e 00 03 00 fe ff 09 00  |........>.......|
00000020  06 00 00 00 00 00 00 00  00 00 00 00 04 00 00 00  |................|
```

Document File Types – PDF

- **Portable Document Format**
  - ✓ **An Adobe open format that provides a container that holds PostScript layout instructions and embedded fonts and graphics**
  - ✓ **PDF ensures that a document prepared on any system will render the same way on another system**
  - ✓ **There are two types of PDF metadata**
    - **Document Information Directory**
      - **Key / Value pairs that provide information about ownership, document titles and content, and MAC times**
    - **XMP**
  - ✓ **Headers**
    - ✓ **0x25 50 44 46 – %PDF**

```
File Types $ hexdump -C Auburn.pdf -s $(( 0*512 )) -n $(( 1*512 ))
00000000  25 50 44 46 2d 31 2e 37  0d 0a 25 b5 b5 b5 b5 0d  |%PDF-1.7..%.....|
00000010  0a 31 20 30 20 6f 62 6a  0d 0a 3c 3c 2f 54 79 70  |.1 0 obj..<</Typ|
00000020  65 2f 43 61 74 61 6c 6f  67 2f 50 61 67 65 73 20  |e/Catalog/Pages |
```

Document Analysis Tools

- **Analysis of document files can be achieved with standard Linux commands and applications**
  - ✓ **file**
    - **Determines the file type of a digital artifact**
    - **Performs filesystem tests, magic number analysis, and language assessment**
  - ✓ **extract**
    - **Command-line tool extracts document metadata**
  - ✓ **exiftool**
    - **Command-line application and Perl library for reading and writing EXIF, GPS, IPTC, XMP data**

Bash Scripting

Scripting File Header Analysis

- **Shell scripting is available on Linux systems to assist with automating repetitive tasks**

- **Using what we have learned up to this point, we will develop a shell script that allows us to conduct bulk analysis of files based on header and / or footer information**

- **For those not yet familiar with shell scripting we will discuss**
  - **Variables**
  - **Loops**
  - **References**
  - **Command Options**
  - **Pipes**

Encoding

Character Encoding

- A mapping between a character set and encoded representation that is commonly used in software and web applications
- Some basic examples of encoding include:
  - ✓ Morse Code
  - ✓ Hamming Code
  - ✓ ASCII
  - ✓ Hex Encoding
  - ✓ ISO-8859
- Character encoding can aid when data parsing, compression, or when displaying application specific content
- RFC 2978 provides an exhaustive list of character sets commonly used in web application

References

- File Signatures
  - ✓ https://www.garykessler.net/library/file_sigs.html
- File Samples
  - ✓ https://filesamples.com
- File Carving Techniques
  - ✓ https://resources.infosecinstitute.com/file-carving/#gref
- Character Encoding
  - ✓ https://www.w3.org/International/questions/qa-what-is-encoding