File Types and File Carving

File Carving *aka* ***Data Carving*** *or* ***Salvaging***

References:

Carrier, Chapter 8

Nelson, Chapter 8

Multimedia Files

Introduction

Much of the work of cyber forensics involves being able to do the following from a forensic image of a mass storage device:

View documents

View pictures, diagrams, tables and videos

Listen to audio files

Know when and where the above were created or recorded

Separate and extract the above from the forensic image

All of the above are files that can usually be opened by various applications – but not always

Nelson & File Carving

In Chapter 8, Nelson et al discuss file carving in the context of recovering graphics files

- *While this is an important part of file carving and fixing corrupted files, it limits carving's applicability*
- *I'll discuss this more during tonight's lecture*

File Types

What is a *File Type*?

File Type
Possible Definition

What's a ***File Type***?

Suggested definition

- *A* **File Type** *is a file name extension that*
  - Identifies what a file **is** and
  - Relates the file to one or more software applications that can use, create or otherwise deal with it
- *Examples*
  - MS Word files, older MSWord files, Excel spreadsheet files, JPEG image files, AVI video files, open database files…

How Do We and Computers Identify File Types

How do forensic analysts identify the type of a file?

Makes one think a bit

How to operating systems identify the type of a file?

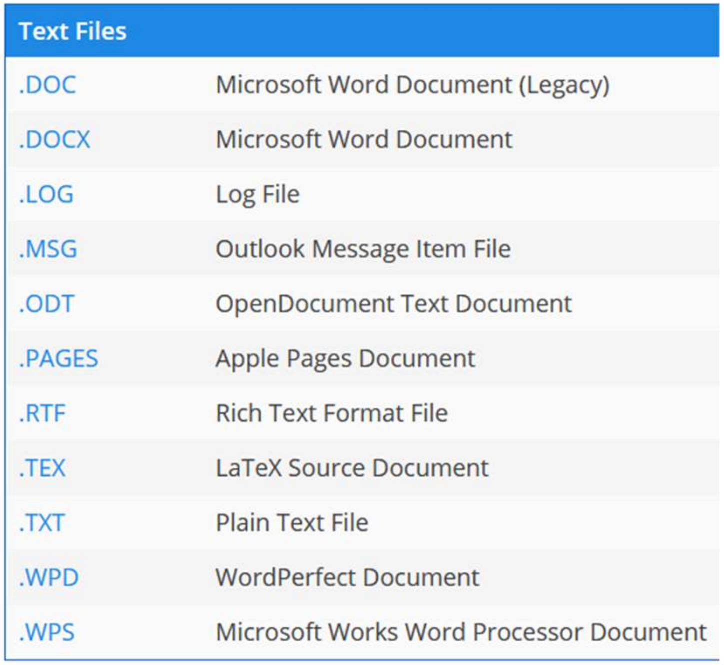
This might be easier to answer in some cases

How does MS Windows do it?

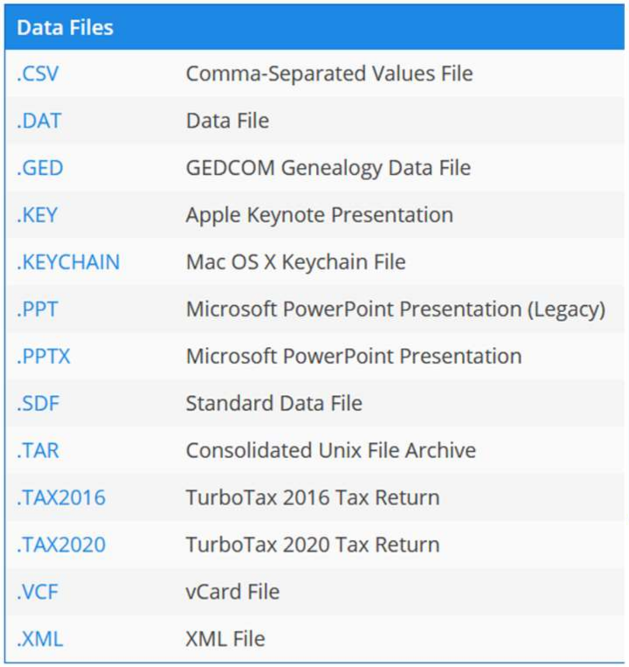
By making the name of the file include the ID of its type

This is accomplished by using file type **extensions**

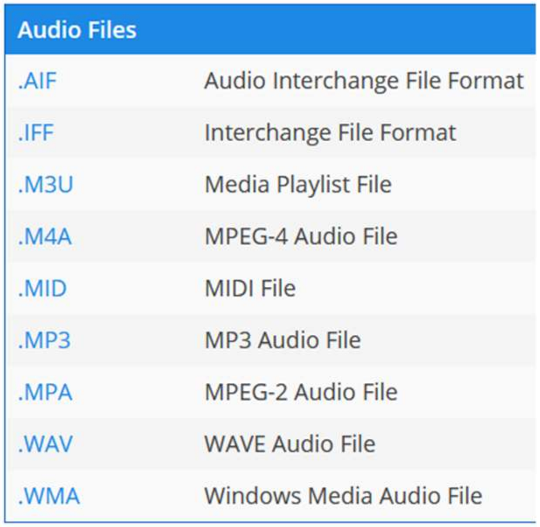
Some Microsoft Official File Type Extensions



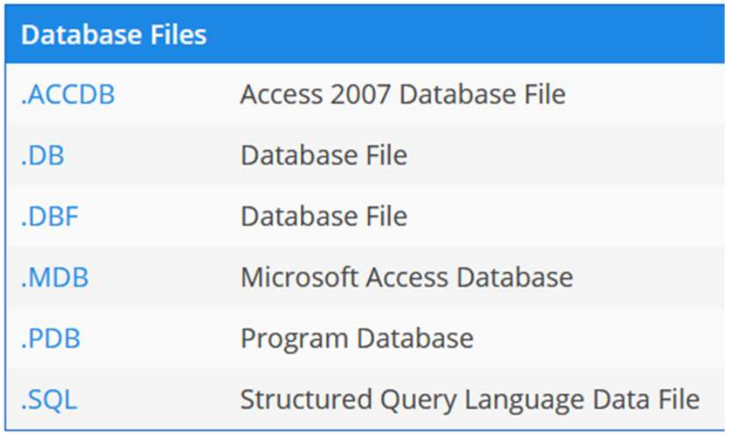
Text Files	
.DOC	Microsoft Word Document (Legacy)
.DOCX	Microsoft Word Document
.LOG	Log File
.MSG	Outlook Message Item File
.ODT	OpenDocument Text Document
.PAGES	Apple Pages Document
.RTF	Rich Text Format File
.TEX	LaTeX Source Document
.TXT	Plain Text File
.WPD	WordPerfect Document
.WPS	Microsoft Works Word Processor Document



Data Files	
.CSV	Comma-Separated Values File
.DAT	Data File
.GED	GEDCOM Genealogy Data File
.KEY	Apple Keynote Presentation
.KEYCHAIN	Mac OS X Keychain File
.PPT	Microsoft PowerPoint Presentation (Legacy)
.PPTX	Microsoft PowerPoint Presentation
.SDF	Standard Data File
.TAR	Consolidated Unix File Archive
.TAX2016	TurboTax 2016 Tax Return
.TAX2020	TurboTax 2020 Tax Return
.VCF	vCard File
.XML	XML File



Audio Files	
.AIF	Audio Interchange File Format
.IFF	Interchange File Format
.M3U	Media Playlist File
.M4A	MPEG-4 Audio File
.MID	MIDI File
.MP3	MP3 Audio File
.MPA	MPEG-2 Audio File
.WAV	WAVE Audio File
.WMA	Windows Media Audio File



Database Files	
.ACCDB	Access 2007 Database File
.DB	Database File
.DBF	Database File
.MDB	Microsoft Access Database
.PDB	Program Database
.SQL	Structured Query Language Data File

This is only a small sample of a more complete list

Such a list is at

https://fileinfo.com/filetypes/common

More Questions

Supposed you change a file extensions to something different or eliminate it completely. What then?

In Linux there are often no file type extensions

How do we ID a file type if we're using Linux?

How does Linux ID a file type?

In earlier versions of Unix and Linux, the user had to open an App and have the App open the file

Some of today's Linux distributions examine the contents of the file itself to try to determine what application to use

One More Question

How do we, as forensic examiners, **<u>correctly</u>** determine a file type?

Suppose in a Windows system, the name extension has been changed or removed

Use file type ***signatures***

Also, other artifacts if needed

File Type Signatures

Almost all file types have ***signatures*** that can be used to identify the type of the file

- *Signatures are located in the first few bytes of a file*
- *Supplementary signature-related information is often located at the end of the file and/or within the file*

File Carving

Sometimes referred to as

Data Carving *or* ***Salvaging***

Introduction

Much of the work of cyber forensics involves being able to do the following from a forensic image of a mass storage device:

Read documents

View pictures, diagrams, tables and videos

Listen to audio from conversations

Know when and where the above were created or recorded

Separate and extract the above from the forensic image

All the above are files that can usually be opened by various applications – but not always

The Problem

But what if you are interested in accessing a file that is:

- *Deleted*
- *Partially overwritten*
- *Type of file changed in the file system?*
- *Don’t have or know the file system*
  - I.e., it comes to you unstructured, like a bunch of bits. There is no file system metadata

What might you have to do?

- *Resort to File Carving*

Some Definitions

File Carving or just ***Carving***

Extracting data from the image of a storage device without the assistance of the file system that originally contained the file

Identifying and recovering files without the use of metadata that sometimes identifies the file

Recovery of files from a digital storage device, especially files that are unrecoverable by conventional means

Reconstructing computer files from file fragments in the absence of file system metadata

File System Metadata

What is the nature of the metadata that file systems separately keep about a file?

File names

In Windows, the file name extension that the OS uses to associate a file with an application

File cluster locations and file size

File cluster fragments

Time and data information

Where is such metadata kept in FAT, NTFS and EXT file systems?

FAT:	*File System **Directory** and **FAT***
NTFS:	*MFT*
EXT:	*Inode tables, Inode Bitmap and Directory contents*

Apps and Their Files

In order to be opened by an application

- *A file must have a known file format*
- *The format is expected by the application*
- *Each format must have*
  - Known bit strings in known file locations (offsets)

Applications won't open a file if the expected strings

- *Don't exist*
- *Aren't where they're supposed to be*

Applications usually look at the file headers

If apps also create files, these files conform to the format

How File Carvers Work

Obtain or create a database of headers and footers (strings of bits at predictable offsets) for many known file types

Old, new and uncommon

Using the DB, search an image for occurrences of the headers and footers

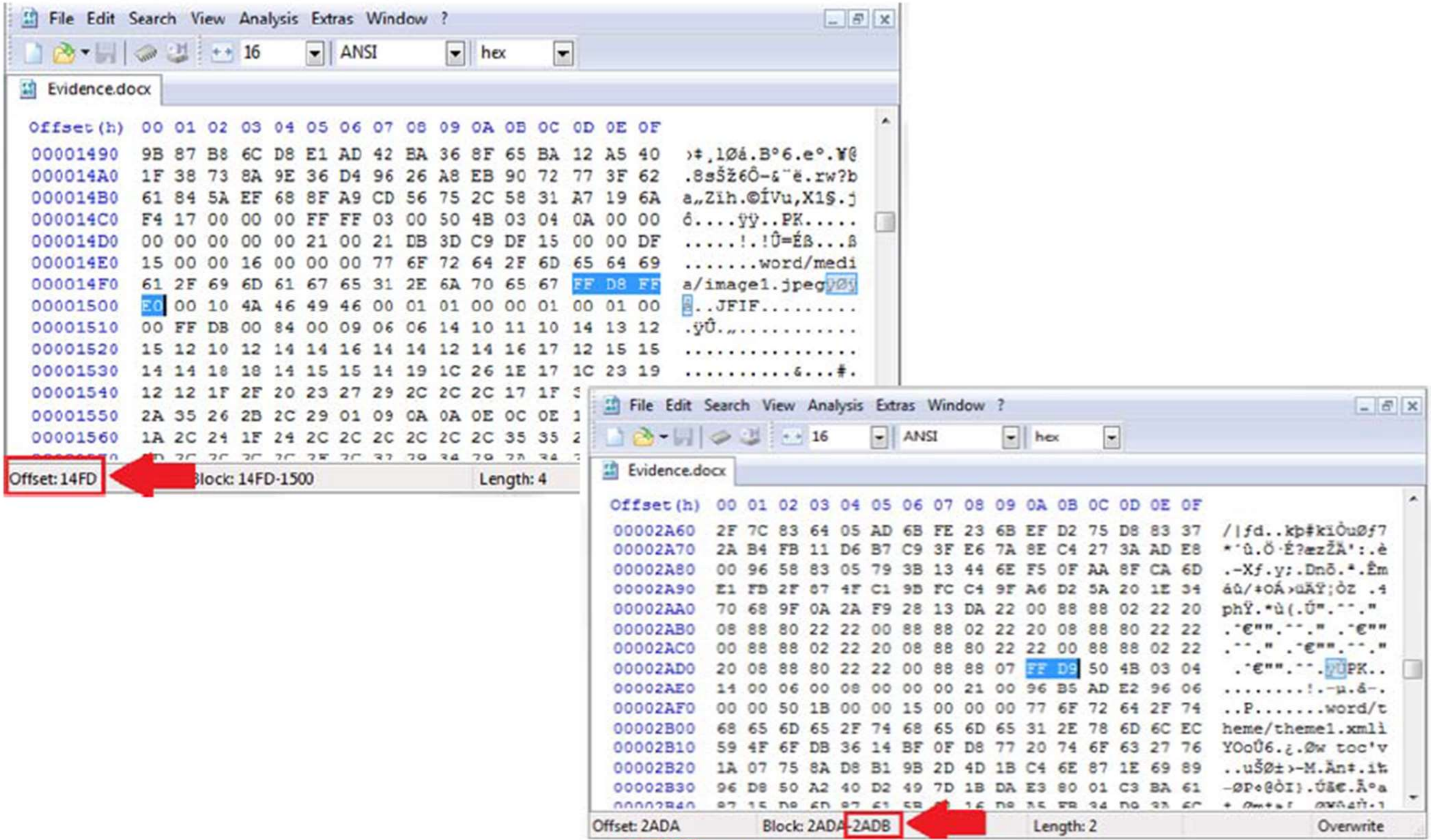
These occurrences might identify the beginning and end of files in the drive image

Retrieve (carve) the files from raw drive images

Regardless of the type of file system in the drive image

The next slides show some known header examples

Searching for a JPEG File



```
Offset(h)  00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F

00001490   9B 87 B8 6C D8 E1 AD 42 BA 36 8F 65 BA 12 A5 40   ›‡¸lØá.B°6.e°.¥@
000014A0   1F 38 73 8A 9E 36 D4 96 26 A8 EB 90 72 77 3F 62   .8sŠž6Ô–&¨ë.rw?b
000014B0   61 84 5A EF 68 8F A9 CD 56 75 2C 58 31 A7 19 6A   a„Zïh.©ÍVu,X1§.j
000014C0   F4 17 00 00 00 FF FF 03 00 50 4B 03 04 0A 00 00   ô....ÿÿ..PK.....
000014D0   00 00 00 00 00 21 00 21 DB 3D C9 DF 15 00 00 DF   .....!.!Û=Éß...ß
000014E0   15 00 00 16 00 00 00 77 6F 72 64 2F 6D 65 64 69   .......word/medi
000014F0   61 2F 69 6D 61 67 65 31 2E 6A 70 65 67 FF D8 FF   a/image1.jpegÿØÿ
00001500   E0 00 10 4A 46 49 46 00 01 01 00 00 01 00 01 00   à..JFIF.........
00001510   00 FF DB 00 84 00 09 06 06 14 10 11 10 14 13 12   .ÿÛ.„...........
00001520   15 12 10 12 14 14 16 14 14 12 14 16 17 12 15 15   ................
00001530   14 14 18 18 14 15 15 14 19 1C 26 1E 17 1C 23 19   ..........&...#.
00001540   12 12 1F 2F 20 23 27 29 2C 2C 2C 17 1F
00001550   2A 35 26 2B 2C 29 01 09 0A 0A 0E 0C 0E
00001560   1A 2C 24 1F 24 2C 2C 2C 2C 2C 2C 35 35
```

Offset: 14FD

Block: 14FD-1500

Length: 4

```
Offset(h)  00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F

00002A60   2F 7C 83 64 05 AD 6B FE 23 6B EF D2 75 D8 83 37   /|ƒd..kþ#kïÒuØƒ7
00002A70   2A B4 FB 11 D6 B7 C9 3F E6 7A 8E C4 27 3A AD E8   *´û.Ö·É?æzŽÄ':.è
00002A80   00 96 58 83 05 79 3B 13 44 6E F5 0F AA 8F CA 6D   .–Xƒ.y;.Dnõ.ª.Êm
00002A90   E1 FB 2F 87 4F C1 9B FC C4 9F A6 D2 5A 20 1E 34   áû/‡OÁ›üÄŸ¦ÒZ .4
00002AA0   70 68 9F 0A 2A F9 28 13 DA 22 00 88 88 02 22 20   phŸ.*ù(.Ú".ˆˆ." 
00002AB0   08 88 80 22 22 00 88 88 02 22 20 08 88 80 22 22   .ˆ€"".ˆˆ." .ˆ€""
00002AC0   00 88 88 02 22 20 08 88 80 22 22 00 88 88 02 22   .ˆˆ." .ˆ€"".ˆˆ."
00002AD0   20 08 88 80 22 22 00 88 88 07 FF D9 50 4B 03 04    .ˆ€"".ˆˆ.ÿÙPK..
00002AE0   14 00 06 00 08 00 00 00 21 00 96 B5 AD E2 96 06   ........!.–µ.â–.
00002AF0   00 00 50 1B 00 00 15 00 00 00 77 6F 72 64 2F 74   ..P.......word/t
00002B00   68 65 6D 65 2F 74 68 65 6D 65 31 2E 78 6D 6C EC   heme/theme1.xmlì
00002B10   59 4F 6F DB 36 14 BF 0F D8 77 20 74 6F 63 27 76   YOoÛ6.¿.Øw toc'v
00002B20   1A 07 75 8A D8 B1 9B 2D 4D 1B C4 6E 87 1E 69 89   ..uŠØ±›-M.Än‡.i‰
00002B30   96 D8 50 A2 40 D2 49 7D 1B DA E3 80 01 C3 BA 61   –ØP¢@ÒI}.Úã€.Ãºa
```

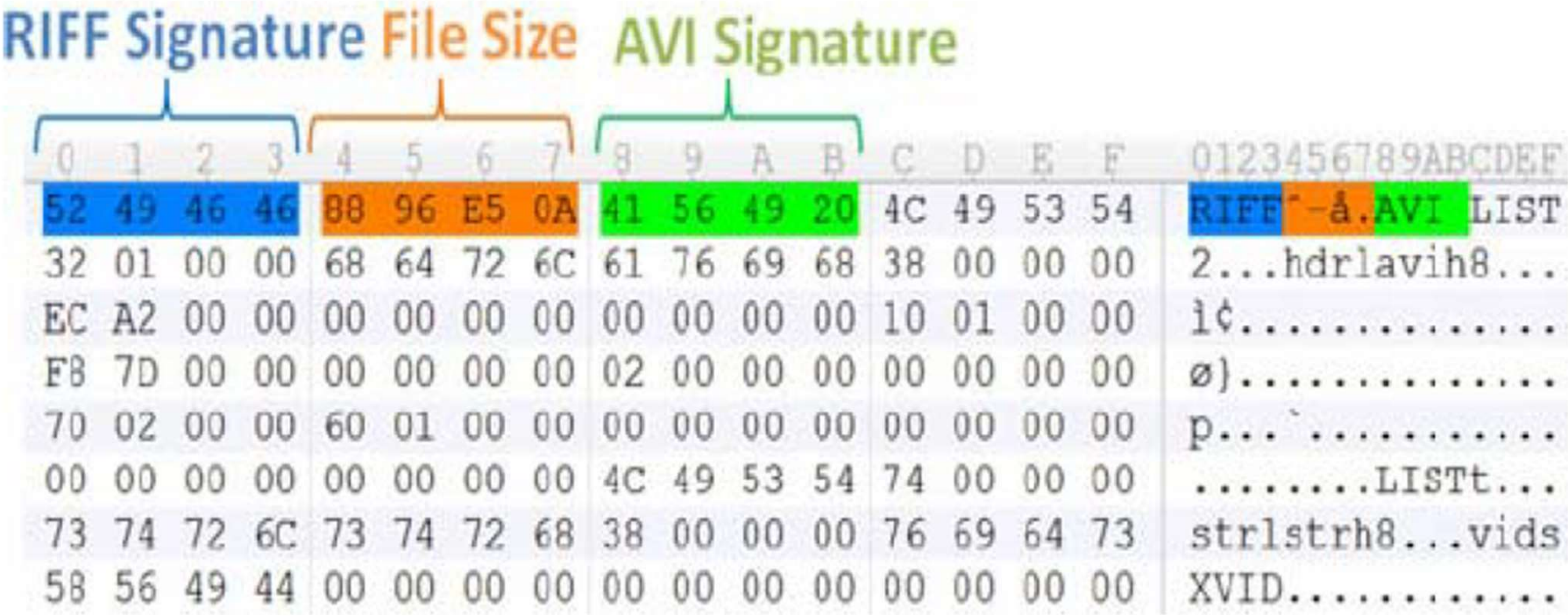
Offset: 2ADA

Block: 2ADA-2ADB

Length: 2

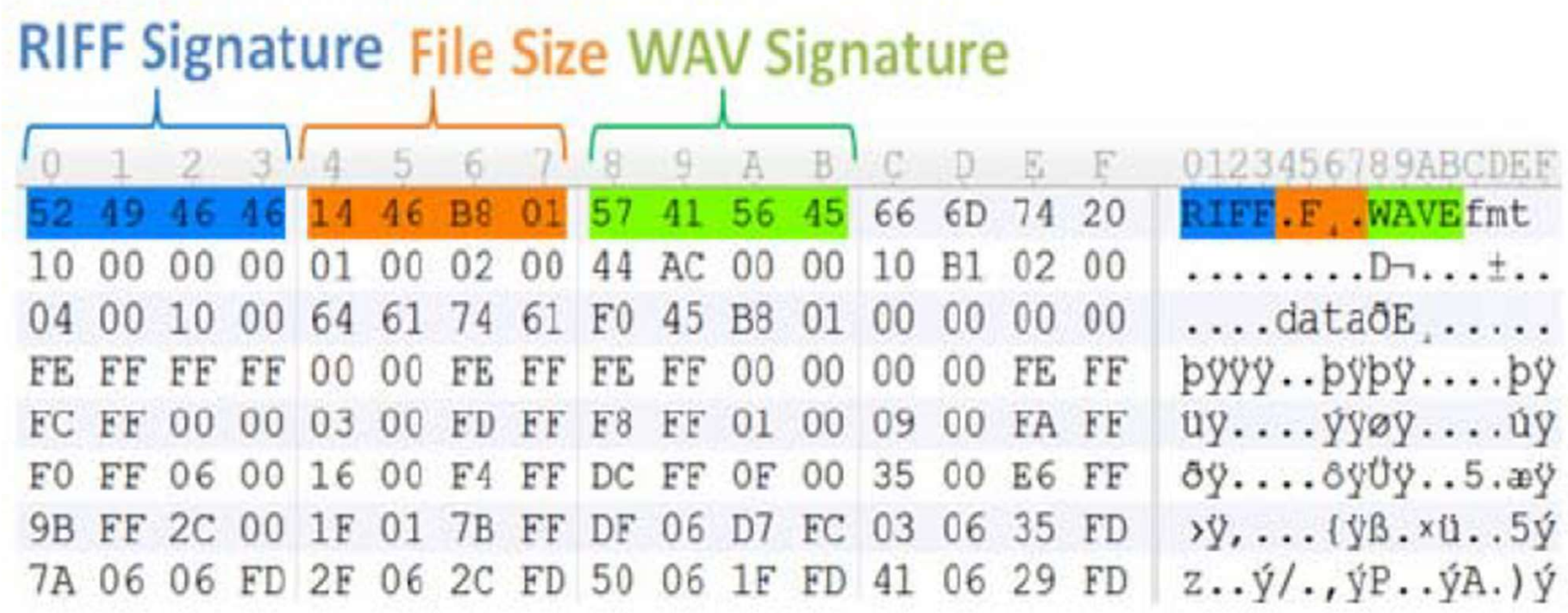
Overwrite

AVI Header & File Length



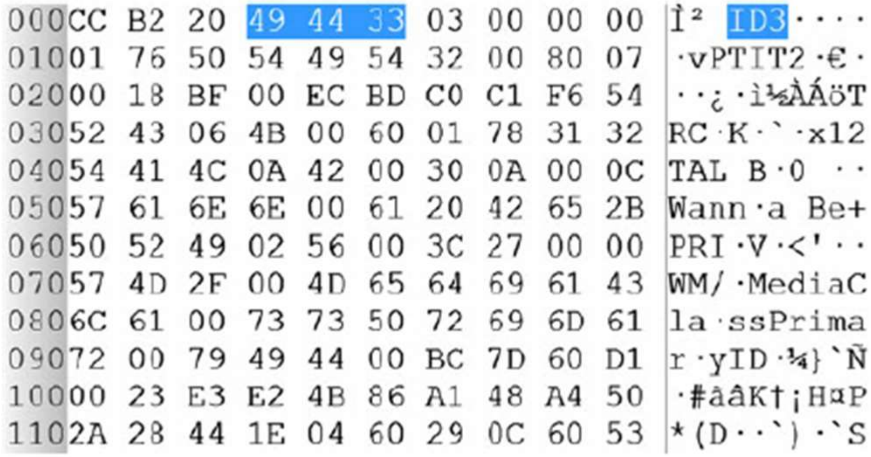
```
 0  1  2  3  4  5  6  7  8  9  A  B  C  D  E  F   0123456789ABCDEF
52 49 46 46 88 96 E5 0A 41 56 49 20 4C 49 53 54   RIFFˆ–å.AVI LIST
32 01 00 00 68 64 72 6C 61 76 69 68 38 00 00 00   2...hdrlavih8...
EC A2 00 00 00 00 00 00 00 00 00 00 10 01 00 00   ì¢..............
F8 7D 00 00 00 00 00 00 02 00 00 00 00 00 00 00   ø}..............
70 02 00 00 60 01 00 00 00 00 00 00 00 00 00 00   p...`...........
00 00 00 00 00 00 00 00 4C 49 53 54 74 00 00 00   ........LISTt...
73 74 72 6C 73 74 72 68 38 00 00 00 76 69 64 73   strlstrh8...vids
58 56 49 44 00 00 00 00 00 00 00 00 00 00 00 00   XVID............
```

WAV Header & File Length



```
 0  1  2  3  4  5  6  7  8  9  A  B  C  D  E  F   0123456789ABCDEF
52 49 46 46 14 46 B8 01 57 41 56 45 66 6D 74 20   RIFF.F¸.WAVEfmt
10 00 00 00 01 00 02 00 44 AC 00 00 10 B1 02 00   ........D¬...±..
04 00 10 00 64 61 74 61 F0 45 B8 01 00 00 00 00   ....dataðE¸.....
FE FF FF FF 00 00 FE FF FE FF 00 00 00 00 FE FF   þÿÿÿ..þÿþÿ....þÿ
FC FF 00 00 03 00 FD FF F8 FF 01 00 09 00 FA FF   üÿ....ýÿøÿ....úÿ
F0 FF 06 00 16 00 F4 FF DC FF 0F 00 35 00 E6 FF   ðÿ....ôÿÜÿ..5.æÿ
9B FF 2C 00 1F 01 7B FF DF 06 D7 FC 03 06 35 FD   ›ÿ,...{ÿß.×ü..5ý
7A 06 06 FD 2F 06 2C FD 50 06 1F FD 41 06 29 FD   z..ý/.,ýP..ýA.)ý
```

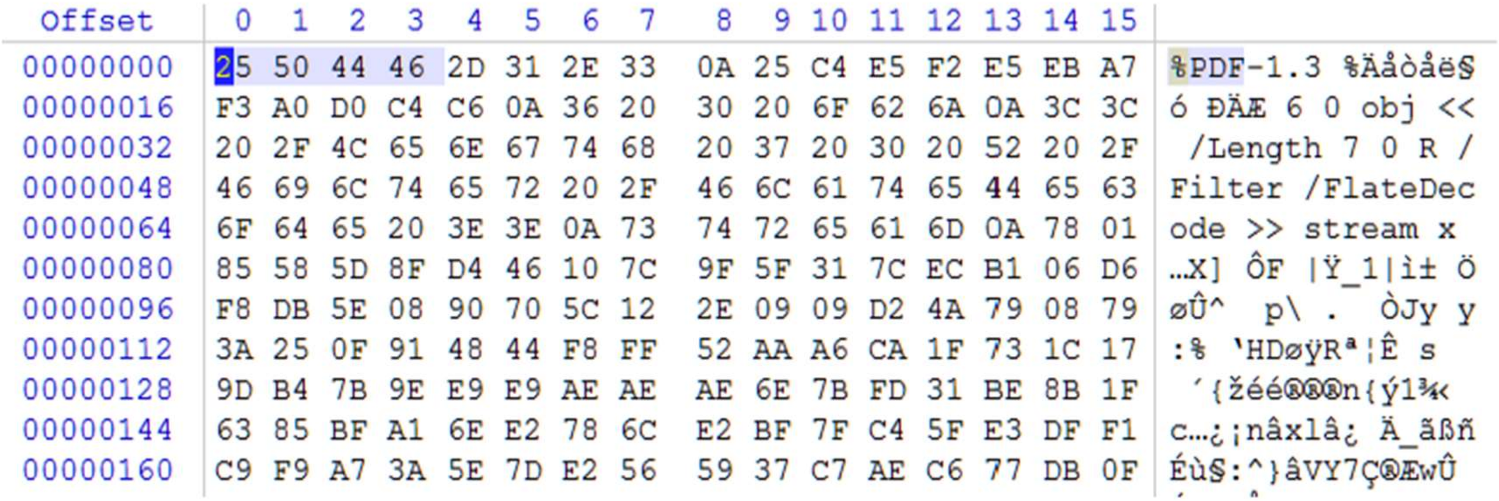
What Have These Searches Found?



```
000CC B2 20 49 44 33 03 00 00 00 Ì²  ID3····
01001 76 50 54 49 54 32 00 80 07  ·vPTIT2·€·
02000 18 BF 00 EC BD C0 C1 F6 54  ··¿·ì½ÀÁöT
03052 43 06 4B 00 60 01 78 31 32 RC·K·`·x12
04054 41 4C 0A 42 00 30 0A 00 0C TAL B·0  ··
05057 61 6E 6E 00 61 20 42 65 2B Wann·a Be+
06050 52 49 02 56 00 3C 27 00 00 PRI·V·<'··
07057 4D 2F 00 4D 65 64 69 61 43 WM/·MediaC
0806C 61 00 73 73 50 72 69 6D 61 la·ssPrima
09072 00 79 49 44 00 BC 7D 60 D1 r·yID·¼}`Ñ
10000 23 E3 E2 4B 86 A1 48 A4 50  ·#ãâK†¡H¤P
1102A 28 44 1E 04 60 29 0C 60 53 *(D··`)·`S
```

MP3

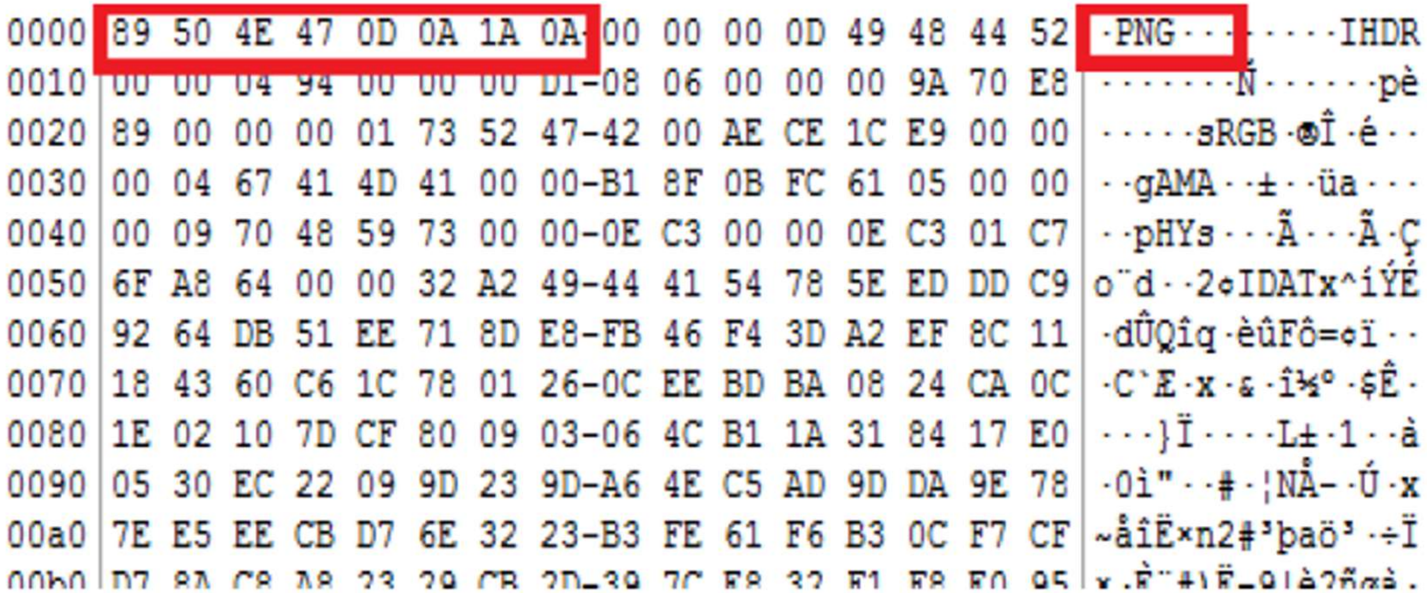
There are other possible MP3 headers



```
Offset     0  1  2  3  4  5  6  7   8  9 10 11 12 13 14 15
00000000  25 50 44 46 2D 31 2E 33  0A 25 C4 E5 F2 E5 EB A7  %PDF-1.3 %ÄåòåëS
00000016  F3 A0 D0 C4 C6 0A 36 20  30 20 6F 62 6A 0A 3C 3C  ó ÐÄÆ 6 0 obj <<
00000032  20 2F 4C 65 6E 67 74 68  20 37 20 30 20 52 20 2F   /Length 7 0 R /
00000048  46 69 6C 74 65 72 20 2F  46 6C 61 74 65 44 65 63  Filter /FlateDec
00000064  6F 64 65 20 3E 3E 0A 73  74 72 65 61 6D 0A 78 01  ode >> stream x
00000080  85 58 5D 8F D4 46 10 7C  9F 5F 31 7C EC B1 06 D6  …X] ÔF |Ÿ_1|ì± Ö
00000096  F8 DB 5E 08 90 70 5C 12  2E 09 09 D2 4A 79 08 79  øÛ^  p\ .  ÒJy y
00000112  3A 25 0F 91 48 44 F8 FF  52 AA A6 CA 1F 73 1C 17  :% 'HDøÿRª¦Ê s
00000128  9D B4 7B 9E E9 E9 AE AE  AE 6E 7B FD 31 BE 8B 1F   ´{žéé®®®n{ý1¾‹
00000144  63 85 BF A1 6E E2 78 6C  E2 BF 7F C4 5F E3 DF F1  c…¿¡nâxlâ¿ Ä_ãßñ
00000160  C9 F9 A7 3A 5E 7D E2 56  59 37 C7 AE C6 77 DB 0F  Éù§:^}âVY7Ç®ÆwÛ
```

PDF

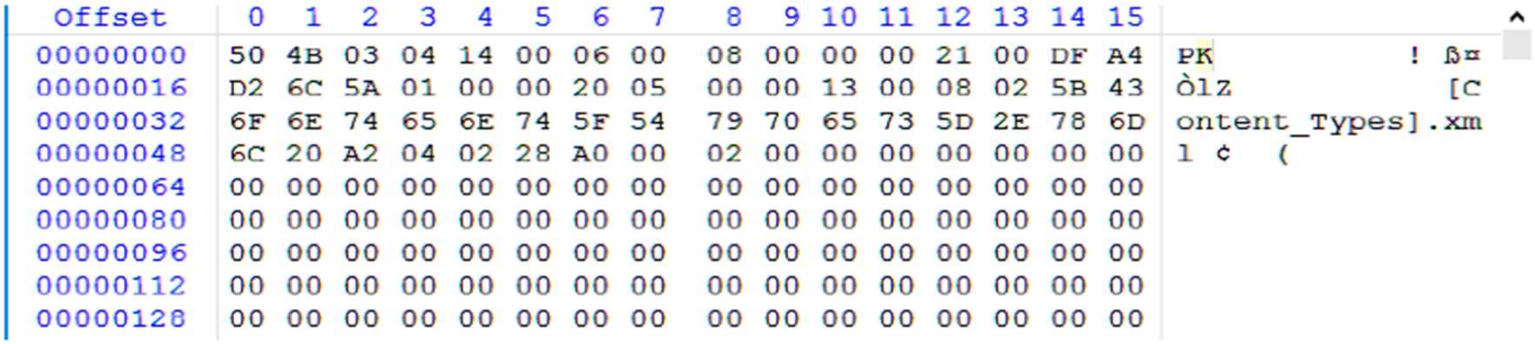
How About This Search?



```
0000 89 50 4E 47 0D 0A 1A 0A-00 00 00 0D 49 48 44 52 ·PNG········IHDR
0010 00 00 04 94 00 00 00 D1-08 06 00 00 00 9A 70 E8 ·······Ñ······pè
0020 89 00 00 00 01 73 52 47-42 00 AE CE 1C E9 00 00 ·····sRGB·®Î·é··
0030 00 04 67 41 4D 41 00 00-B1 8F 0B FC 61 05 00 00 ··gAMA··±··üa···
0040 00 09 70 48 59 73 00 00-0E C3 00 00 0E C3 01 C7 ··pHYs···Ã···Ã·Ç
0050 6F A8 64 00 00 32 A2 49-44 41 54 78 5E ED DD C9 o¨d··2¢IDATx^íÝÉ
0060 92 64 DB 51 EE 71 8D E8-FB 46 F4 3D A2 EF 8C 11 ·dÛQîq·èûFô=¢ï··
0070 18 43 60 C6 1C 78 01 26-0C EE BD BA 08 24 CA 0C ·C`Æ·x·&·î½º·$Ê·
0080 1E 02 10 7D CF 80 09 03-06 4C B1 1A 31 84 17 E0 ···}Ï····L±·1··à
0090 05 30 EC 22 09 9D 23 9D-A6 4E C5 AD 9D DA 9E 78 ·0ì"··#·¦NÅ-·Ú·x
00a0 7E E5 EE CB D7 6E 32 23-B3 FE 61 F6 B3 0C F7 CF ~åîË×n2#³þaö³·÷Ï
00b0 D7 8A C8 A8 23 29 CB 2D-39 7C E8 32 F1 F8 F0 95 ×·È¨#)Ë-9|è2ñøð·
```

PNG: Portable Network Graphics

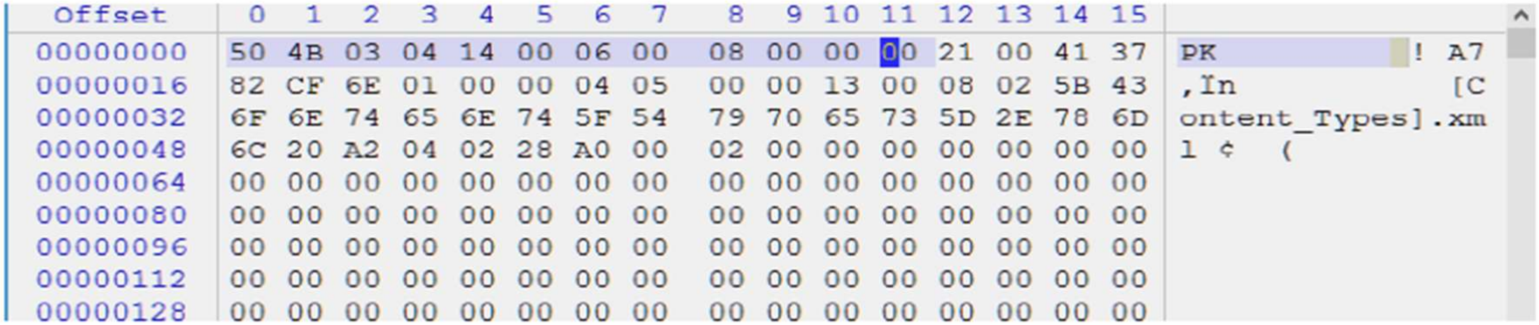
What Are These?



```
Offset     0  1  2  3  4  5  6  7   8  9 10 11 12 13 14 15
00000000  50 4B 03 04 14 00 06 00  08 00 00 00 21 00 DF A4   PK          ! ß¤
00000016  D2 6C 5A 01 00 00 20 05  00 00 13 00 08 02 5B 43   ÒlZ            [C
00000032  6F 6E 74 65 6E 74 5F 54  79 70 65 73 5D 2E 78 6D   ontent_Types].xm
00000048  6C 20 A2 04 02 28 A0 00  02 00 00 00 00 00 00 00   l ¢  (
00000064  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
00000080  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
00000096  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
00000112  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
00000128  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
```

MS Word
(.docx)

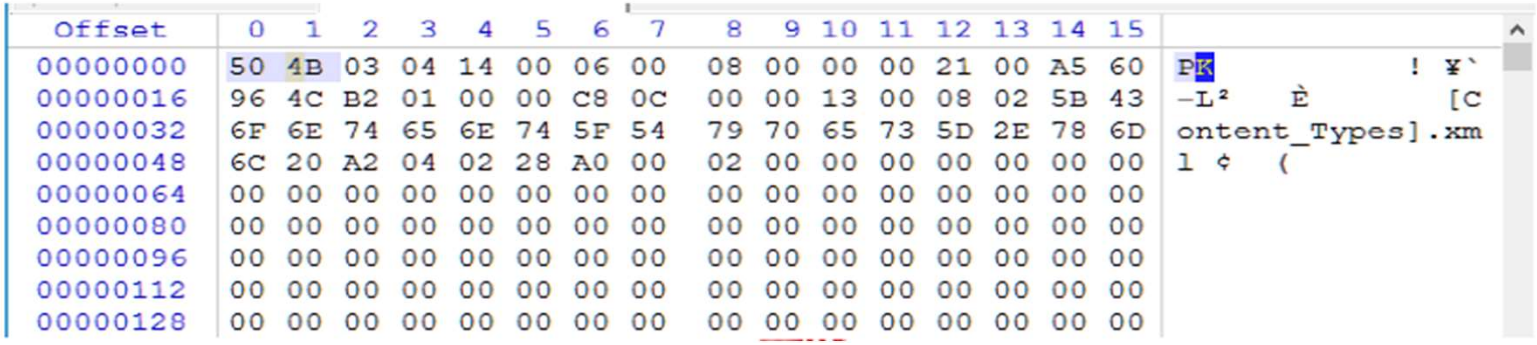
0x504b0304140006…



```
Offset     0  1  2  3  4  5  6  7   8  9 10 11 12 13 14 15
00000000  50 4B 03 04 14 00 06 00  08 00 00 00 21 00 41 37   PK          ! A7
00000016  82 CF 6E 01 00 00 04 05  00 00 13 00 08 02 5B 43   ,In            [C
00000032  6F 6E 74 65 6E 74 5F 54  79 70 65 73 5D 2E 78 6D   ontent_Types].xm
00000048  6C 20 A2 04 02 28 A0 00  02 00 00 00 00 00 00 00   l ¢  (
00000064  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
00000080  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
00000096  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
00000112  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
00000128  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
```

MS Excel
(.xlsx)

0x504b0304140006…



```
Offset     0  1  2  3  4  5  6  7   8  9 10 11 12 13 14 15
00000000  50 4B 03 04 14 00 06 00  08 00 00 00 21 00 A5 60   PK          ! ¥`
00000016  96 4C B2 01 00 00 C8 0C  00 00 13 00 08 02 5B 43   –L²    È       [C
00000032  6F 6E 74 65 6E 74 5F 54  79 70 65 73 5D 2E 78 6D   ontent_Types].xm
00000048  6C 20 A2 04 02 28 A0 00  02 00 00 00 00 00 00 00   l ¢  (
00000064  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
00000080  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
00000096  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
00000112  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
00000128  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
```

MS Powerpoint
(.pptx)

0x504b0304140006…

Other Headers

BMP file:	**BM.**	or	**0x424D2E5C**
	BM	or	**0x424D**
GIF file:	**GIF**	or	**0x474946**

There are many other known file headers

Some Other Carving Capabilities

File carving can be done if the file system metadata is unavailable

Importantly, file carving can often be done even if:

- *File parts have been overwritten or changed*
- *The file is fragmented*

But the above probably requires some manual intervention

Carving is Great!

File Carving is great. Right?

Just turn a file carver loose on a drive image.

Go to lunch.

When you return, a nice report is ready for you.

Comments?

Issues

You are given an apparently intact 1TB drive from a computer that was mostly destroyed

- *You image the drive and find that it was from an NTFS file system. How?*
- *But the MFT and its backup are corrupted*
- *So you have to carve*

Cool! You need to do some shopping during lunch anyway

Issues

But in order to carve while you're at lunch, the files need to be contiguous

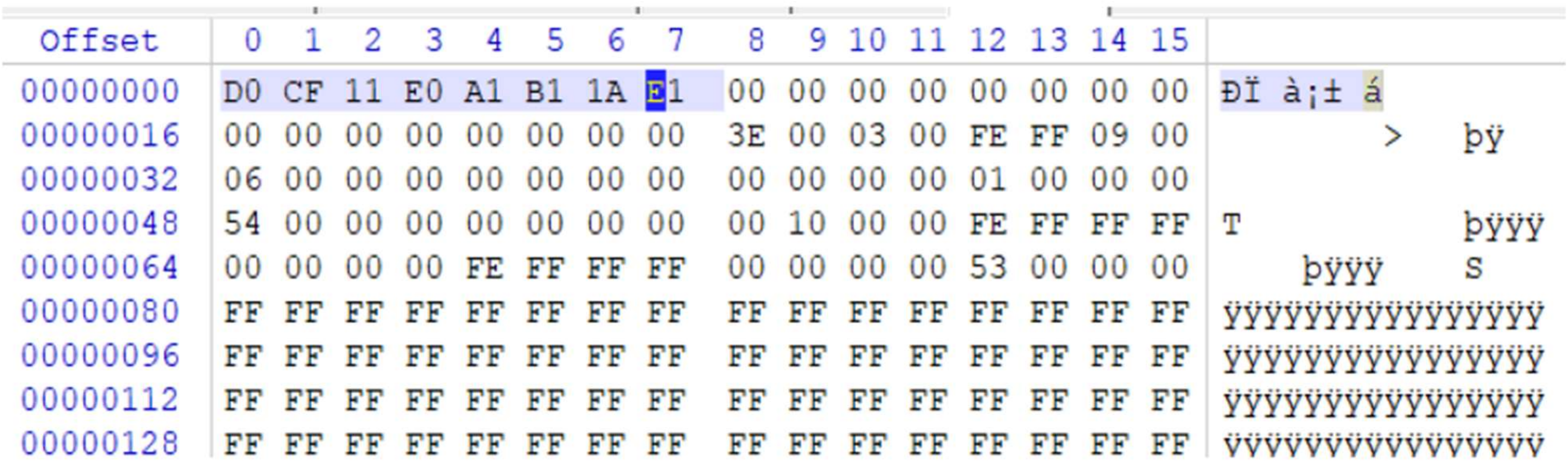
Otherwise you've got work to do

And carving may require searching the entire 1 TB drive image multiple times

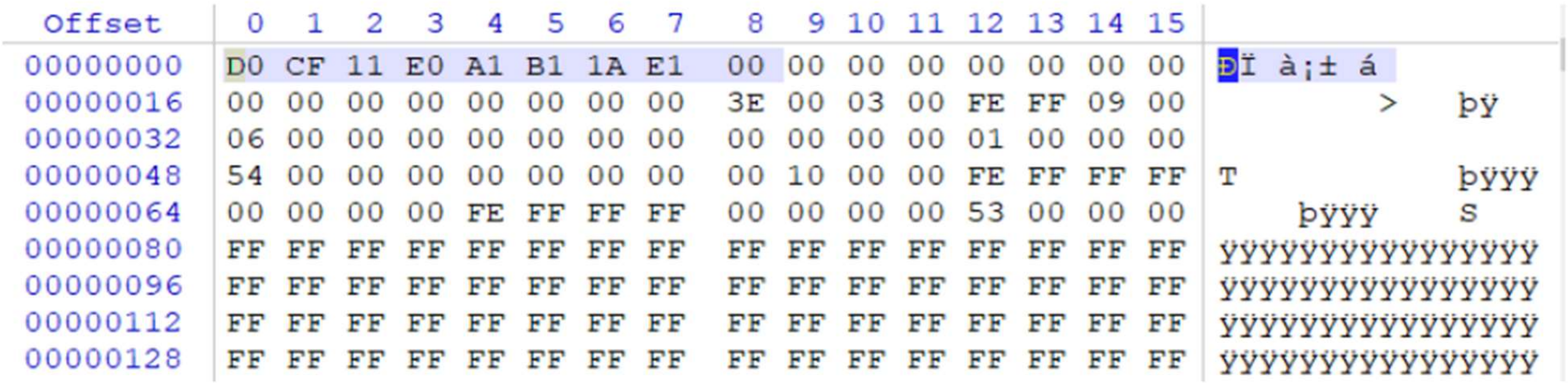
Make it a looooonnnggg lunch

Maybe even a weekend or holiday

More Signature Examples

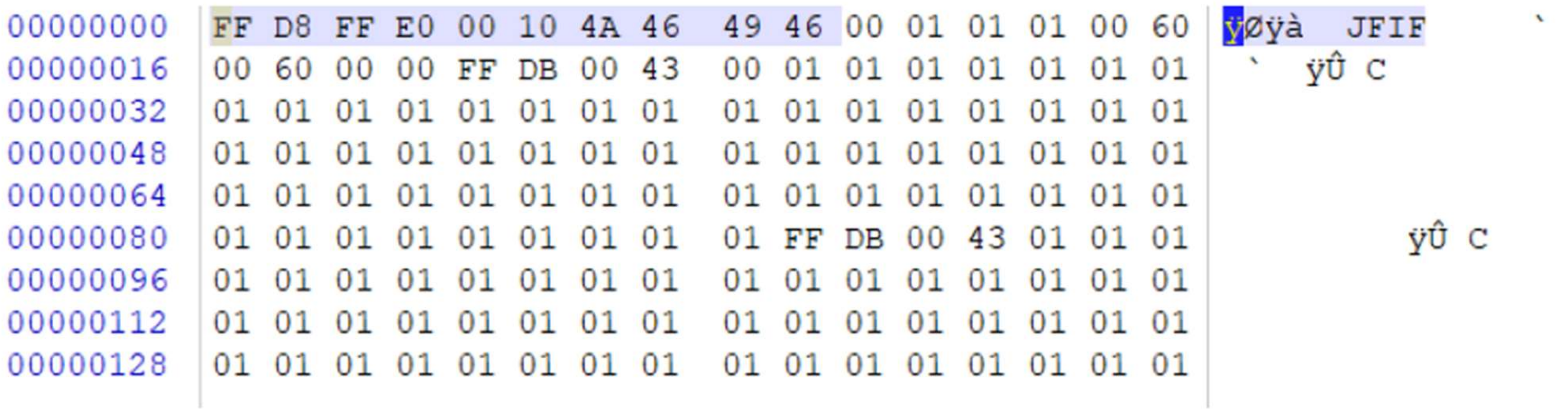


```
Offset     0  1  2  3  4  5  6  7   8  9 10 11 12 13 14 15
00000000  D0 CF 11 E0 A1 B1 1A E1  00 00 00 00 00 00 00 00  ÐÏ à¡± á
00000016  00 00 00 00 00 00 00 00  3E 00 03 00 FE FF 09 00          >   þÿ
00000032  06 00 00 00 00 00 00 00  00 00 00 00 01 00 00 00
00000048  54 00 00 00 00 00 00 00  00 10 00 00 FE FF FF FF  T           þÿÿÿ
00000064  00 00 00 00 FE FF FF FF  00 00 00 00 53 00 00 00      þÿÿÿ    S
00000080  FF FF FF FF FF FF FF FF  FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
00000096  FF FF FF FF FF FF FF FF  FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
00000112  FF FF FF FF FF FF FF FF  FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
00000128  FF FF FF FF FF FF FF FF  FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
```

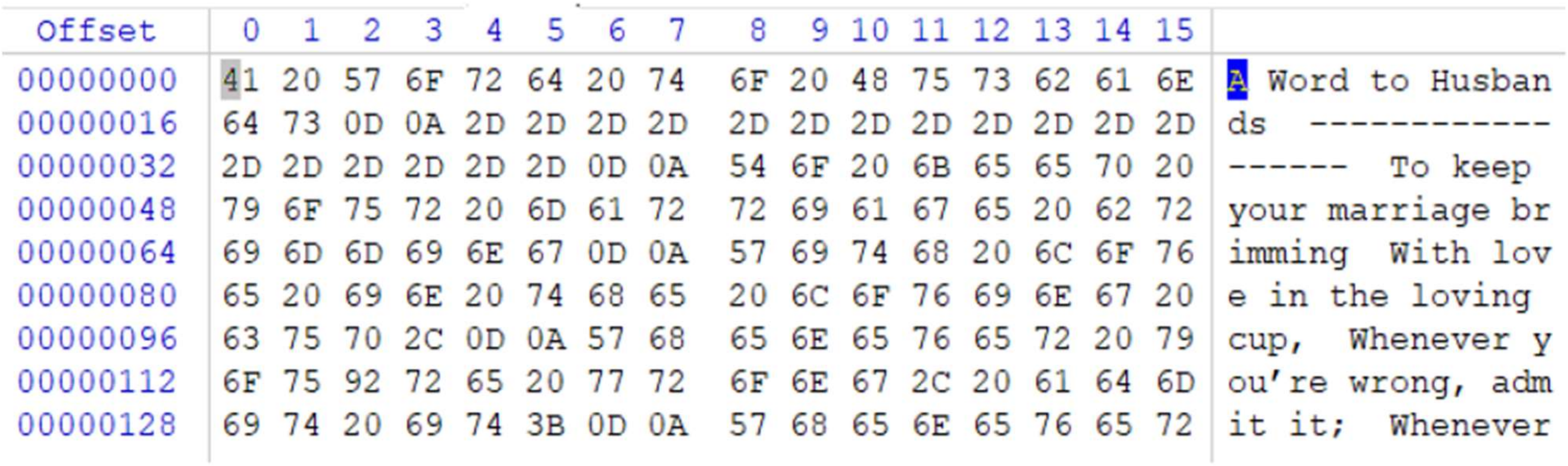


```
Offset     0  1  2  3  4  5  6  7   8  9 10 11 12 13 14 15
00000000  D0 CF 11 E0 A1 B1 1A E1  00 00 00 00 00 00 00 00  ÐÏ à¡± á
00000016  00 00 00 00 00 00 00 00  3E 00 03 00 FE FF 09 00          >   þÿ
00000032  06 00 00 00 00 00 00 00  00 00 00 00 01 00 00 00
00000048  54 00 00 00 00 00 00 00  00 10 00 00 FE FF FF FF  T           þÿÿÿ
00000064  00 00 00 00 FE FF FF FF  00 00 00 00 53 00 00 00      þÿÿÿ    S
00000080  FF FF FF FF FF FF FF FF  FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
00000096  FF FF FF FF FF FF FF FF  FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
00000112  FF FF FF FF FF FF FF FF  FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
00000128  FF FF FF FF FF FF FF FF  FF FF FF FF FF FF FF FF  ÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿÿ
```

More Signature Examples

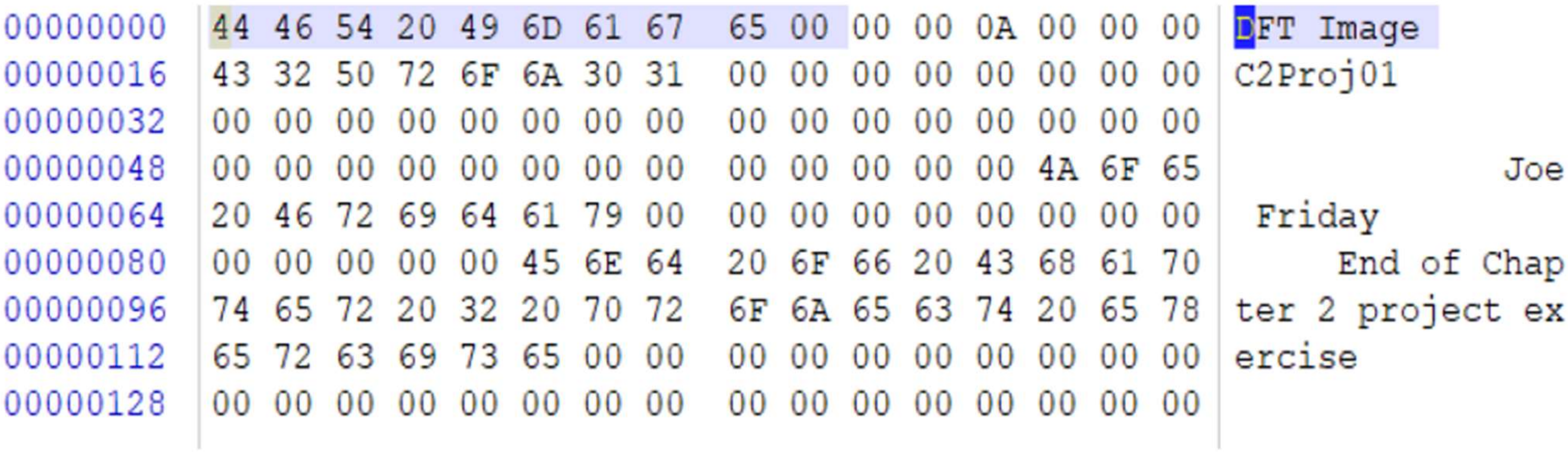


```
00000000  FF D8 FF E0 00 10 4A 46  49 46 00 01 01 01 00 60  ÿØÿà  JFIF      `
00000016  00 60 00 00 FF DB 00 43  00 01 01 01 01 01 01 01   `  ÿÛ C
00000032  01 01 01 01 01 01 01 01  01 01 01 01 01 01 01 01
00000048  01 01 01 01 01 01 01 01  01 01 01 01 01 01 01 01
00000064  01 01 01 01 01 01 01 01  01 01 01 01 01 01 01 01
00000080  01 01 01 01 01 01 01 01  01 FF DB 00 43 01 01 01           ÿÛ C
00000096  01 01 01 01 01 01 01 01  01 01 01 01 01 01 01 01
00000112  01 01 01 01 01 01 01 01  01 01 01 01 01 01 01 01
00000128  01 01 01 01 01 01 01 01  01 01 01 01 01 01 01 01
```



```
  Offset   0  1  2  3  4  5  6  7   8  9 10 11 12 13 14 15
00000000  41 20 57 6F 72 64 20 74  6F 20 48 75 73 62 61 6E  A Word to Husban
00000016  64 73 0D 0A 2D 2D 2D 2D  2D 2D 2D 2D 2D 2D 2D 2D  ds  ------------
00000032  2D 2D 2D 2D 2D 2D 0D 0A  54 6F 20 6B 65 65 70 20  ------  To keep 
00000048  79 6F 75 72 20 6D 61 72  72 69 61 67 65 20 62 72  your marriage br
00000064  69 6D 6D 69 6E 67 0D 0A  57 69 74 68 20 6C 6F 76  imming  With lov
00000080  65 20 69 6E 20 74 68 65  20 6C 6F 76 69 6E 67 20  e in the loving 
00000096  63 75 70 2C 0D 0A 57 68  65 6E 65 76 65 72 20 79  cup,  Whenever y
00000112  6F 75 92 72 65 20 77 72  6F 6E 67 2C 20 61 64 6D  ou’re wrong, adm
00000128  69 74 20 69 74 3B 0D 0A  57 68 65 6E 65 76 65 72  it it;  Whenever
```

More Signature Examples



```
00000000  44 46 54 20 49 6D 61 67  65 00 00 00 0A 00 00 00  DFT Image
00000016  43 32 50 72 6F 6A 30 31  00 00 00 00 00 00 00 00  C2Proj01
00000032  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
00000048  00 00 00 00 00 00 00 00  00 00 00 00 00 4A 6F 65               Joe
00000064  20 46 72 69 64 61 79 00  00 00 00 00 00 00 00 00   Friday
00000080  00 00 00 00 00 45 6E 64  20 6F 66 20 43 68 61 70       End of Chap
00000096  74 65 72 20 32 20 70 72  6F 6A 65 63 74 20 65 78  ter 2 project ex
00000112  65 72 63 69 73 65 00 00  00 00 00 00 00 00 00 00  ercise
00000128  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00
```

List of File Signatures

Wikipedia:

- *https://en.wikipedia.org/wiki/List_of_file_signatures*
- *Not complete, but it does have most of the more common file sigatures*

Filesignatures.net

- *https://www.filesignatures.net/*
- *Give pretty comprehensive list of signatures*
- *Can input signature and get file type(s)*
- *Can input file type and get signature(s)*

Data Carving
What Does it Do

Searches for signatures in unknown data units that correspond to the beginning and end of known file types

Often is used on unallocated data units in order to recover files that do not have metadata structures pointing to them

Used to recover files based on their headers, footers, and internal data structures

A File Carver

One of the most well-known file carvers is ***Scalpel***

- *Developed for Linux*
- *Ported to Windows*

Contains a database of file signatures

- *Headers, footers, and other info that can be used to ID a file type*

Other Carving Tools

Linux

foremost : Heavy duty carving based upon signatures

- *Analyzes entire file system (raw or image)*
- *Signatures contain*
  - Known header info
  - Header case sensitivity
  - Usual file name extensions
  - Max. file size
  - Known footer info

Other Carving Tools

TSK Application Category

`file` : Can identify the structure on many unknown files

- *Based upon a self-contained database of signature values*
- *Sort of a lightweight carving tool*

`lazarus` : Processes an entire file system image, executing **`file`** on each sector

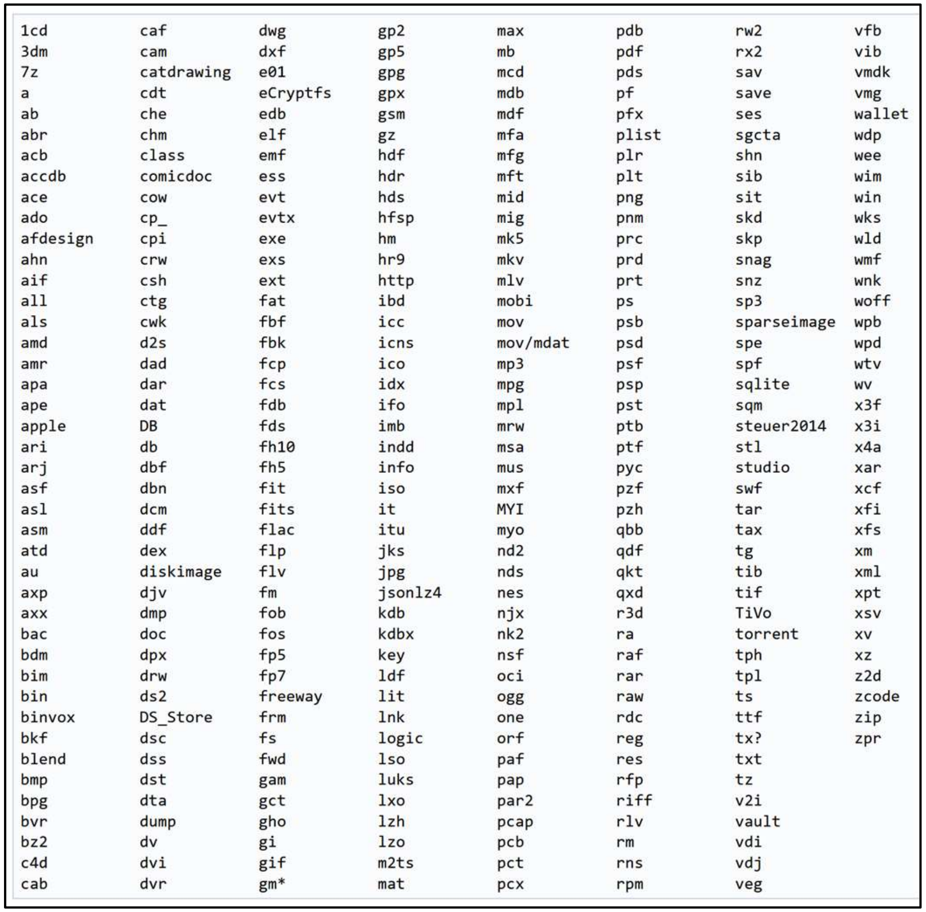
- *Contiguous sectors having the same signature values are grouped*
- *Lists each sector or group and its signature value*

Other Carving Tools

Platform-independent

Autopsy

- *Most of the file carving capability comes from the PhotoRec Carver ingest module*
  - It does much more than carve graphics images
  - It can be customized to add new file signatures
- *List of files scanned by default*



```
1cd
3dm
7z
a
ab
abr
acb
accdb
ace
ado
afdesign
ahn
aif
all
als
amd
amr
apa
ape
apple
ari
arj
asf
asl
asm
atd
au
axp
axx
bac
bdm
bim
bin
binvox
bkf
blend
bmp
bpg
bvr
bz2
c4d
cab
caf
cam
catdrawing
cdt
che
chm
class
comicdoc
cow
cp_
cpi
crw
csh
ctg
cwk
d2s
dad
dar
dat
DB
db
dbf
dbn
dcm
ddf
dex
diskimage
djv
dmp
doc
dpx
drw
ds2
DS_Store
dsc
dss
dst
dta
dump
dv
dvi
dvr
dwg
dxf
e01
eCryptfs
edb
elf
emf
ess
evt
evtx
exe
exs
ext
fat
fbf
fbk
fcp
fcs
fdb
fds
fh10
fh5
fit
fits
flac
flp
flv
fm
fob
fos
fp5
fp7
freeway
frm
fs
fwd
gam
gct
gho
gi
gif
gm*
gp2
gp5
gpg
gpx
gsm
gz
hdf
hdr
hds
hfsp
hm
hr9
http
ibd
icc
icns
ico
idx
ifo
imb
indd
info
iso
it
itu
jks
jpg
jsonlz4
kdb
kdbx
key
ldf
lit
lnk
logic
lso
luks
lxo
lzh
lzo
m2ts
mat
max
mb
mcd
mdb
mdf
mfa
mfg
mft
mid
mig
mk5
mkv
mlv
mobi
mov
mov/mdat
mp3
mpg
mpl
mrw
msa
mus
mxf
MYI
myo
nd2
nds
nes
njx
nk2
nsf
oci
ogg
one
orf
paf
pap
par2
pcap
pcb
pct
pcx
pdb
pdf
pds
pf
pfx
plist
plr
plt
png
pnm
prc
prd
prt
ps
psb
psd
psf
psp
pst
ptb
ptf
pyc
pzf
pzh
qbb
qdf
qkt
qxd
r3d
ra
raf
rar
raw
rdc
reg
res
rfp
riff
rlv
rm
rns
rpm
rw2
rx2
sav
save
ses
sgcta
shn
sib
sit
skd
skp
snag
snz
sp3
sparseimage
spe
spf
sqlite
sqm
steuer2014
stl
studio
swf
tar
tax
tg
tib
tif
TiVo
torrent
tph
tpl
ts
ttf
tx?
txt
tz
v2i
vault
vdi
vdj
veg
vfb
vib
vmdk
vmg
wallet
wdp
wee
wim
win
wks
wld
wmf
wnk
woff
wpb
wpd
wtv
wv
x3f
x3i
x4a
xar
xcf
xfi
xfs
xm
xml
xpt
xsv
xv
xz
z2d
zcode
zip
zpr
```

File Carving Lab

We will examine the unallocated region of a disk image with four different file carving tools and compare the results:

Kali Linux

- foremost*
- scalpel
- magicrescue

Windows 10

- Autopsy

* Needs to be installed first

File Carving Lab

Prerequisite Activities:

Copy over disk image to your Documents directory

Image location:

`R:\Share\Labs\File\File Carving Lab`

File name:

`L0_Graphic.dd`

Install foremost on Kali Linux

File Carving in Kali Linux

foremost

scalpel

magicrescue

File Carving

foremost

Log onto Kali Linux through VMWorkstation Pro:

Username: kali

Password: kali

Open a terminal window

Try to run foremost by typing in “foremost”

It will prompt you with the correct command to install

Follow the directions to install foremost

File Carving

foremost

Once foremost is installed, we need to update the configuration file so that it will carve the files we want

Configuration file location:

/etc

Configuration file name:

foremost.conf

Specifically, we'll need to comment out the lines in the file that correspond to the files we want foremost to find

We're looking for graphics file types:

jpg, png, bmp, gif, tif, pcx

Here are a few command lines for reference:

Editing the configuration file:

```
sudo mousepad foremost.conf
```

Running foremost:

```
foremost –v –i ./L0_Graphic.dd –o ./foremost_recov
```

Follow the activities in class to complete this lab

File Carving

scalpel

We need to update the scalpel configuration file so that it will carve the files we want

Configuration file location:

/etc/scalpel

Configuration file name:

scalpel.conf

Specifically, we'll need to comment out the lines in the file that correspond to the files we want foremost to find

We're looking for graphics file types:

jpg, png, bmp, gif, tif, pcx

Here are a few command lines for reference:

Editing the configuration file:

```
sudo mousepad scalpel.conf
```

Running scalpel:

```
scalpel –b –v –o ./scalpel_recov L0_graphic.dd
```

Follow the activities in class to complete this lab

File Carving

magicrescue

magicrescue uses recipes that provide instructions for carving different file types

The recipes are located in the following directory:

/usr/share/magicrescue/recipes

We'll use the following recipes:

jpg-jfif, jpeg-exit, png, canon-cr2, gimp-xcf

Here are a few command lines for reference:

Running magicrescus:

```
magicrescue –r jpeg-jfif –r jpeg-exif –r png –r canon-cr2  -r
  nikon-raw –d ./magicrescue_recov L0_Grapic.dd
```

Follow the activities in class to complete this lab

Autopsy for File Carving

File Carving

Autopsy

On your Windows 10 desktop:

- *Open Autopsy*
- *Start a new case*
- *Select the L0_Graphics.dd file*
  - Type: Unallocated Space Image File
- *Select only the following ingest file:*
  - Photorec Carving
- *Find out what deleted files are found*

Follow the direction in class for detailed instructions