Basic File Carving With FTimes

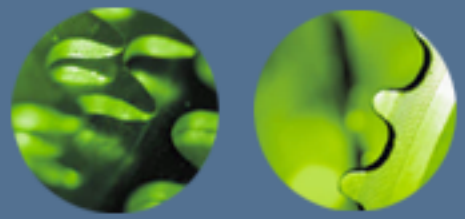
CEIC 2007

May 8, 2007

KoreLogic, Inc:

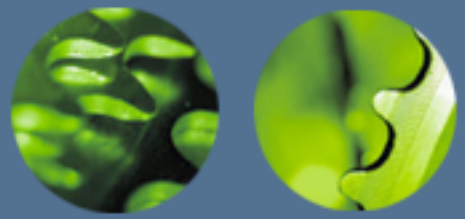
Andy Bair
pab-ceic@korelogic.com

Jay Smith
jsmith-ceic@korelogic.com



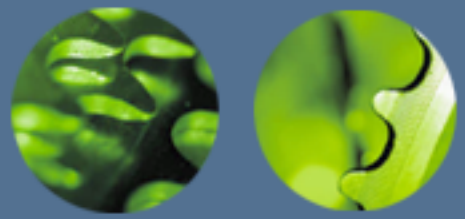
Overall Agenda

- Basic Section
  - Introduction - File Carving Overview
  - Background – Terminology & Methodology
  - FTimes – Introduction to FTimes Map Mode
  - Lab #1 – Extract JPEG Image From A Tar Ball
- Advanced Section
  - XMagic – Fundamentals
  - XMagic & Carving
  - Lab #2 – Extract Word Doc From A Packet Capture
  - Conclusions



Basic Section Agenda

- Introduction - File Carving Overview
- Background – Terminology & Methodology
- FTimes – Introduction to FTimes Map Mode
- Lab #1 – Extract JPEG Image From A Tar Ball



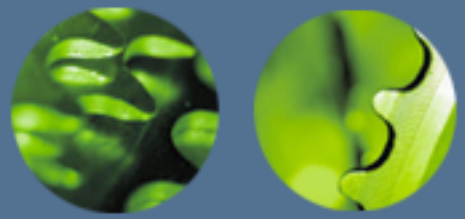
File Carving – Overview

- Identify and recover data/files based on analysis of file formats
- Carving is a powerful technique because it can
  - Identify and recover data and/or files of interest from raw, deleted, or damaged file system, memory, or swap space data
  - Assist in recovering files and data that may not be accounted for by the operating system and file system
  - Assist in simple data recovery
- Not just limited to data/files, can use techniques on memory, etc

Reality Check! #1

Your subject was hacked and you believe some evidence resides within the machine's memory. You have dumped the running memory and need to look for evidence.

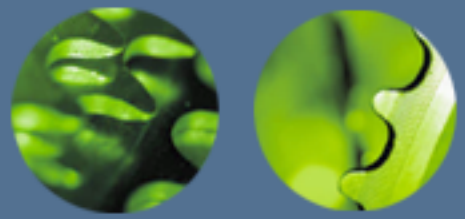
Using data carving, you can extract the processes and other relevant data from the memory dump.



File Carving – Details

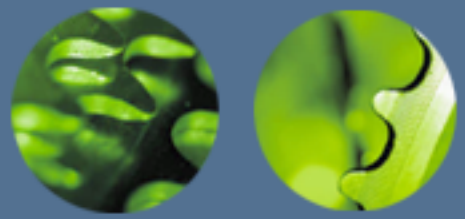
- Many file types have well-known values or magic(5)* numbers in the 4-8 KB of the file header
- Typical file carvers
  - Identify specific types of file headers and/or footers
  - Carve out blocks between these two boundaries
  - Stop carving after a user-specified or set limit has been reached

(*)Denotes section 5 of the magic man page.



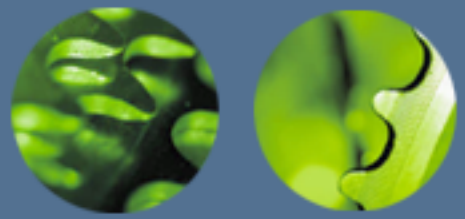
File Carving - Challenges

- Number of forensic cases growing exponentially today
- Investigations can be lengthy
  - Machines tied up for days during investigations
  - Forensic targets with GB or TB of storage
  - Still need rapid turnaround, especially in time-sensitive cases involving potential loss of life or property -- think terrorists
- Not all file types have standard footer signature, determining end can be difficult
- Existing file carving tools typically produce many false positives and can miss key evidence
  - Need file carving algorithms that identify more files and reduce the number of false positives
- Tools don't handle fragmentation very well
  - A 'validator' can assist in testing the carved data



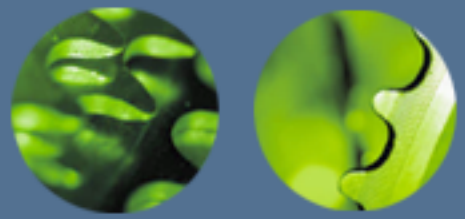
Basic Section Agenda

- Introduction - File Carving Overview
- Background – Terminology & Methodology
- FTimes – Introduction to FTimes Map Mode
- Lab #1 – Extract JPEG Image From A Tar Ball



Background - Terminology

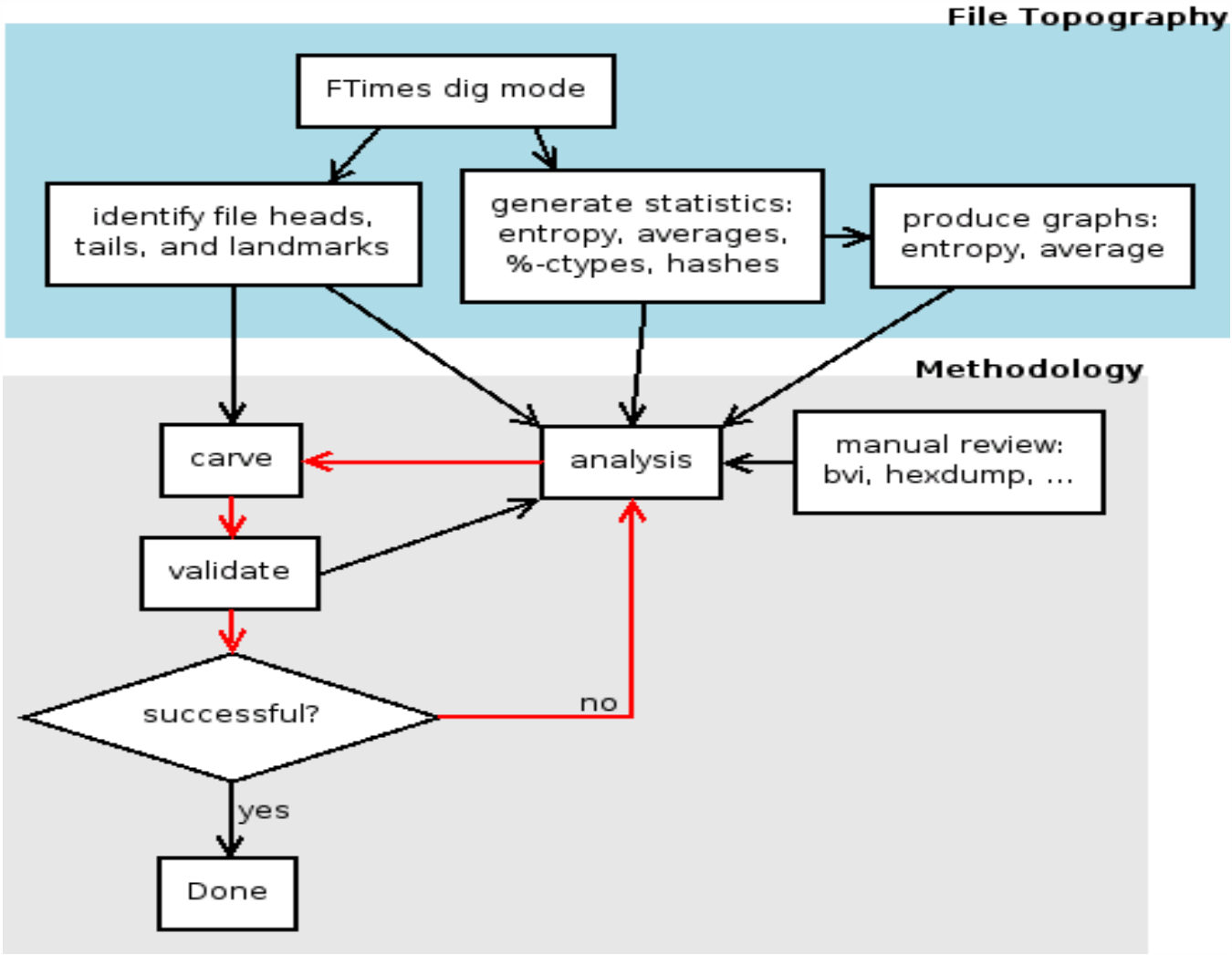
- **SOF/EOF** - start/end of file
- **SOI/EOI** - start/end of image
- **FAT** - file allocation table (also referred to as SAT blocks)
- **OLE** - Object Linking and Embedding, Microsoft's compound documents
- **Magic** - constant used to identify file or data type (also know as file typing)
- **XMagic -** Extended Magic
- **Zero-storage -** Zero additional space needed to carve out files. Information about sector information is used to keep track of location of data to be carved

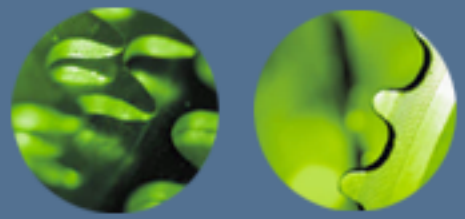


Background - Terminology

- **Entropy (1-byte)**
  - Measure of randomness
  - Range = 0-8; 8 = most random; 0 = least random
  - Dramatic entropy changes can indicate file boundary
- **Sliding Entropy**
  - Calculating entropy for each sequential file data block
- **Sliding Average**
  - Calculating average for each sequential file data block
- **Sliding Hash (MD5 ,SHA1 and SHA256)**
  - Calc message digests for each sequential file data block
  - Block size is variable
  - Bashed against 1+ subject images
  - Can use to locate duplicate blocks

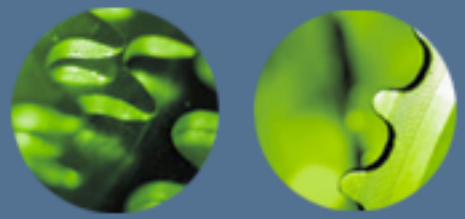
Background - Methodology





Basic Section Agenda

- Introduction - File Carving Overview
- Background – Terminology & Methodology
- FTimes – Introduction to FTimes
- Lab #1 – Extract JPEG Image From A Tar Ball

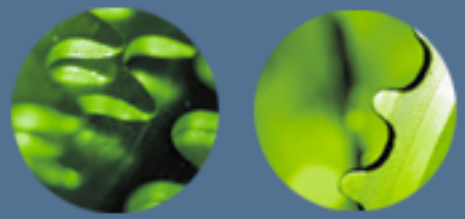


FTimes - Overview

http://ftimes.sourceforge.net/FTimes/index.shtml

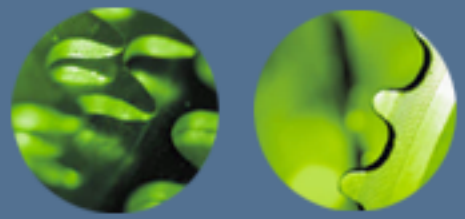
- System baselining and evidence collection tool
- Gather/develop topographical information & attributes about directories and files in a manner conducive to intrusion and forensic analysis
- Lightweight: small footprint, command line interface
- Used dig (“search”) mode in conjunction with XMagic to develop topography
- 2 major modes of operation: Map and Dig mode

File **T**opography and **I**ntegrity **M**onitoring on an **E**nterprise **S**cale



FTimes usage

Usage: ftimes --cfgtest file mode [-s]
ftimes --compare mask baseline snapshot [-l level]
ftimes --decoder snapshot [-l level]
ftimes --digauto file [-l level] [list]
ftimes --digfull file [-l level] [list]
ftimes --diglean file [-l level] [list]
ftimes --getmode file [-l level]
ftimes --mapauto mask [-l level] [list]
ftimes --mapfull file [-l level] [list]
ftimes --maplean file [-l level] [list]
ftimes --version



File Carving – Tool Comparison

- Foremost : Tool used to extract files based off of know SOF and EOF.
- Scalpel : Complete rewrite of foremost and recommended by the authors of foremost.
- ftimes-crv2raw.pl : Can utilize known SOF and EOF. XMagic can be used to extend the carving functionality to carve out files that do not have an EOF.
- CarvFS/LibCarvPath : Provides virtual file system to forensics tools of sectors to be carved.

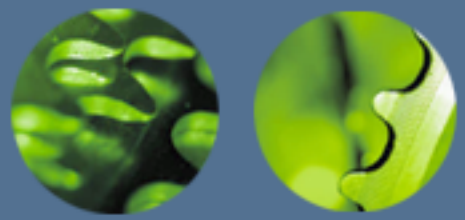
Reality Check! #2

You have been called in to perform an onsite analysis of a running system that can not be taken offline. You want to preserve all the MAC/MACH times as well as get information on the types of files that reside on the subject.

….hmmm….

FTimes in map mode with a FieldMask of "all" will collect all file attributes as well as collect file type information. When run under Windows FTimes will also collect the CH time.

* Windows NTFS files have an additional CH time that corresponds to the change time of a file.

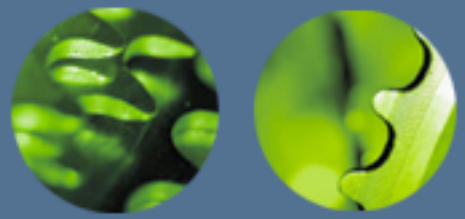


FTimes example output

C:\ftimes>ftimes --mapauto none -l 6 c:\temp
name
“C:\temp\employee.xls"
“C:\temp“

C:\ftimes>ftimes --mapauto none+size+md5 -l 6 c:\temp
name|size|md5
“C:\temp\employee.xls"|13824|bae16190d0527edaed2b42fdeb90bd1f
“C:\temp"|0|DIRECTORY

C:\ftimes>ftimes --mapauto none+hashes -l 6 c:\temp
name|md5|sha1|sha256
“C:\temp\employee.xls"|bae1…bd1f|0dd0…acef|eb4d…d1fc
“C:\temp"|DIRECTORY|DIRECTORY|DIRECTORY

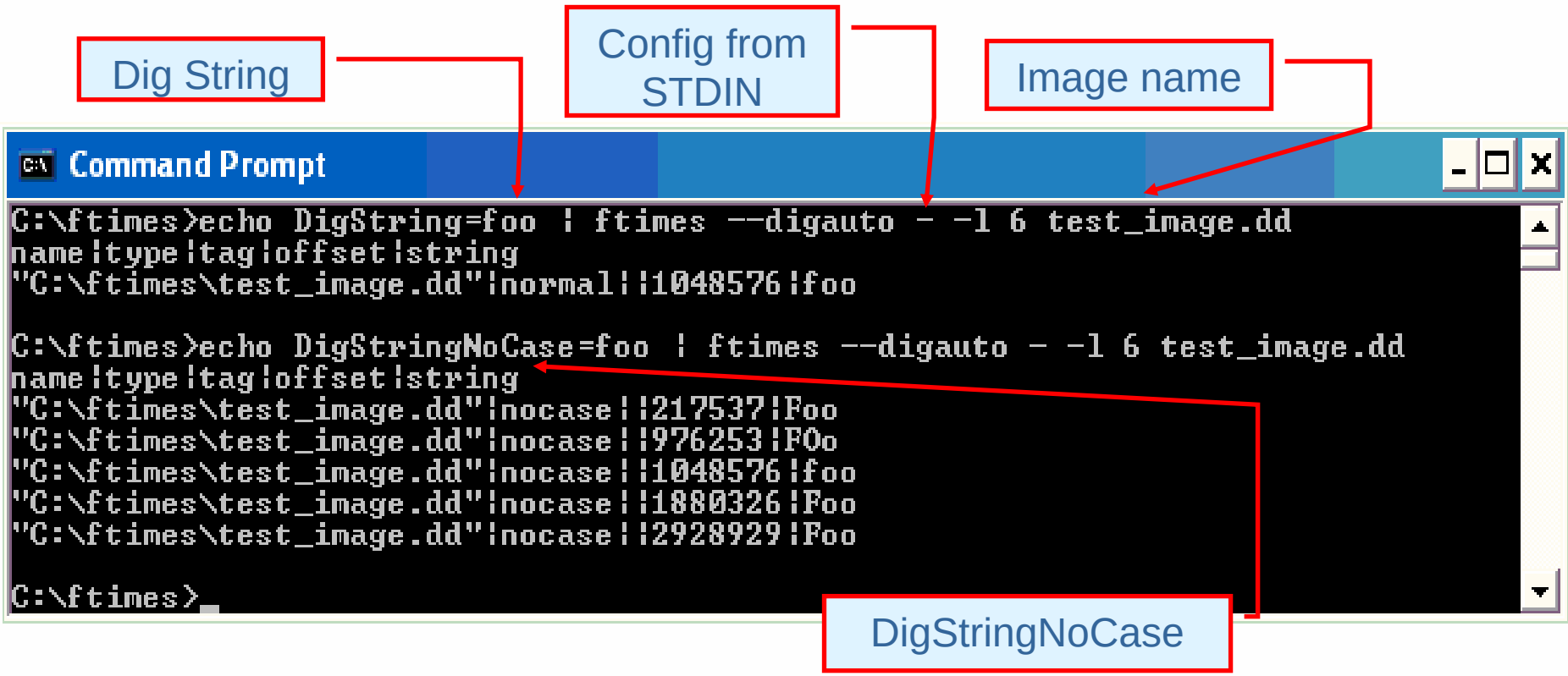


FTimes – Dig Mode

- Search through directories and files recursively
  - User-specified regular expressions
  - Sequence(s) of bytes
- 3 tiers of searching
  - Basic – DigStringNormal, DigStringNoCase
    - DigStringNormal=foo
  - Advanced – DigStringRegExp
    - DigStringRegExp=\x66\x6f\x6f
  - Expert – DigStringXMagic
- Can be used to search **all** old image cases for information of interest
- Case insensitive
- Regular expressions most powerful but not fastest
- Specify byte by byte patterns

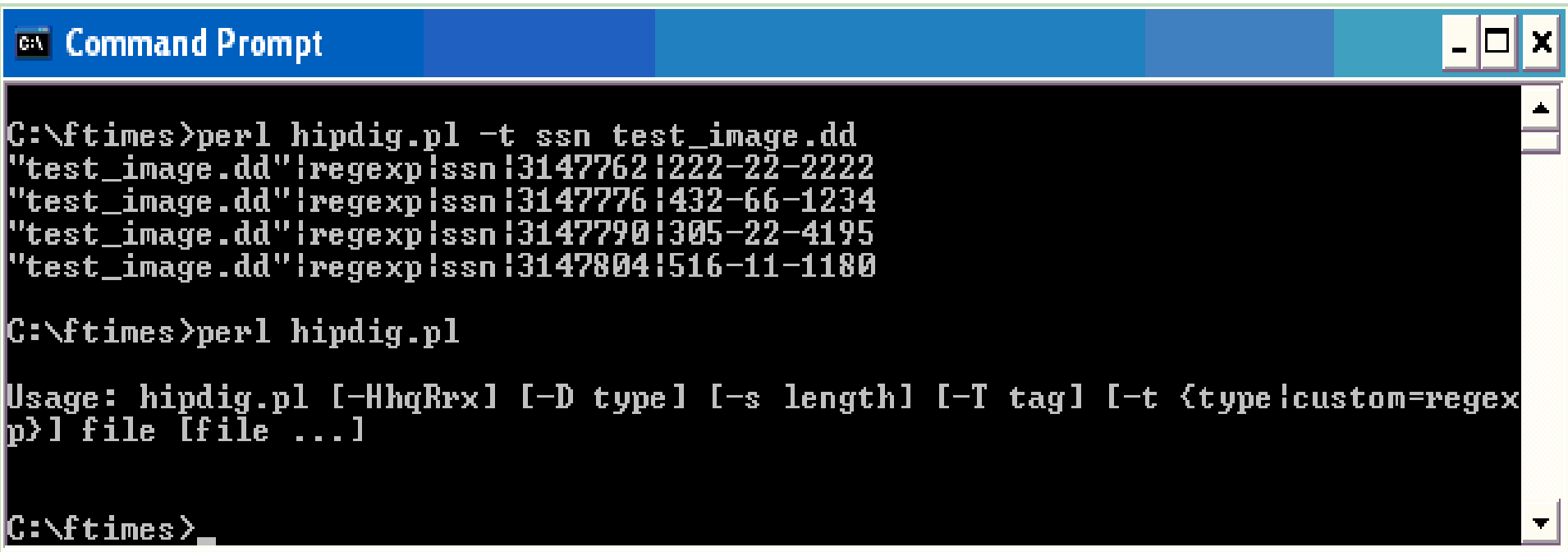
Command Line Sample

- Using different DigString options you can achieve different results



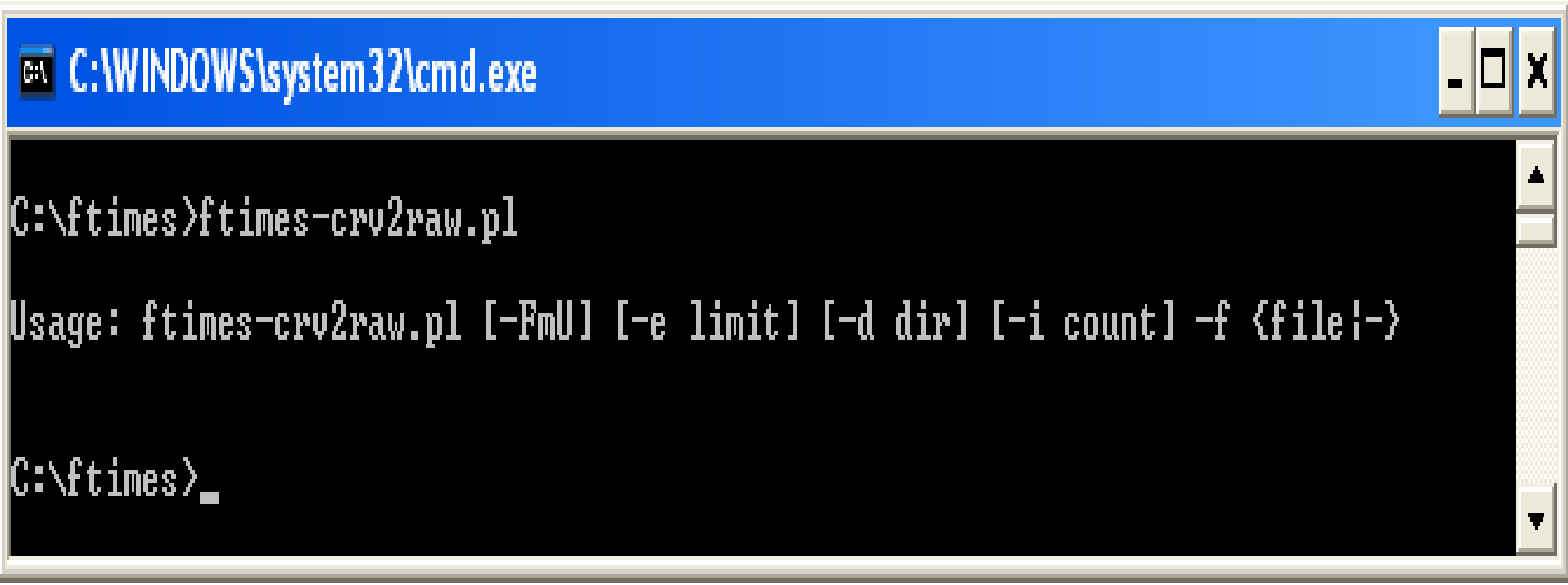
Hipdig.pl

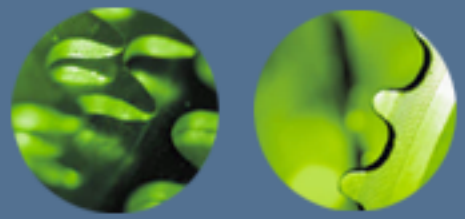
- hipdig.pl – Utility part of the FTimes project that can be used as a standalone application to search for terms within files or a block of data.



ftimes-crv2raw.pl

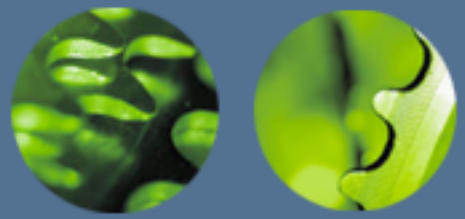
- ftimes-crv2raw.pl – Utility that is part of the FTimes project that is used to carve out data/files using input from hipdig, ftimes, custom tools, or manually.





Why FTimes?

- XMagic can be used to locate points of interest within a block of data
- Currently available tools do not assist in identifying fragments and issues associated with straight carving
- Command line driven and can be incorporated into scripts and other tools
- Won the 2006 DFRWS File Carving Challenge



File Carving – Basic Example

- JPEG files start (SOI) with **`0xffd8`** and end (EOI) with **`0xffd9`**
- To recover a JPEG file:
  - Find the locations of its header (SOI) and footer (EOI)
  - And carve everything between those two endpoints (inclusive)

Hexdump of sample.jpg

```
ff d8 ff e0 00 10 4a 46  49 46 00 01 01 01 00 50  |......JFIF.....P|
... Data ...
28 a2 80 3f ff d9                                 |(..?..|
```

FTimes – EOFs and SOFs

JPEG with 2 thumbnails:
Example of regexs and byte patterns

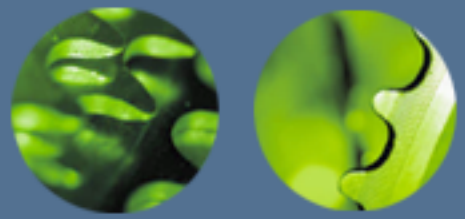
combined.cfg

```
DigStringRegExp=(?s)(\xff\xd8....JFIF)   sof.jpeg
DigStringNormal=%ff%d9                   eof.jpeg
```

ftimes –diglean combined.cfg challenge.raw

combined.dig

```
"challenge.raw"|regexp|sof.jpeg|1980416|%ff%d8%ff%e0%00%10JFIF
"challenge.raw"|regexp|sof.jpeg|1980748|%ff%d8%ff%e0%00%10JFIF
"challenge.raw"|normal|eof.jpeg|1986297|%ff%d9
"challenge.raw"|regexp|sof.jpeg|1995443|%ff%d8%ff%e0%00%10JFIF
"challenge.raw"|normal|eof.jpeg|2000992|%ff%d9
"challenge.raw"|normal|eof.jpeg|2267600|%ff%d9
```



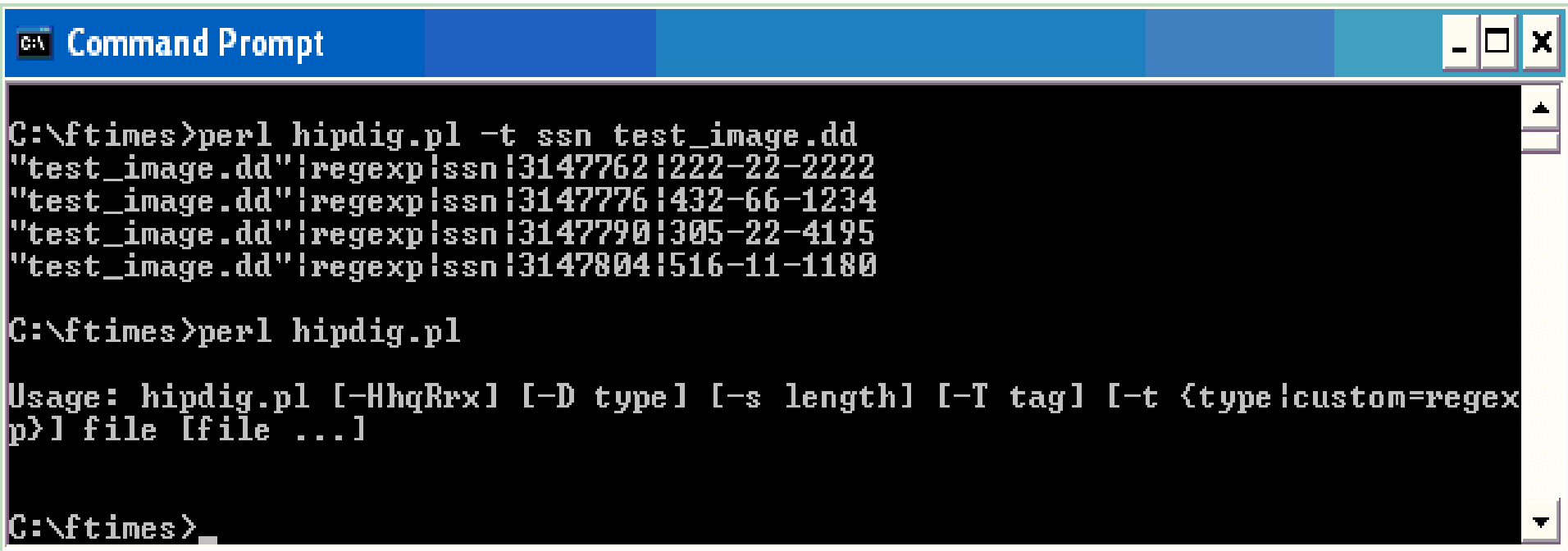
Reality Check

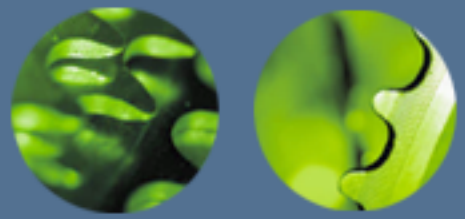
- Example Case: Subject is stealing confidential information (e.g., SSNs or personal information) and storing it in an XL spreadsheet that is hidden from the Operating System.

- How would you go about finding and extracting those spreadsheets?

Look for the data

- We first run hipdig.pl looking for the SSNs on the subject image.





Reality Check

Run FTimes with the XMagic we want to look for:

```
ftimes --diglean ftimes-diglean-ole.cfg -l 6 test_image.dd > test_image.dig
```

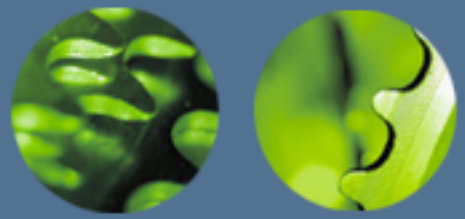
This will produce results like this:

```
name|type|tag|offset|string
"C:\ftimes\flow\test_image.dd"|xmagic|sof.ole|3145759|18
"C:\ftimes\flow\test_image.dd"|xmagic|fat.ole|3158559|00000001:aaaf...
```

```
ole-dig2crv.pl -f test_image.dig 2>&1 |egrep "valid FAT block #”
```

This will produce results like this:

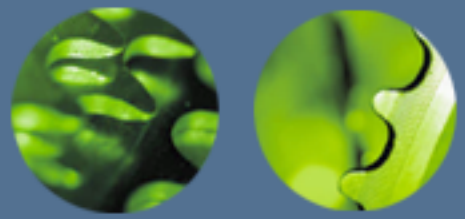
```
file at 3145759, offset 3158559, block 6169.060546875 -- vld FAT blck #1, 0x18
```



Reality Check

- ole-dig2crv.pl -f test_image.dig

```
C:\ftimes\flow>ole-dig2crv.pl -f test_image.dig

file at 3145759, offset 3145759, block 6144.060546875 – hdr FAT block pointers (in hex): 18

file at 3145759, offset 3158559, block 6169.060546875 -- valid FAT block #1, 0x18

file at 3145759 -- PredictedFileSize = 13824 = (((1-1) * 128) + 26 + 1) * 512

file at 3145759 -- ExactEndOffset = 3159582 = (3145759 + (((1-1) * 128) + 26 + 1) * 512) – 1

file at 3145759 -- CrvData="C:\ftimes\flow\test_image.dd"|ole|3145759|1|3145759-3159582
"C:\ftimes\flow\test_image.dd"|ole|3145759|1|3145759-3159582

file at 3145759 -- CrvDataTemplate="C:\ftimes\flow\test_image.dd"|ole|3145759|1|3145759-3159583
```



Reality Check

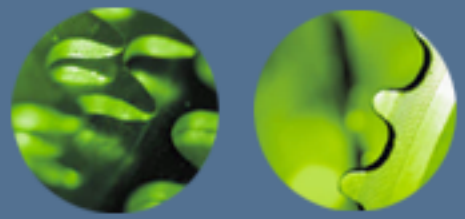
```
C:\ftimes\flow>ole-dig2crv.pl -f test_image.dig 2>&1 | egrep
   CrvData= | cut -d= -f2- > test_image.crv
```

This will produce results like this:

```
"C:\ftimes\flow\test_image.dd"|ole|3145759|1|3145759-3159582
```

C:\ftimes\flow>ftimes-crv2raw.pl -m -f test_image.crv

```
name|size|md5|sha1
"carve_tree/ftimes\flow\test_image.dd.3145759.ole"|1382
   4|5771…042fa|5fa5…6f85
```



Reality Check

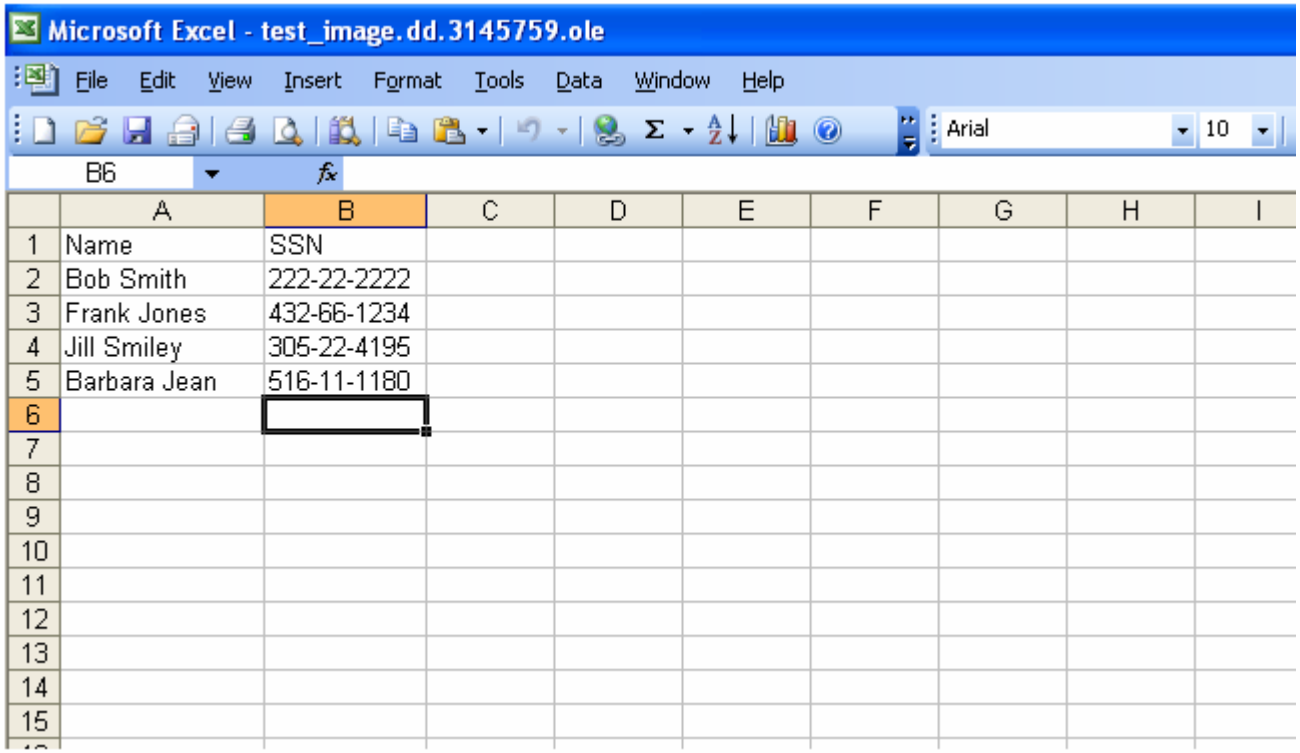
```
C:\ftimes\flow>ftimes-crv2raw.pl -m -f test_image.crv
name|size|md5|sha1
"carve_tree/ftimes\flow\test_image.dd.3145759.ole"|13
   824|5771…42fa|5fa5…6f85
```

- Lets check to make sure the carved files hashes match the original:

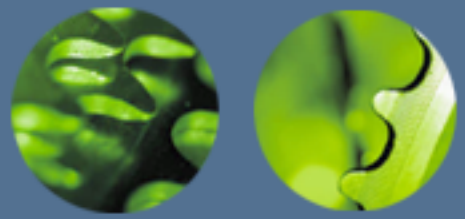
```
C:\ftimes\flow>ftimes --mapauto none+sha1+md5 -l 6
   c:\ftimes\employee.xls

name|md5|sha1
"c:\ftimes\employee.xls"|5771…42fa|5fa5…6f85
```

Reality Check

The final extracted XL spreadsheet

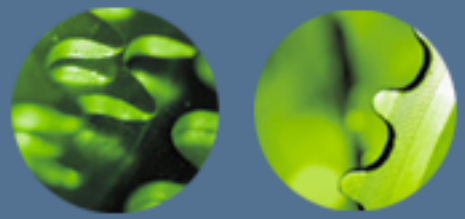


	A	B	C	D	E	F	G	H	I
1	Name	SSN							
2	Bob Smith	222-22-2222							
3	Frank Jones	432-66-1234							
4	Jill Smiley	305-22-4195							
5	Barbara Jean	516-11-1180							
6									
7									
8									
9									
10									
11									
12									
13									
14									
15									



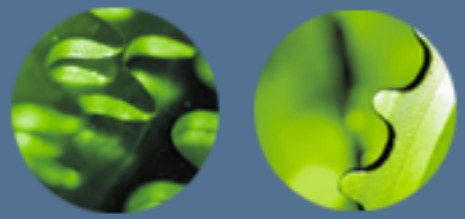
Basic Section Agenda

- Introduction - File Carving Overview
- Background – Terminology & Methodology
- FTimes – Introduction to FTimes Map Mode
- Lab #1 – Extract JPEG Image From A Tar Ball



Basic Lab #1

- Extract JPEG Image From A Tar Ball
  - Not about tools
  - It’s about techniques
  - Using your brain



Links and Info

http://www.korelogic.com

http://ftimes.sourceforge.net

http://www.korelogic.com/Resources/Projects/dfrws_challenge_2006/

Email: pab_ceic@korelogic.com
jsmith_ceic@korelogic.com