HackerFrogs Afterschool
Digital Forensics: Digital Pictures

Class:
Digital Forensics

Workshop Number:
AS-FOR-01

Document Version:
1.2

Special Requirements:
Registered account at picoctf.org

Welcome to HackerFrogs Afterschool!

Hey there HackerFrogs!

This is the first intro to digital forensics workshop.

Let us begin!

Digital Image Files

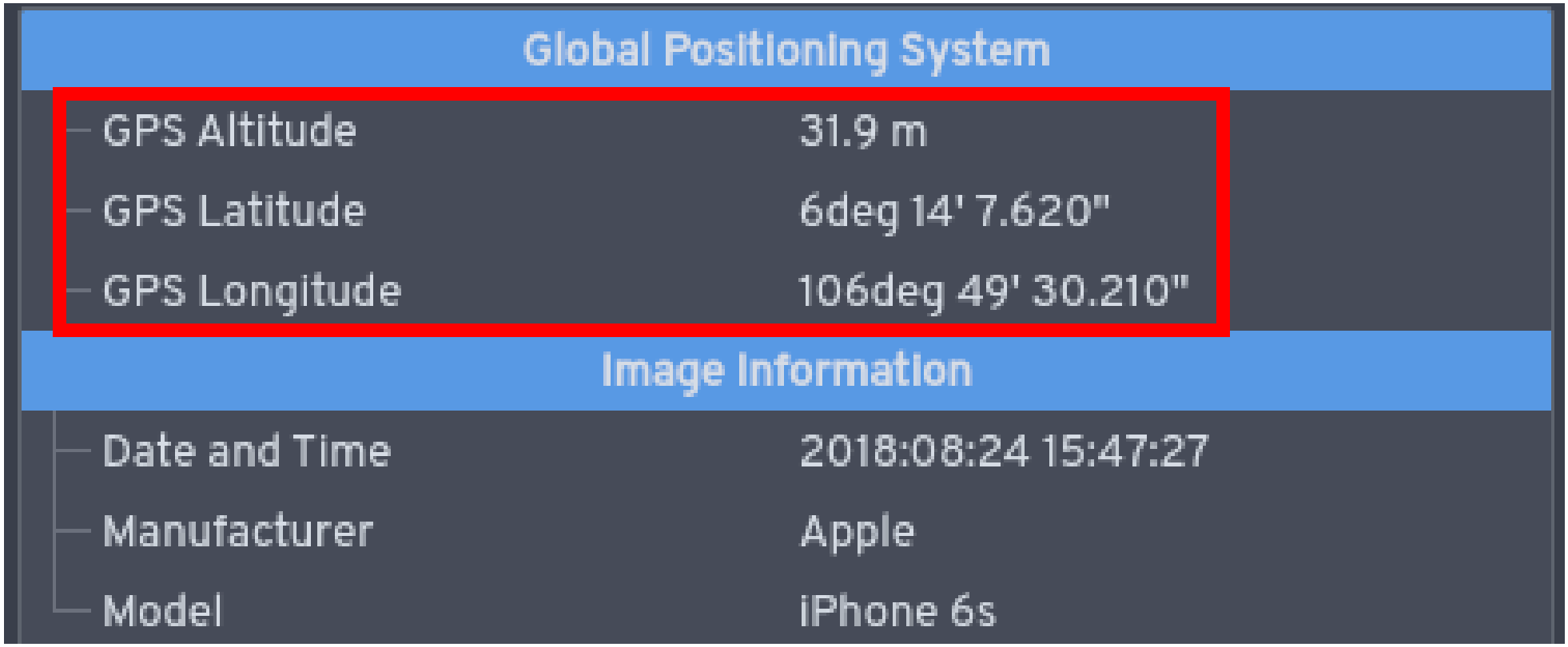
The topic of this workshop is image file forensics, i.e., digital pictures. There are a few different ways information can be hidden in these types of files.

EXIF and Image File Metadata

Metadata is data that provides information about other data.
In the context of digital image files, each file has a plethora of metadata information which is generated when the picture is taken, including:

Global Positioning System	
GPS Altitude	31.9 m
GPS Latitude	6deg 14' 7.620"
GPS Longitude	106deg 49' 30.210"
Image Information	
Date and Time	2018:08:24 15:47:27
Manufacturer	Apple
Model	iPhone 6s
Photograph Information	
Aperture	F2.2
Exposure Bias	0 EV
Exposure Mode	Auto
Exposure Program	Auto
Exposure Time	1/874 s
Flash	No, auto
FNumber	F2.2
Focal Length	4.2 mm
ISO Speed Ratings	25
Metering Mode	Multi-segment
Shutter speed	1/874 s
White Balance	Auto

EXIF and Image File Metadata



Where the image was taken (GPS Coordinates)

EXIF and Image File Metadata

When the image was taken (date and time),

EXIF and Image File Metadata

And the type of device used to create the image (manufacturer and model).

EXIF and Image File Metadata

For photos, this metadata is called Exchangeable Image File Format (EXIF). The EXIF standard is associated with digital cameras, smartphones and scanners.

Global Positioning System	
GPS Altitude	31.9 m
GPS Latitude	6deg 14' 7.620"
GPS Longitude	106deg 49' 30.210"
Image Information	
Date and Time	2018:08:24 15:47:27
Manufacturer	Apple
Model	iPhone 6s
Photograph Information	
Aperture	F2.2
Exposure Bias	0 EV
Exposure Mode	Auto
Exposure Program	Auto
Exposure Time	1/874 s
Flash	No, auto
FNumber	F2.2
Focal Length	4.2 mm
ISO Speed Ratings	25
Metering Mode	Multi-segment
Shutter speed	1/874 s
White Balance	Auto

EXIF and Image File Metadata

While EXIF data is useful for archival purposes, this metadata poses a serious security risk when image files are shared, especially when uploaded to the internet.

EXIF and Image File Metadata

The most obvious example of risk is the abuse of EXIF data which could allow third-parties to discover a person's residence or neighborhood through through GPS metadata.

Part 1 – Image Metadata With Pico Information

Let's learn more about image metadata, and how to access it by solving the **Information** challenge in Pico CTF:

https://play.picoctf.org/practice/challenge/186

Base64 Encoding

```
[root@localhost ~]# echo -n 'this will be encoded into base64' | base64
dGhpcyB3aWxsIGJlIGVuY29kZWQgaW50byBiYXNlNjQ=
```

Base64 is a widely-used encoding method which is commonly used to transform file and picture data for transmission over computer networks.

Base64 Encoding

```
[root@localhost ~]# echo -n 'this will be encoded into base64' | base64
dGhpcyB3aWxsIGJlIGVuY29kZWQgaW50byBiYXNlNjQ=
```

It transforms text and / or data bytes into strings like the one illustrated above. Base64 also features in a lot of CTF challenges, and is used to disguise information.

Base64 Encoding

```
[root@localhost ~]# echo -n 'this will be encoded into base64' | base64
dGhpcyB3aWxsIGJlIGVuY29kZWQgaW50byBiYXNlNjQ=
```

In the example above, the text 'this will be encoded into base64' is transformed into the string underlined in red.

Base64 Encoding

Base64 encoded strings can be identified as a string of characters composed of any of the characters located in the encoding table seen here.

Base64 Encoding Table

Value	Char	Value	Char	Value	Char	Value	Char
0	A	16	Q	32	g	48	w
1	B	17	R	33	h	49	x
2	C	18	S	34	i	50	y
3	D	19	T	35	j	51	z
4	E	20	U	36	k	52	0
5	F	21	V	37	l	53	1
6	G	22	W	38	m	54	2
7	H	23	X	39	n	55	3
8	I	24	Y	40	o	56	4
9	J	25	Z	41	p	57	5
10	K	26	a	42	q	58	6
11	L	27	b	43	r	59	7
12	M	28	c	44	s	60	8
13	N	29	d	45	t	61	9
14	O	30	e	46	u	62	+
15	P	31	f	47	v	63	/

Base64 Encoding

```
[root@localhost ~]# echo -n 'this will be encoded into base64' | base64
dGhpcyB3aWxsIGJlIGVuY29kZWQgaW50byBiYXNlNjQ=
```

The encoded strings can also contain the equals (=) sign if the encoded string character length is not evenly divisible by 4.

Base64 Encoding

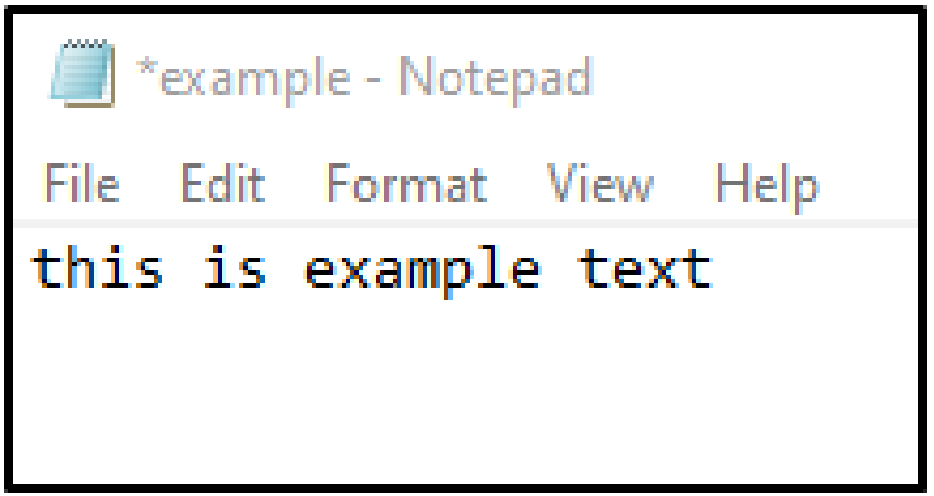
```
[root@localhost ~]# echo -n 'this will be encoded into base64' | base64
dGhpcyB3aWxsIGJlIGVuY29kZWQgaW50byBiYXNlNjQ=
```

In the underlined string, the character length is 43 without the equals character, so one is added so that the length equals 44, a length that divides evenly by 4.

The Contents of Image Files

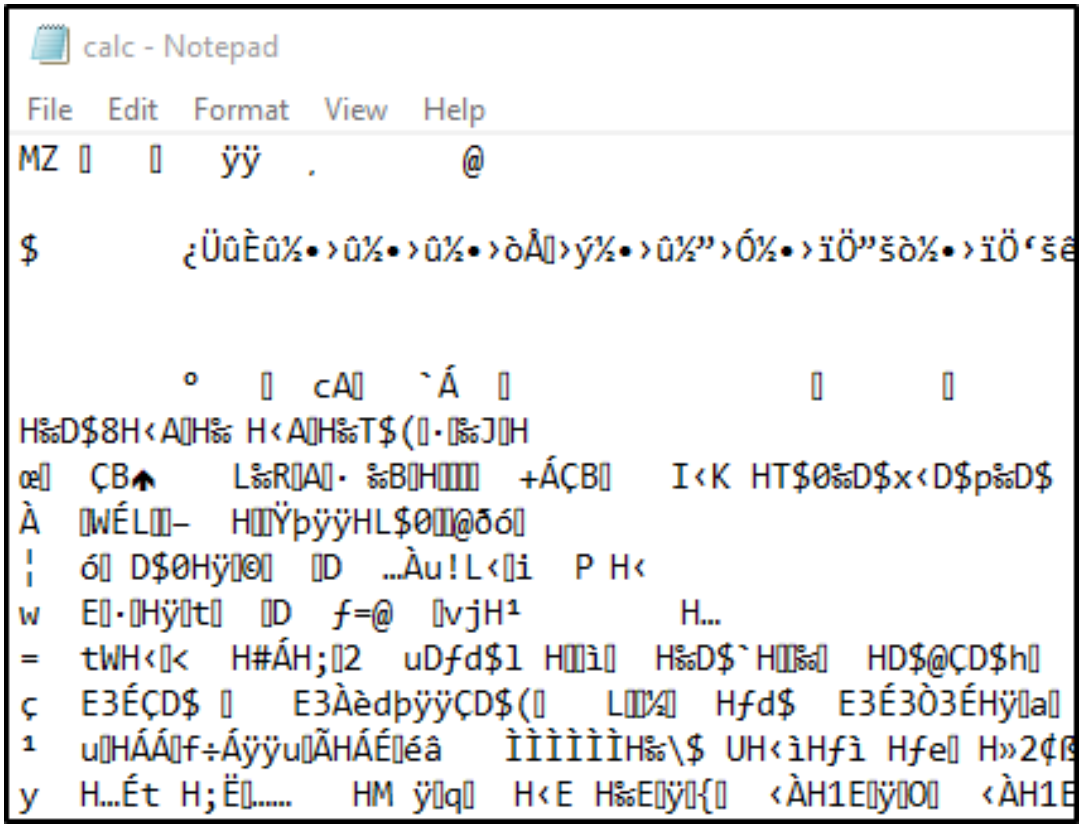
At an abstract level, files contain either human-readable **text** (letters, numbers, and symbols), and / or non-human-readable **data** (program code / instructions, etc), also called **binary data**.

The Contents of Image Files



This is an example of text file contents. It is intended to be read by users (people).

The Contents of Image Files



This is an example of data file contents. It is not meant to be read by people, rather read and executed by machines (computers / software).

The Contents of Image Files

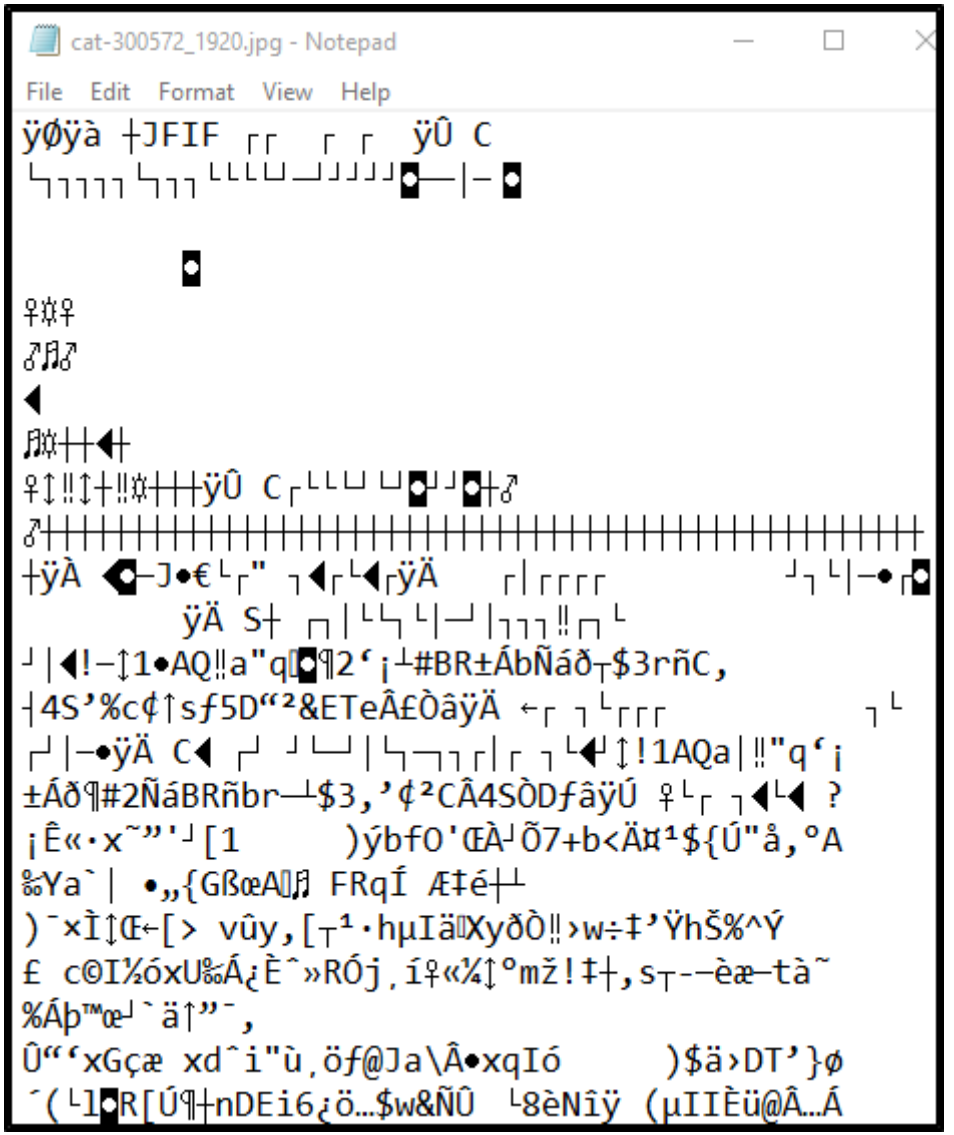
In digital picture files, the majority of their contents are non-human-readable data that is intended to be ready by an image viewing program or web browser.

The Contents of Image Files

When opened with one of these programs, the file's data is read and the image appears on the screen. However, as we learned earlier, there are often text strings present inside image files.

The Contents of Image Files

However, if image files are opened with a program that isn't designed to read image files, then the contents of the file will appear to be gibberish.



Part 2 – Extracting Text With Pico Glory of the Garden

Let's learn more about text versus data, and how to extract text by solving the **Glory of the Garden** challenge in Pico CTF:

https://play.picoctf.org/practice/challenge/44

The Strings Command

The Strings command is used to return human-readable text (strings) from files. It's used to return text from files that contain both data bytes and text (e.g., picture files).

The Strings Command

The Strings command is used to return human-readable text (strings) from files. It's used to return text from files that contain both data bytes and text (e.g., picture files).

The Strings Command

```
theshyhat-picoctf@webshell:~/tmp$ strings garden.jpg | head
JFIF
XICC_PROFILE
HLino
mntrRGB XYZ
acspMSFT
```

If run by itself, it returns all strings, which often isn't useful, because a lot of random bytes in files can also be interpreted as ASCII characters.

The Strings Command

```
theshyhat-picoctf@webshell:~/tmp$ strings -16 garden.jpg
Copyright (c) 1998 Hewlett-Packard Company
sRGB IEC61966-2.1
sRGB IEC61966-2.1
IEC http://www.iec.ch
IEC http://www.iec.ch
.IEC 61966-2.1 Default RGB colour space - sRGB
```

But if we run the command with a number argument, we can specify the minimum-length string to display in the output.

Embedded Files

Embedded files are files that exist inside of other files. Embedded files are often associated with archive file types, such as Zip and Rar.

Embedded Files

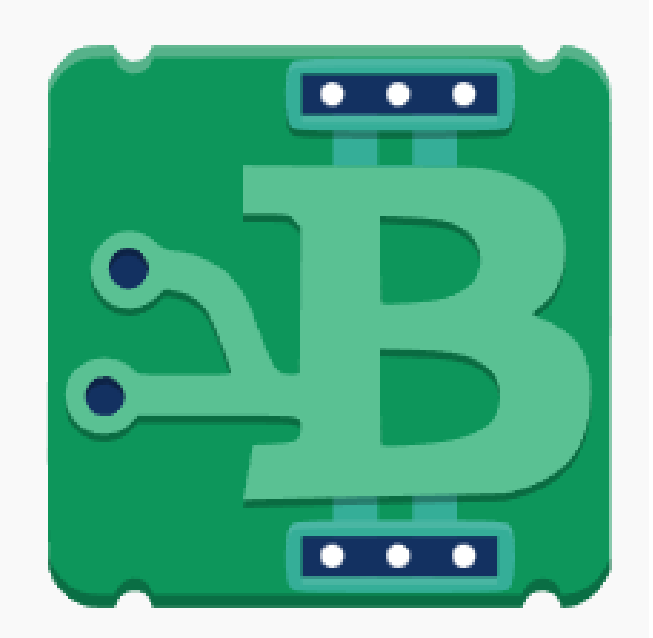
However, embedded files can be found in many common file types, such as Word, Excel, PDF, MP3 audio, and various picture files.

Part 3 – Embedded Files With Pico Matroyshka Doll

Let's explore embedded files with another Pico CTF challenge:

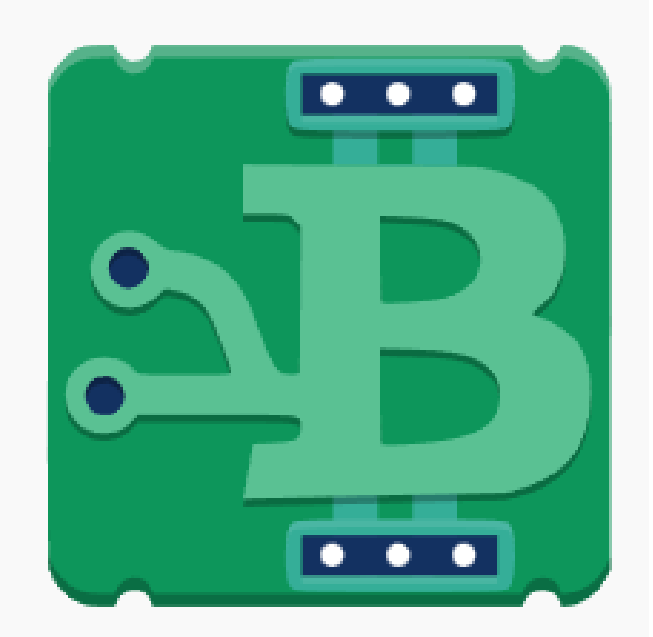
https://play.picoctf.org/practice/challenge/129

The Binwalk Program



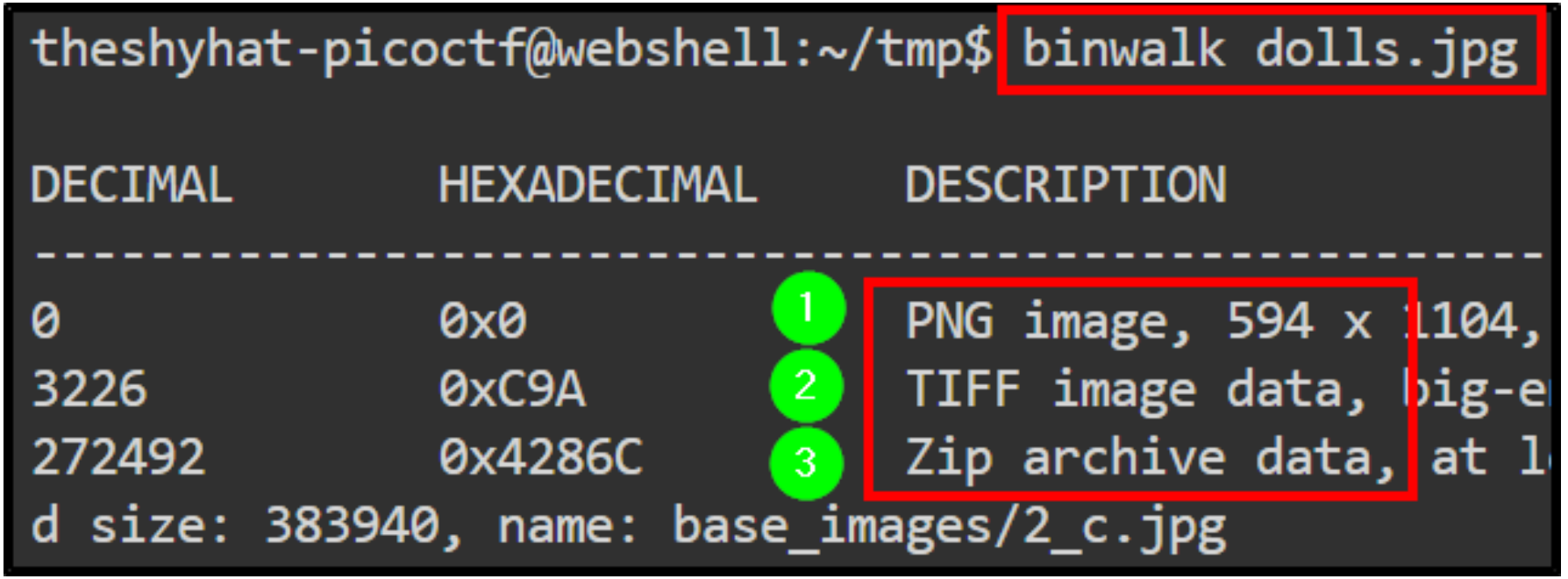
Binwalk is a CLI program used for analyzing and extracting data from files.

The Binwalk Program



It is used by cybersecurity students and security researchers to analyze and reverse-engineer binary executable files as well as other file types.

The Binwalk Program



When run with a file, it returns the contents of that file, identifying embedded files.

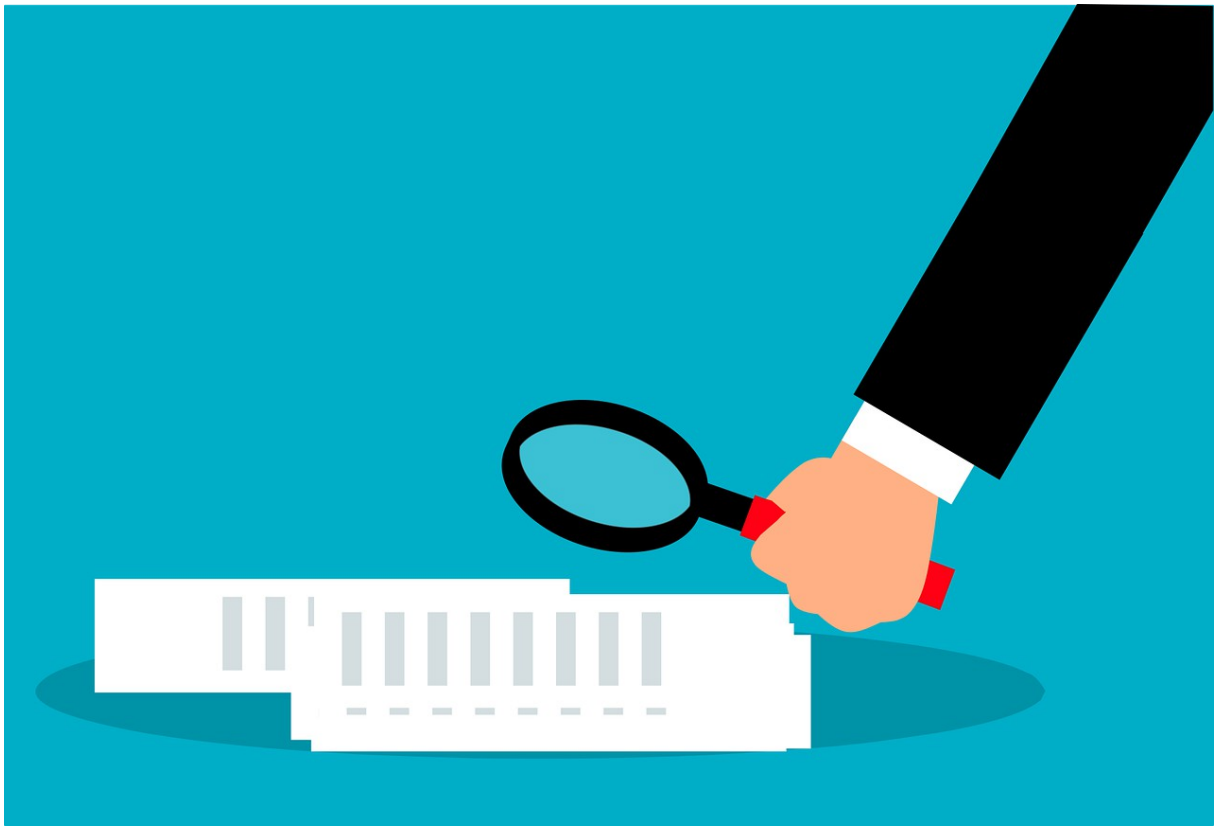
The Binwalk Program

```
binwalk -e dolls.jpg
```

```
_dolls.jpg.extracted  dolls.jpg
```

If run with the -e switch, binwalk extracts all files from the target file to a directory with the .extracted extension.

Summary



Let's review the digital forensics concepts we learned in this workshop:

EXIF and Image File Metadata

Metadata is data that provides information about other data.
In the context of digital image files, each file has a plethora of metadata information which is generated when the picture is taken.

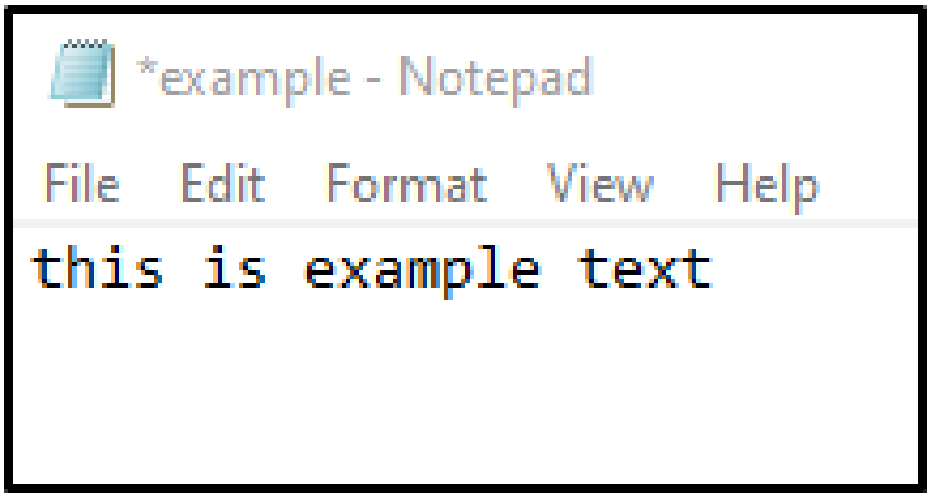
Global Positioning System	
GPS Altitude	31.9 m
GPS Latitude	6deg 14' 7.620"
GPS Longitude	106deg 49' 30.210"
Image Information	
Date and Time	2018:08:24 15:47:27
Manufacturer	Apple
Model	iPhone 6s
Photograph Information	
Aperture	F2.2
Exposure Bias	0 EV
Exposure Mode	Auto
Exposure Program	Auto
Exposure Time	1/874 s
Flash	No, auto
FNumber	F2.2
Focal Length	4.2 mm
ISO Speed Ratings	25
Metering Mode	Multi-segment
Shutter speed	1/874 s
White Balance	Auto

EXIF and Image File Metadata

Photo metadata, called EXIF, is easily accessible, and can reveal information about the device used to take the photo, or even GPS location data about where the photo was taken.

Global Positioning System	
GPS Altitude	31.9 m
GPS Latitude	6deg 14' 7.620"
GPS Longitude	106deg 49' 30.210"
Image Information	
Date and Time	2018:08:24 15:47:27
Manufacturer	Apple
Model	iPhone 6s
Photograph Information	
Aperture	F2.2
Exposure Bias	0 EV
Exposure Mode	Auto
Exposure Program	Auto
Exposure Time	1/874 s
Flash	No, auto
FNumber	F2.2
Focal Length	4.2 mm
ISO Speed Ratings	25
Metering Mode	Multi-segment
Shutter speed	1/874 s
White Balance	Auto

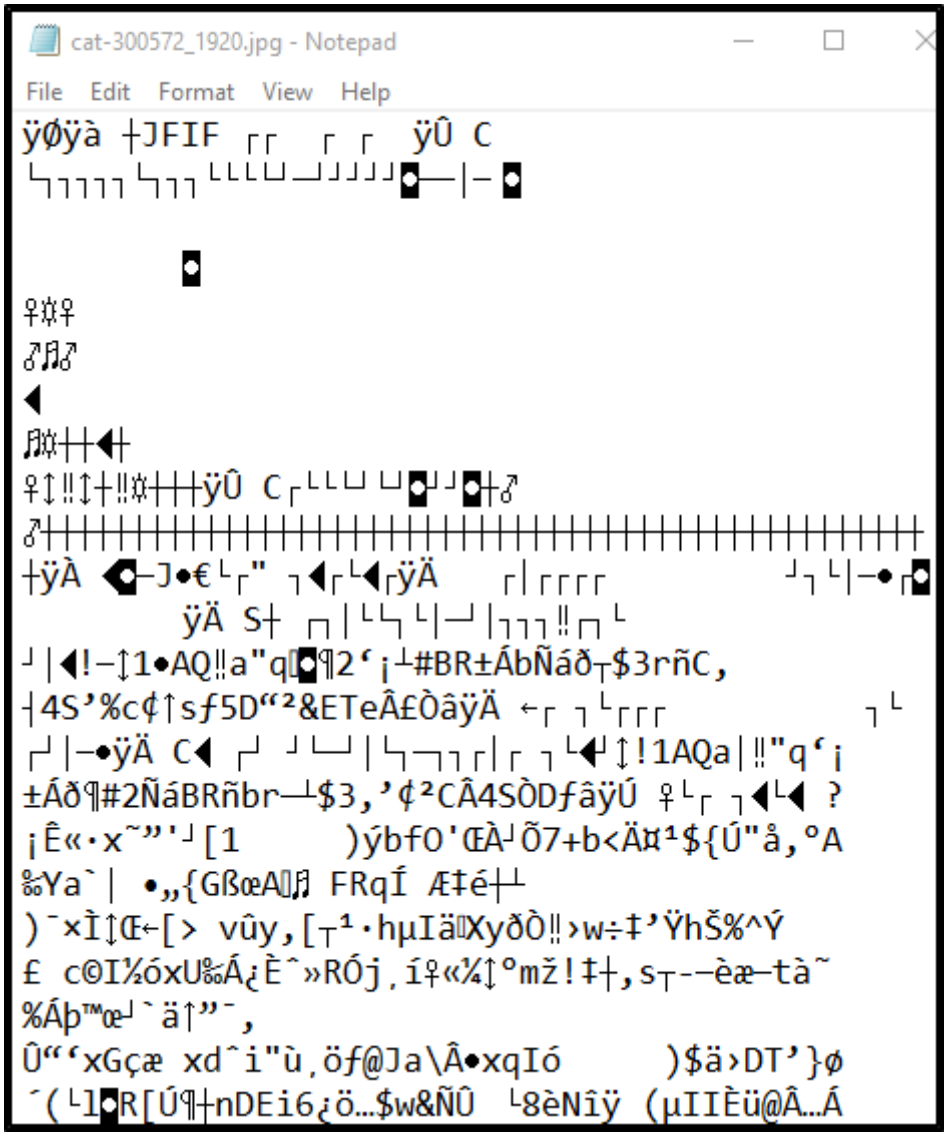
The Contents of Image Files



We learned about the contents of picture files, and the difference between text and data. Text is computer content that is meant to be read by humans.

The Contents of Image Files

And data content is meant to be read by computers / software, and is illegible if not accessed through the proper program / app.



Embedded Files

Lastly, we learned about embedded files, and how embedded files can be extracted and examined using specialized software.

What's Next?

In the next digital forensics workshop, we'll introduce network traffic analysis with the TryHackMe platform.

Extra Credit

Looking for more study material on this workshop's topics?

See this video's description for links to supplemental documents and exercises!

Until Next Time, HackerFrogs!