Karandeep Jaswal

3/22/2023

INCS 712 – Computer Forensics

Assignment 1

Table of Contents

Part 1: Research Based

1. **Explain what is file signature and file header** - Refer to https://www.garykessler.net/library/file_sigs.html

A file signature (or magic number) is a sequence of bytes at the beginning of the file that gives information about the file's type or format (Threatdotmedia, 2022). Typically, different file types have different file signatures. For example, the file signature for a generic JPEG image file is "FF D8".

FF D8		ÿØ
	JPE, JPEG, JPG	Generic JPEGimage file
		Trailer: FF D9 (ÿÙ)

Whereas a PDF file has a file signature of "25 50 44 46".

25 50 44 46		%PDF
	PDF, FDF, AI	Adobe Portable Document Format, Forms Document Format, and Illustrator graphics files

The file signature helps the operating system identify a file's type so that an appropriate application can be used to handle the file. This is helpful because different file types have different internal structures and require different applications to open and edit them. For example, when opening a file that contains a signature of "FF D8", Windows will analyze the file's signature and will know to open the file with the Photos application since the file's signature will match the signature of a JPEG image file. However, if the file has a file signature of "25 50 44 46", Windows will open the Adobe Reader application to open the PDF file.

The file signatures are located in the file header. A file header is a sequence of bytes at the beginning of the file that contains the metadata about the file (NIST, n.d.). Different files have different file headers. For example, the file header for a PDF file can contain information about the file type, size, date created, etc. In contrast, an MP3 audio file header can contain information about the name, tagging format, compression information, etc. (Christensson, 2012).

2. **Explain Data Carving and its techniques**

Data Carving is a forensic technique used to extract data (file) from undifferentiated blocks (raw data) (Merola, 2008). It's called data "carving" because the files are extracted or "carved" from a larger block of data. The unallocated file system space on a disk (for example, a HDD, SSD, USB stick, etc) is analyzed to extract files. The unallocated space is the area of a disk that is not currently being used by any files. So, this technique helps to recover any hidden, deleted or corrupted files on a disk. It does this by scanning the raw bytes of the unallocated space and reassembling them to reconstruct files (Warlock, 2018). The file signatures (including the headers – the first few bytes and footers- the last few bites) are examined to identify and reconstruct and recover any damaged files.

There are multiple techniques for data carving (Warlock, 2018):

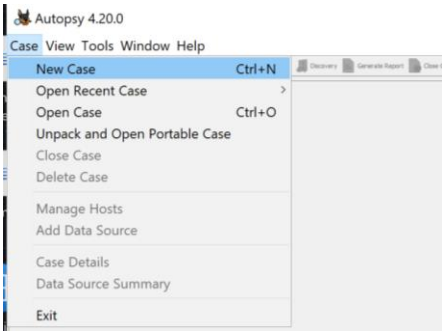
1) Header-Footer Carving: Recovers file contents by looking for known headers and footers (to recognize the start and end of a file) or maximum file size. For example, JPEG files have a header of "xFFxD8″ and "xFFxD9" as footer, and GIF files have a header of "x47x49x46x38x37x6" and "x00x3B" as footer.

2) File Structure based carving: Recovers file contents by searching for the internal layout of the file. Elements used to recover files using this data carving technique include header, footer, identifier strings, and size information.
3) Content based carving: Elements used to recover files using this data carving technique include character count, text/language recognition, white and black listing of data, statistical attributes (Chi^2), etc. This is often used for files that do not have a well-defined signature, such as multimedia files.

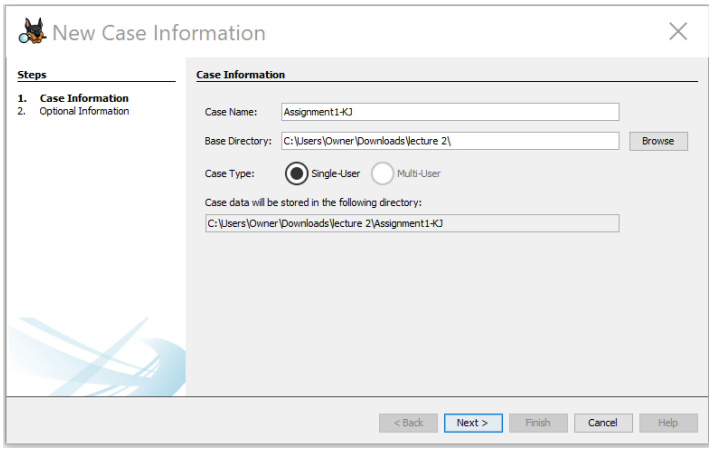
Part 2: Practice-based

3. **Import two dd image files to Autopsy as 'unallocated space disk image' and run ‘Ingest Module’ with ‘PhotoRec Carver' enabled**

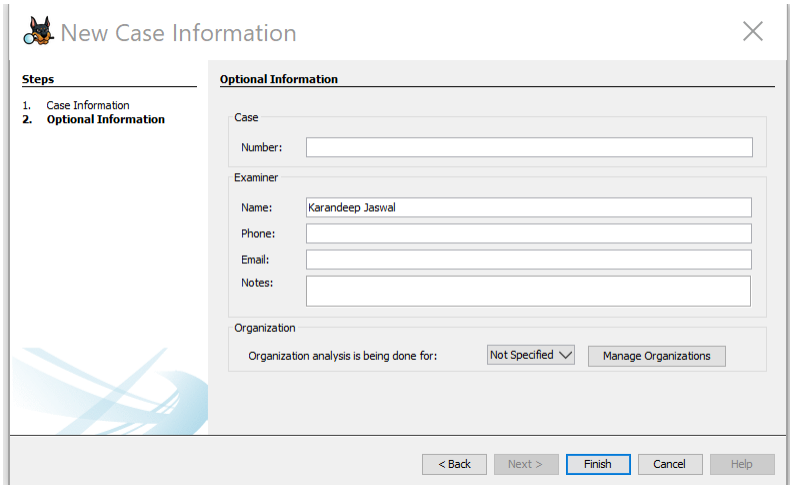
- Launch Autopsy and open a new case.



- Give the case a name and select a base directory. Then, click next.



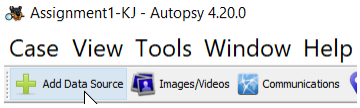
- Enter any optional information. In this case, the examiner’s name was filled out.



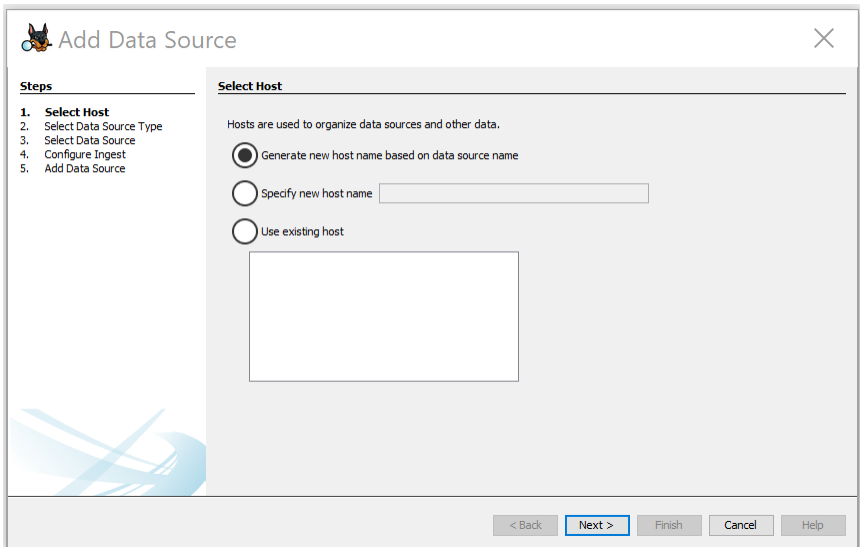
Import L0 dd image file

NOTE: For simplicity, in this paper L0_Graphic.dd file will be called L0 and L2_Graphic.dd file will be called L2.

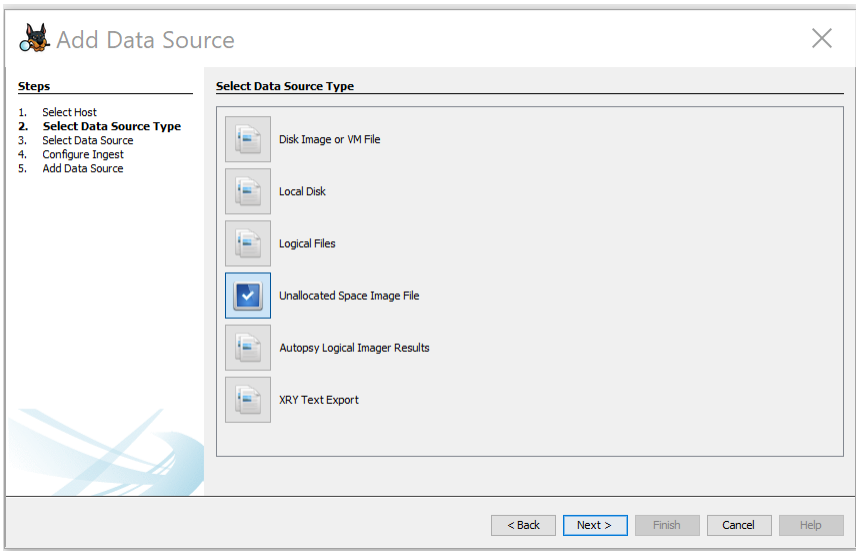
- Click on Add Data Source.



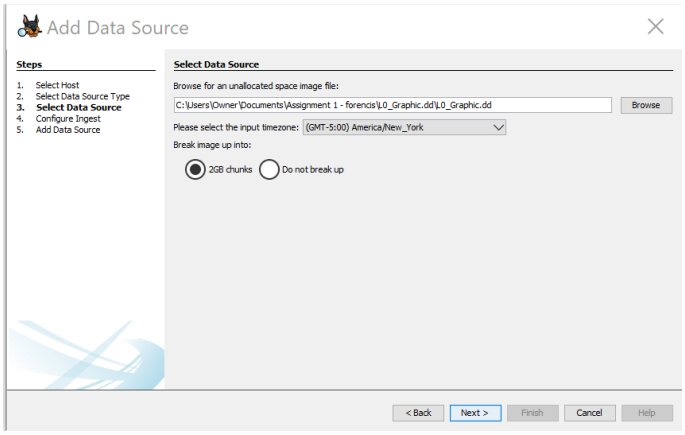
- Click on "Generate new host name based on source name" and click Next.



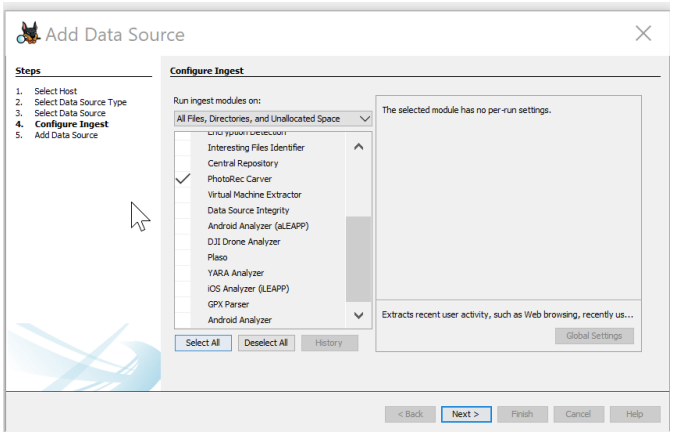
- Click on "Unallocated Space Image File" to import L0 and click Next.



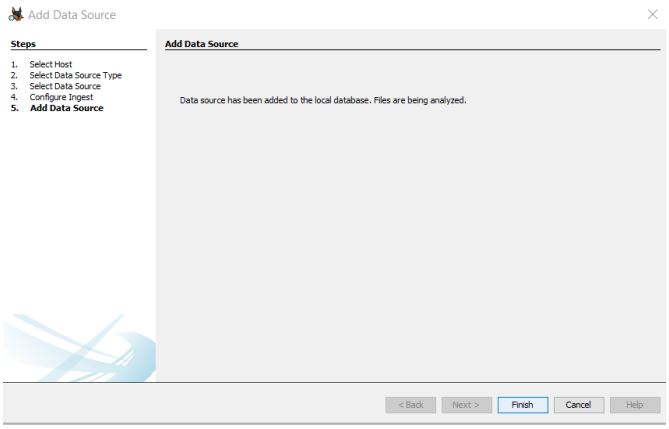
- Select the unzipped L0 file and click Next.



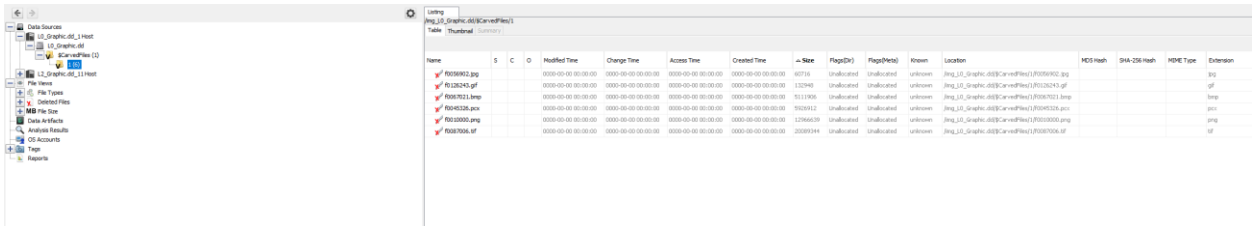
- Select the PhotoRec Carver and click Next.



- Click Finish

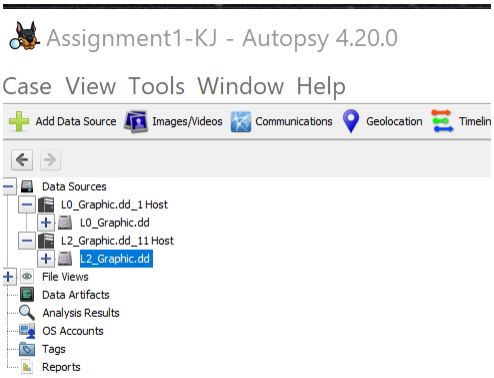


- The L0 dd image file has been successfully added to Autopsy.



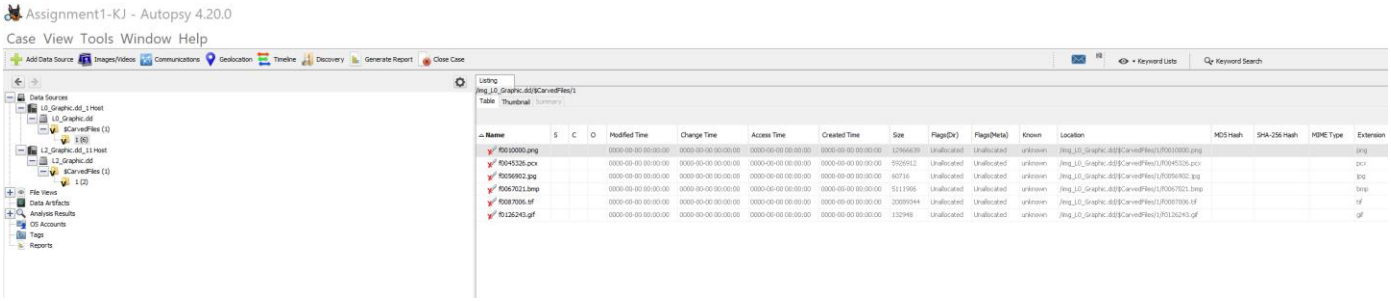
- Following the same procedure, the L2 dd image file was also added to Autopsy.

Both L0 and L2 were added to Autopsy.



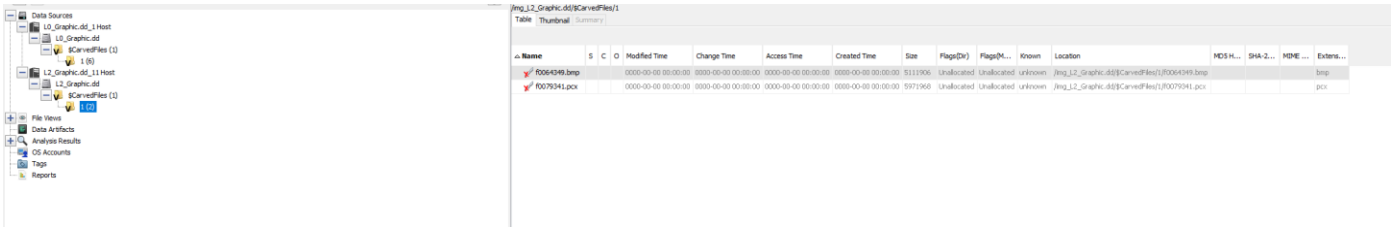
a) List all carved files from each dd image file

- The following 6 files were carved from the L0 dd image file:



Name	S	C	O	Modified Time	Change Time	Access Time	Created Time	Size	Flags(Dir)	Flags(Meta)	Known	Location	MD5 Hash	SHA-256 Hash	MIME Type	Extension
f0010000.png				0000-00-00 00:00:00	0000-00-00 00:00:00	0000-00-00 00:00:00	0000-00-00 00:00:00	12966639	Unallocated	Unallocated	unknown	/img_L0_Graphic.dd/$CarvedFiles/1/f0010000.png				png
f0045326.pcx				0000-00-00 00:00:00	0000-00-00 00:00:00	0000-00-00 00:00:00	0000-00-00 00:00:00	5926912	Unallocated	Unallocated	unknown	/img_L0_Graphic.dd/$CarvedFiles/1/f0045326.pcx				pcx
f0056902.jpg				0000-00-00 00:00:00	0000-00-00 00:00:00	0000-00-00 00:00:00	0000-00-00 00:00:00	60716	Unallocated	Unallocated	unknown	/img_L0_Graphic.dd/$CarvedFiles/1/f0056902.jpg				jpg
f0067021.bmp				0000-00-00 00:00:00	0000-00-00 00:00:00	0000-00-00 00:00:00	0000-00-00 00:00:00	5111906	Unallocated	Unallocated	unknown	/img_L0_Graphic.dd/$CarvedFiles/1/f0067021.bmp				bmp
f0087006.tif				0000-00-00 00:00:00	0000-00-00 00:00:00	0000-00-00 00:00:00	0000-00-00 00:00:00	20089344	Unallocated	Unallocated	unknown	/img_L0_Graphic.dd/$CarvedFiles/1/f0087006.tif				tif
f0126243.gif				0000-00-00 00:00:00	0000-00-00 00:00:00	0000-00-00 00:00:00	0000-00-00 00:00:00	132948	Unallocated	Unallocated	unknown	/img_L0_Graphic.dd/$CarvedFiles/1/f0126243.gif				gif

- The following 2 files were carved from the L2 dd image file:



/img_L2_Graphic.dd/$CarvedFiles/1

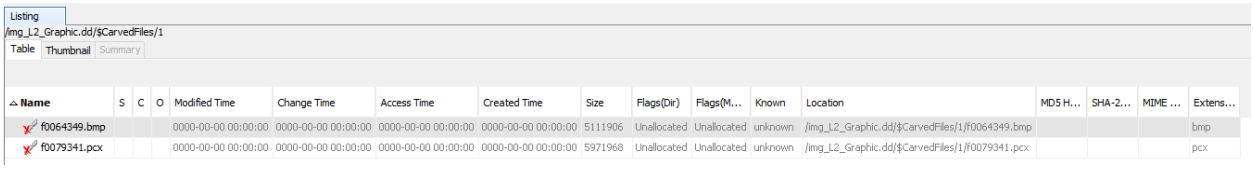
Table | Thumbnail | Summary

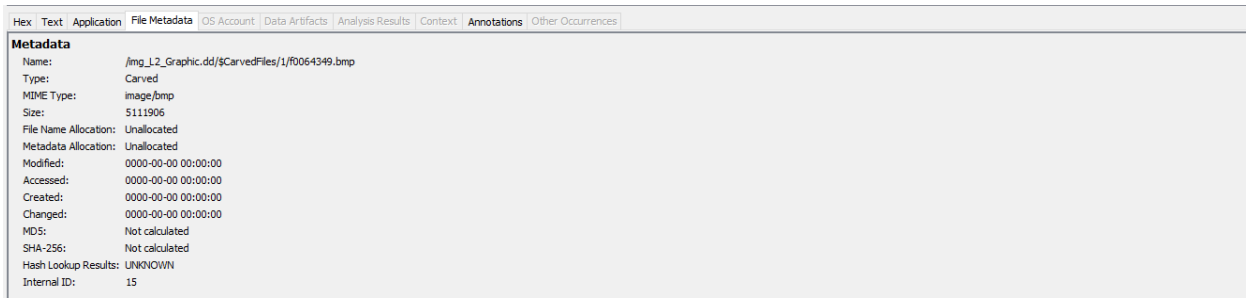
Name	S	C	O	Modified Time	Change Time	Access Time	Created Time	Size	Flags(Dir)	Flags(M...	Known	Location	MD5 H...	SHA-2...	MIME ...	Extens...
f0064349.bmp				0000-00-00 00:00:00	0000-00-00 00:00:00	0000-00-00 00:00:00	0000-00-00 00:00:00	5111906	Unallocated	Unallocated	unknown	/img_L2_Graphic.dd/$CarvedFiles/1/f0064349.bmp				bmp
f0079341.pcx				0000-00-00 00:00:00	0000-00-00 00:00:00	0000-00-00 00:00:00	0000-00-00 00:00:00	5971968	Unallocated	Unallocated	unknown	/img_L2_Graphic.dd/$CarvedFiles/1/f0079341.pcx				pcx

b) Choose a carved file from both dd images that has the same extension and file size

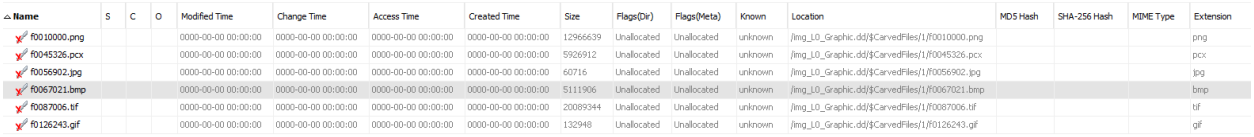
From both dd images, the file **f0064349.bmp** in L2 and the file **f0067021.bmp** in L0 have the same extension (bmp) and file size (5111906). From the File Metadata section, it can be observed the File MIME type (media type) is image/bmp and File Size is 5111906 for both files.

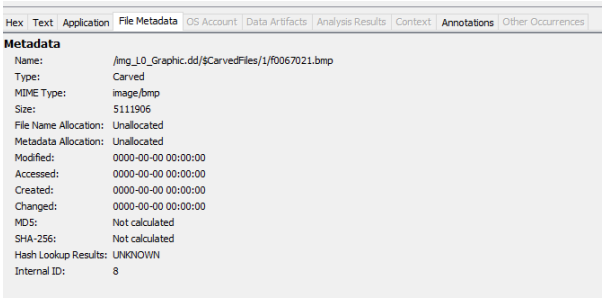
- The file **f0064349.bmp** in L2:

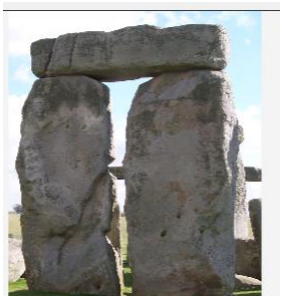




- The file **f0067021.bmp** in L0:







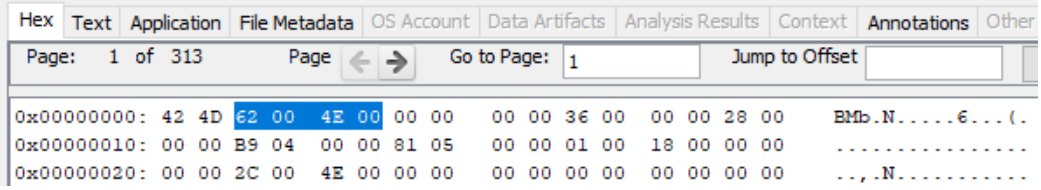
c) Show the header value indicating file size in Hex. Show its size in decimal using a converter.

From the resource give in part one (https://www.garykessler.net/library/file_sigs.html), we know BMP files' bytes 2-5 give information about the file's length (or size) in little-endian order.

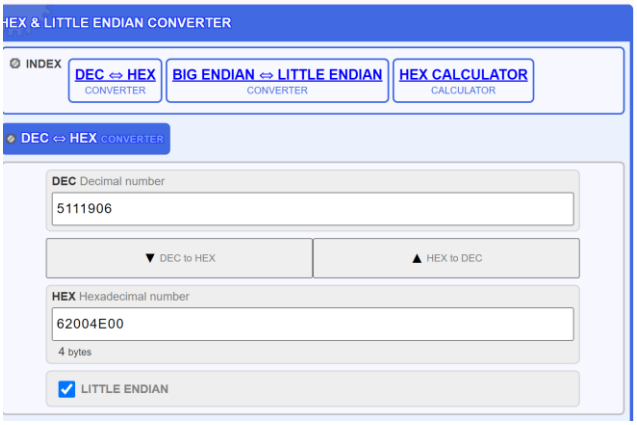
42 4D	BM
BMP, DIB	Windows (or device-independent) bitmap image **NOTE:** Bytes 2-5 contain the file length in little-endian order.

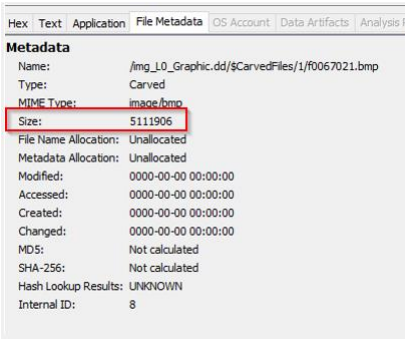
- Go to the hex section and analyze the header value of the file **f0067021.bmp** in L0. Bytes 2 to 5 give information about the file size.

Hex for bytes 2-5: "62 00 4E 00" (this was the same hex value found in **f0064349.bmp** in L2 – since the file size of both the files is the same, this observation was justified).



- Using an online converter, we confirmed the hex value (in little-endian order) "62 00 4E 00" is 5111906 in decimal. 5111906 is the same file size found in part (b).

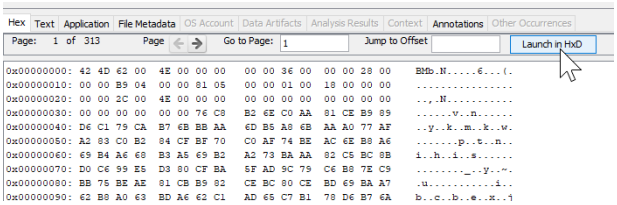




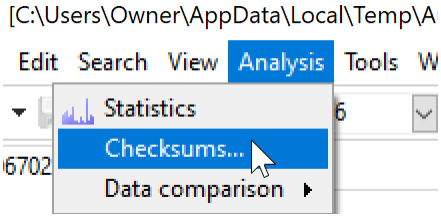
d) Do you think whether these 2 files are originally same or not? Why?

These files are different because when analyzing the checksum of both files, the checksum values were different.

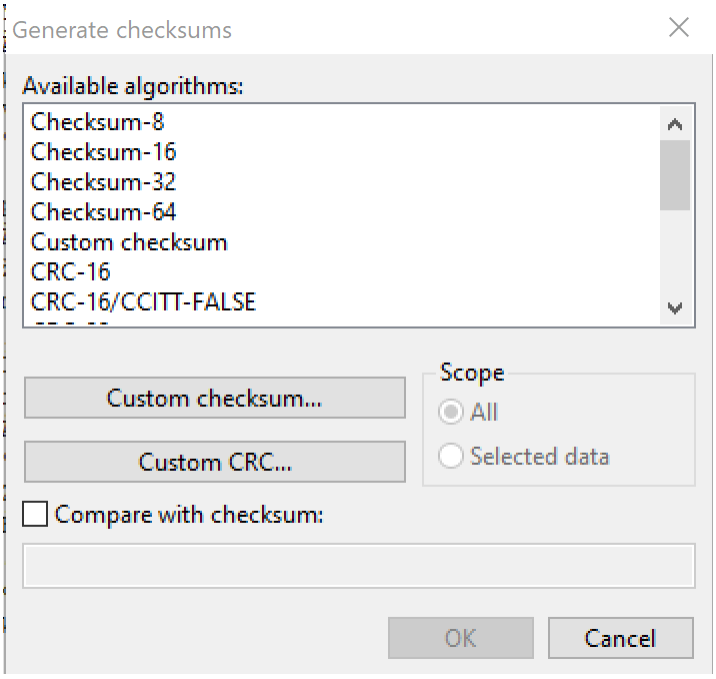
- To find the checksum, click on the bmp file and click “Launch in HxD.” Open both bmp files.



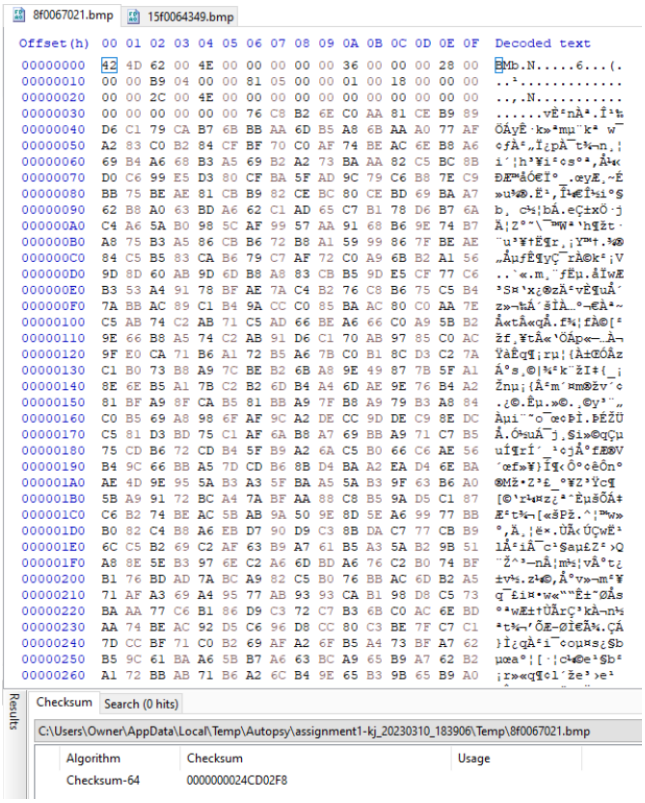
- Click on Checksum

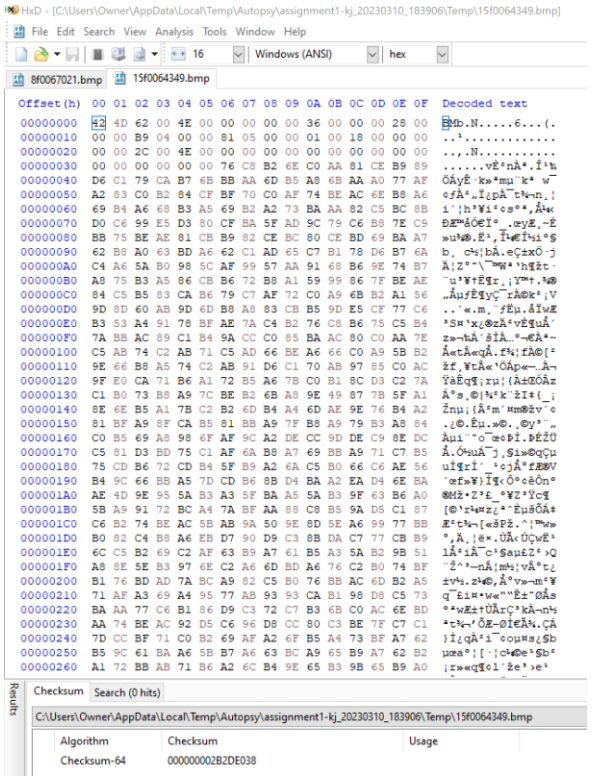


- Chose a checksum algorithm – I chose Checksum – 64.

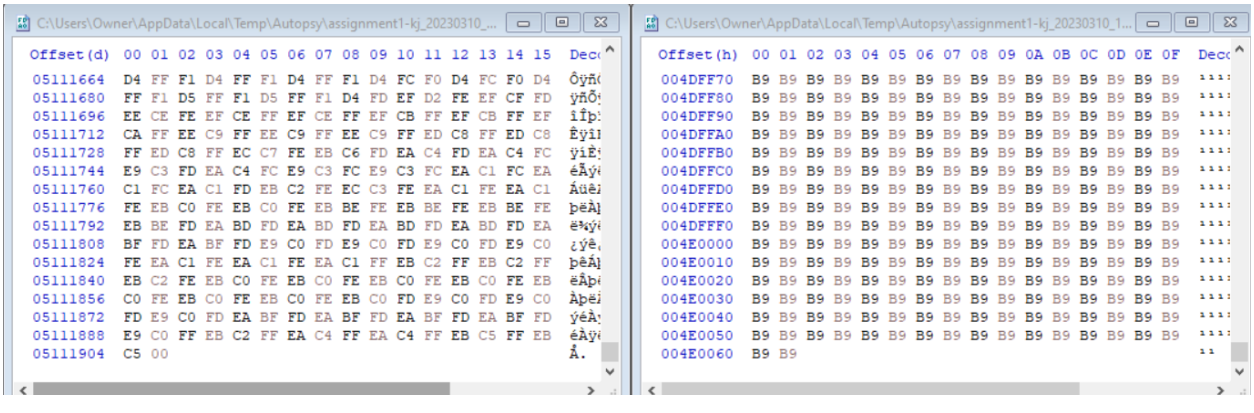


- The checksum for both files was different, meaning the files were not the same.





Also, by going down to the footers of these bmp files, it can be observed the files have different byte values:



References

Christensson, P. (2012, October 2). *Header*. Definition. Retrieved from https://techterms.com/definition/header#:~:text=File%20Header,-A%20file%20header&text=For%20example%2C%20the%20file%20header,tagging%20format%2C%20and%20compression%20information.

Merola, A. (2008). *Data Carving Concepts*. Retrieved from https://www.giac.org/paper/gcfa/1161/data-carving-concepts/110685

NIST. (n.d.). *File header - glossary: CSRC*. COMPUTER SECURITY RESOURCE CENTER. Retrieved from https://csrc.nist.gov/glossary/term/file_header

Threatdotmedia. (2022). *What is a file signature? - definition by threatdotmedia*. ThreatDotMedia - Cyber Explained in Simple Terms. Retrieved from https://threat.media/definition/what-is-a-file-signature/

Warlock. (2021, May 27). *File carving*. Infosec Resources. Retrieved from https://resources.infosecinstitute.com/topic/file-carving/