Module	CMP209 – Digital Forensics
Module Lecturer	Dr Karl van der Schyff
Lab No	6
Due Date	complete before your next lab.

LAB INSTRUCTIONS

WHERE (and what) TO SUBMIT?	• Note that this lab is formative in nature. In other words, although you are expected to complete it, I do not require you to hand it in for assessment. Having said this, I am more than happy to give feedback if you wish to receive it.
WHAT ABOUT USING AI?	• Use of Generative AI Tools such as ChatGPT, DALL-E, Bard etc. is explicitly prohibited for this module's assessment (i.e., the court report). The output gleaned or generated from the tools mentioned and others that are similar relates to sources already published/available. **Also, be aware that the information obtained can be inaccurate or incomplete. Thus, all work should be your own.** If the assessment (i.e., court report) is found to have been plagiarised or to have used unauthorised AI tools, you will be referred to the Student Disciplinary Officer within the School and this may result in an Academic Misconduct charge.

LAB REQUIREMENTS

WHAT DO I NEED TO COMPLETE IT?	• Access to the analysis workstation, which is an Ubuntu VM that can be accessed via VMware Workstation. • You can also download the analysis VM from MLS via the tile labelled "*OneDrive link to analysis VM download files*". • Access to MLS to download the above (and the supplementary docs), but also the John Doe and Windows XP image file. The image files are labelled *johnDoe.dd.gz* and *winXPPostInstall.dd.gz* respectively.

AIM OF THIS LAB

The aim of this lab is to:

- Develop a whitelist to reduce the possible evidence on John Doe's disk image.

CORE LEARNING OUTCOMES

- Increased understanding and competence using the Linux operating system to perform digital forensics.
- Familiarization with the processes required to conduct a digital forensic investigation (read the supplementary docs for this).

- **How to reduce the number of files in your investigation that need to be manually inspected by filtering out those that are on a “whitelist” of known good files.**

Logical searching – Part 1 – Hashing (and mounting) the johnDoe.dd image

1. Check the MD5 hash of your johnDoe.dd file. It should be unchanged from when you first created it.

2. Create a mount point directory within your home folder. You will need this directory to exist before you execute the third command in step 3 below.

```
cd ~
mkdir suspectDrive
```

3. Use the loopback mounting technique to make the files in your johnDoe.dd image available. Note that the command assumes that you have placed the image file into `~/jd`. Adjust the path if you placed the file somewhere else. Also, note that each line below contains a separate command.

```
cd ~/jd
sudo losetup -o 32256 /dev/loop30 johnDoe.dd
sudo mount -o ro -t ntfs /dev/loop30 ~/suspectDrive
```

4. You can now access John Doe's c: drive via `~/suspectDrive` by running the following commands:

```
cd ~/suspectDrive
ls
```

5. Explore the disk (i.e., the image file) via the command line or a file manager.

6. Using the utility `md5deep` create a file containing the MD5 hash and full path/filename of every file on John Doe's disk:

```
cd ~/suspectDrive
md5deep -r * > /home/cmp209/jd/johnDoe.nameAndmd5List.txt
```

Note that the above command has to the run from within the mounted suspect drive. If not, you will be creating a hash of the local filesystem.

7. If md5deep isn't installed on your system, you can add it as follows. Having said this, it should be installed on the Ubuntu analysis VM.

```
sudo apt install hashdeep
sudo ln -s /usr/bin/hashdeep /usr/local/bin/md5deep
```

8. Always remember to unmount your evidence drive, which in this case is your johnDoe image. Also remove the loopback device. If you receive an error message that target device is busy, you could also restart your VM.

```
sudo umount ~/suspectDrive
sudo losetup -d /dev/loop30
```

Logical searching – Part 2 – Creating a whitelist

In this part, you will produce a list of MD5 hashes of all the files within a fresh installation of Windows XP, which happens to be the OS that John Doe used.

9. As stated in the lab requirements, download the Windows XP image (labelled `winXPPostInstall.dd.gz`) from MLS. Copy the file to the `~/jd` directory and extract it there. Follow the same extraction procedures we used in step 2 of Week 5's lab.

10. Using the same approach as in part 1 above (i.e., loopback mounting and md5deep) create a file (called `winXP.nameAndmd5List.txt`) in your `jd` directory which contains the hash and path of every file in the clean Windows XP installation. Note that, instead of mounting the John Doe image you should be mounting the Windows XP image.

11. A whitelist file should contain just the hashes and not the paths/filename. Use the Linux `cut` command to remove the filenames from `winXP.nameAndmd5List.txt`.

    ```
    cut -f 1 -d ‘ ‘ winXP.nameAndmd5List.txt > winXP.md5List.txt
    ```

    Note that if you simply copy the above command to the Ubuntu command line you may receive an error due to the fact that single quotes are represented differently in Ubuntu. If this happens, copy the command as usual and manually retype the single quotes from within Ubuntu (i.e., at the command line).

Logical searching – Part 3 – Narrowing the search using a whitelist

12. Compare the two hash lists to find all the jpg and bmp image files (*.jpg, *.bmp) in the suspect image that aren't part of a standard windows installation. Note that the command below will have to be expanded to be fully effective.

    ```
    grep -v -f winXP.md5List.txt johnDoe.nameAndmd5List.txt
    ```

Logical searching – Part 4 – Doing the above with Autopsy

13. Once you have completed the above, investigate the hashing, whitelisting and filtering capabilities of `Autopsy`.

Additional questions

14. How would you use the above approach to filter out the files of applications that an analyst will typically encounter (e.g., Microsoft Office, Adobe Reader, etc.)?

15. How would you adapt this approach to identify which files have changed (or been added) after installing a new application?

16. How would you adapt the approach to assess the forensic soundness of a new forensic tool?