R V COLLEGE OF ENGINEERING

Name: HASIFA A S **USN:** 1RV18IS016 **Dept/Lab:** ISE/CSDF **Expt No.:** 08 a
Date: 14/12/2021 **Title:** INFORMATION GATHERING TOOL

---

a. METAGOOFIL

INTRODUCTION

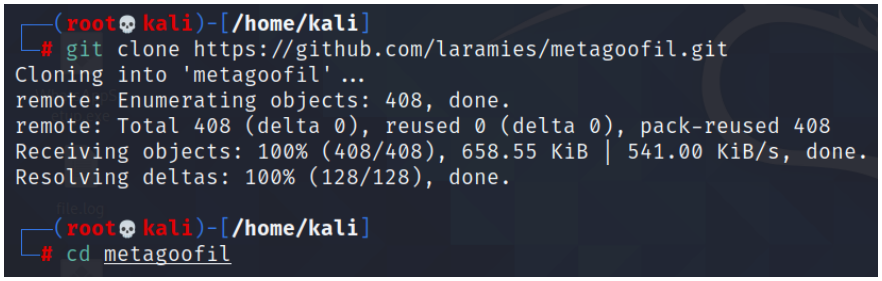
Metagoofil is an information-gathering tool. It is a tool for extracting metadata of public documents (pdf,doc,xls,ppt,etc) availables in the target websites.This information could be useful because you can get valid usernames, people names, for using later in brute force password attacks (vpn, ftp, webapps), the tool will also extracts interesting “paths” of the documents, where we can get shared resources names, server names, etc.
The tool first perform a query in Google requesting different file types that can have useful metadata (pdf, doc, xls,ppt,etc), then will download those documents to the disk and extracts the metadata of the file using specific libraries for parsing different file types (Hachoir, Pdfminer, etc).
It will generate a html page with the results of the metadata extracted, plus a list of potential usernames very useful for preparing a bruteforce attack on open services like ftp, pop3,web applications, vpn and so on. Also it will extract a list of disclosed PATHs in the metadata, with this information you can guess OS, network names, shared resources etc.
This new version extracts MAC addresses from Microsoft Office documents.

Objectives - To perform a search belonging to a target company to identify and download the documents to local disk.

EXECUTION STEPS

1. **Installing Metagoofil from a package**
   Command - *sudo apt install metagoofil*

2. **Installing Metagoofil using Git**
   Command - *git clone https://github.com/laramies/metagoofil.git*



3. **To run Metagoofil**
   Command - *python metagoofil.py*

```
┌──(root💀kali)-[/home/kali/metagoofil]
└─# python metagoofil.py

*******************************************************
*  /\/\   ___| |_ __ _  __ _  ___   ___  / _(_) | *
*    /    \ / _ \ __/ _` |/ _` |/ _ \ / _ \| |_| | | *
*   / /\/\ \  __/ || (_| | (_| | (_) | (_) |  _| | | *
*   \/    \/\___|\__\__,_|\__, |\___/ \___/|_| |_|_| *
*                         |___/                       *
* Metagoofil Ver 2.2                                  *
* Christian Martorella                                *
* Edge-Security.com                                   *
* cmartorella_at_edge-security.com                    *
*******************************************************

 Usage: metagoofil options

        -d: domain to search
        -t: filetype to download (pdf,doc,xls,ppt,odp,ods,docx,xlsx,pptx)
        -l: limit of results to search (default 200)
        -h: work with documents in directory (use "yes" for local analysis)
        -n: limit of files to download
        -o: working directory (location to save downloaded files)
        -f: output file
```

Example cases

1. **To extract pdfs and information from a website** -
   Scan for documents from a domain (*-d*) that are PDF files (*-t*), searching 10 results (*-l*), download 2 files (*-n*), saving the downloads to a directory (*-o*), and saving the output to a file (*-f*):
   Syntax - *python metagoofil.py -d <URL> -l <Number_of_Results> -n <Number_of_results_to_download> -t <file_type> -o <target_directory>*

   Command - *python metagoofil.py -d rvce.edu.in -l 10 -n 2 -t pdf -o rvceOutput*

```
┌──(root💀kali)-[/home/kali/metagoofil]
└─# python metagoofil.py -d rvce.edu.in -l 10 -n 2 -t pdf -o rvceOutput

*******************************************************
*  /\/\   ___| |_ __ _  __ _  ___   ___  / _(_) | *
*    /    \ / _ \ __/ _` |/ _` |/ _ \ / _ \| |_| | | *
*   / /\/\ \  __/ || (_| | (_| | (_) | (_) |  _| | | *
*   \/    \/\___|\__\__,_|\__, |\___/ \___/|_| |_|_| *
*                         |___/                       *
* Metagoofil Ver 2.2                                  *
* Christian Martorella                                *
* Edge-Security.com                                   *
* cmartorella_at_edge-security.com                    *
*******************************************************
['pdf']

[-] Starting online search...

[-] Searching for pdf files, with a limit of 10
        Searching 100 results...
Results: 102 files found
Starting to download 2 of them:
_____________________________________
```

2. **To perform local directory analysis**
   Scan documents in a directory. Use “yes” for local analysis.
   Syntax - *python metagoofil.py -h yes -o <directory> -f <output_filename>*

   Command - *python metagoofil.py -h yes -o rvceOutput -f results.html*

CONCLUSION

1. Metagoofil is an information gathering tool designed for extracting metadata of public documents (pdf,doc,xls,ppt,odp,ods) available on the target/victim website.
2. As seen above, the tool will find all the details of the website, if any, available on the website.

REFERENCES

1. Metagoofil - https://www.kali.org/tools/metagoofil/
2. Metagoofil – Tool to Extract Information from Docs, Images in Kali Linux - https://www.geeksforgeeks.org/metagoofil-tool-to-extract-information-from-docs-images-in-kali-linux/
3. Metagoofil Cyber Security Tool - https://iemlabs.com/metagoofil/