

# Search Engines [GHDB] & Metadata Tools



**LinkedIn** : <https://www.linkedin.com/in/mohamed-elsayaad>

**GitHub** : <https://github.com/0xDos>



# Search Engines (1/2)

- Google Hacking Data Base [GHDB]
- The common operators (*AND, OR, +, -, ""*)
- [*link:www.website.com*]
- [*site:www.website.com*]
- [*intext:www.website.com*]
- [*cache:www.website.com*]
- [*filetype:pdf*] -filetype:html
- **intitle:"index of" "parent directory"**

# Footprinting using Advanced Google Hacking Techniques



- ❑ **cache:**  
Displays Google's cached version
- ❑ **link:** هاجبلي أي لينك موجود فيه الدومين ال عاوزه  
Show a list of web pages that have links to your target
- ❑ **related:**  
Similar web pages
- ❑ **info:**  
View information Google has on the target
- ❑ **site:** هاجبلي Subdomains الخاصة ب Target  
Limits results to just the domain listed

- ❑ **allintitle:**  
Limits results to those websites with ALL the search words in the title
- ❑ **intitle:**  
Limits results to documents that contain the search word in the title
- ❑ **allinurl:**  
Limits results to only those webpages with ALL search words in the URL
- ❑ **inurl:**  
Limits results to documents that contain the search word in the URL

يمكن تقسيم وظائف Google Dorks إلى قسمين: أساسي ومتقدم.

## بعض وظائف Google Dorks الأساسية:

هي عبارة عن أوامر مقدمة من طرف غوغل لتعميق عملية البحث وجعلها أكثر فعالية.

أمثلة:

- **Site:** يستخدم هذا الاستعلام للبحث في موقع محدد، مثلا لو كتبت في مربع البحث عبارة site:pcfatile1.com سوف يعرض لك جميع نتائج التي تخص مدونة حاسب سهل فقط.

- **Link:** يستخدم هذا الاستعلام للكشف عن جميع الصفحات التي تحتوي على رابط الموقع، مثلا link:pcfatile1.com هنا سوف يعرض جميع الصفحات التي تحتوي على رابط مدونة حاسب سهل.

- **Info:** يعرض هذا الاستعلام المواقع التي تقدم معلومات عن الموقع الذي تبحث عنه، مثلا لو أردت معلومات عن موقع ياهو أكتب info:yahoo.com.

- **Define:** هذا الاستعلام لعرض تعريف المصطلحات، مثلا Define:hackers لعرض تعريف الهكرز.

- **Filetype:** هذا الاستعلام للبحث عن ملفات محددة بامتداد معين (- ppt - xls - pdf - rtf - swf - asp - txt - doc...), مثلا filetype:nature.ppt هنا سوف يعرض جميع الملفات التي تتكلم عن الطبيعة داخل ملفات بوربونت.

- **Intitle:** هذا الاستعلام يقوم بالبحث في وسوم عناوين الصفحات، مثلا intitle:security يعرض جميع الصفحات التي تحمل في عناوينها كلمة security.

- **Inurl:** هذا الاستعلام يقوم بالبحث في روابط المواقع، مثلا inurl:admin سوف يعرض جميع الروابط التي تحتوي على كلمة admin.

- **And (+):** هذا الاستعلام للبحث عن كلمات مختلفة، مثلا nature+world+earth أو nature And world And earth سوف يعرض الصفحات التي تحتوي على 3 كلمات المفتاحية المذكورة (بدل كتابة جملة طويلة، استخدم هذا الاستعلام وأكتب كلمات مفتاحية فقط).

- **OR (|):** يخبر هذا الاستعلام غوغل بعرض على نتائج أحد الكلمات التي قد يعثر عليها.

- **NOT:** هذا الاستعلام لاستبعاد كلمة من الظهور في نتائج البحث.

## بعض وظائف Google Dorks المتقدمة:

هي دمج أوامر أو الاستعلامات الأساسية المقدمة من غوغل للوصول إلى صفحات حساسة أو روابط تقود إلى صفحة الأدمن أو الكشف عن الثغرات.

أمثلة:

- مثلا للبحث عن الكاميرات النشطة على الإنترنت أكتب في شريط البحث: inurl:view/view.shtml أو inurl:ViewerFrame?Mode= أو inurl:/view.shtml أو inurl:"Active Webcam Page" inurl:8080 وهناك العديد غيرها.

- مثلا هذا الاستعلام يقود إلى صفحات تسجيل الدخول بالنسبة للأدمن: intext:"Please Authenticate" intitle:Peakflow

- هذا الاستعلام للإطلاع على الحسابات وكلمات السر المخترقة بالنسبة لحسابات فايسبوك وحسابات البريد الإلكتروني عن طريق [الصفحات المزورة](#):

inurl:"passes" OR inurl:"pass" OR inurl:"passwords" OR  
inurl:"credentials"-search-download-techsupt-git-games-gz-bypass  
-exe filetype:txt @yahoo.com OR @gmail OR @hotmail OR @rediff

هناك مواقع متخصصة في عرض آخر أوامر Google Dorks يمكنك الاطلاع عليها من أحد المواقع التالية:

- [exploit-db](#)

- [hackersforcharity](#)

Examples:

site:learningnetwork.cisco.com intext:yasser pdf

site:\* inurl:etc -intext:etc ext:passwd

intitle:webcam 7 inurl:8080 -intext:8080 intitle:"live

view / -axis"

**Intext:**

هاخذ التوقيع ال بيكون في نهاية صفحة site واسيرش بيه مع intext علشان دا فايده :  
هايجبلي المواقع التابعه للهدف والمتاح ليها استخدام نفس التوقيع دا



<https://www.exploit-db.com/google-hacking-database/> <https://www.offensive-security.com/community-projects/google-hacking-database/>  
[https://www.google.com/advanced\\_search](https://www.google.com/advanced_search)  
<http://www.googleguide.com/phonebook.html>



ASN Lookup Tool | UltraTool | ASN discovery - GitBook | Hurricane Electric BGP Tool | Google Hacking Database | Grab - Transport, Food Deli | استكشف تطبيق أوبر | أوبر

exploit-db.com/google-hacking-database

EXPLOIT DATABASE

GET CERTIFIED

Google Hacking Database

Filters Reset All

Show 15 Quick Search

Date Added Dork

		Category	Author
2020-07-27	inurl:/"vam/index_vam_op.php"	Advisories and Vulnerabilities	Alexandros Pappas
2020-07-27	"Share Link" inurl:/share.cgi?ssid=	Sensitive Directories	Alexandros Pappas
2020-07-26	Index of : wp-content/plugins/wpmudev-updates/	Advisories and Vulnerabilities	Pratik Khalane
2020-07-26	inurl:/+CSCOE+/logon.html?	Pages Containing Login Portals	Supun Halangoda
2020-07-26	intext:"Frame rate" inurl:/home/homej.html	Various Online Devices	Alexandros Pappas
2020-07-26	inurl:wp-content/plugins/my-calendar	Advisories and Vulnerabilities	Lokesh S
2020-07-26	intitle:"index of" /lsass.exe	Sensitive Directories	Prasad Lingamaiah
2020-07-26	inurl:wp-content/plugins/updraftplus	Advisories and Vulnerabilities	Lokesh S
2020-07-26	inurl:wp-content/plugins/redirection	Advisories and Vulnerabilities	Lokesh S
2020-07-26	intitle:ePMP 1000 intext:Log In -site:*.com -site:com.*	Advisories and Vulnerabilities	cyb3rmx0
2020-07-26	intext:"Device Name"   intext:"Host Name" inurl:mainFrame.cgi	Various Online Devices	Alexandros Pappas
2020-07-26	site:com "sap netweaver portal"	Pages Containing Login Portals	berat isler
2020-07-26	inurl:/webconsole/webpages/login.jsp	Pages Containing Login Portals	Dharmveer Singh
2020-07-26	inurl:axis-cgi/mjpg/video.swf	Various Online Devices	Sachin Kattimani
2020-07-26	inurl:/home/homej.html	Various Online Devices	Alexandros Pappas

Showing 1 to 15 of 5,959 entries

FIRST PREVIOUS 1 2 3 4 5 ... 398 NEXT LAST

## Advanced Search

## Find pages with...

all these words:

this exact word or phrase:

any of these words:

none of these words:

numbers ranging from:

to

## To do this in the search box.

Type the important words: **tri-colour rat terrier**Put exact words in quotes: **"rat terrier"**Type **OR** between all the words you want: **miniature OR standard**Put a minus sign just before words that you don't want:  
**-rodent, -"Jack Russell"**Put two full stops between the numbers and add a unit of measurement:  
**10..35 kg, £300..£500, 2010..2011**

## Then narrow your results by...

language:

Find pages in the language that you select.

region:

Find pages published in a particular region.

last update:

Find pages updated within the time that you specify.

site or domain:

Search one site (like **wikipedia.org**) or limit your results to a domain like **.edu, .org** or **.gov**

terms appearing:

Search for terms in the whole page, page title or web address, or links to the page you're looking for.

SafeSearch:

Tell **SafeSearch** whether to filter sexually explicit content.

file type:

Find pages in the format that you prefer.

usage rights:

Find pages that you are free to use yourself.

Advanced Search

# Meta Data Extraction tools

- FOCA

**FOCA (Fingerprinting Organizations with Collected Archives)** is a tool used mainly to find metadata and hidden information in the documents its scans

Using FOCA, it is possible to undertake multiple attacks and analysis techniques such as **metadata extraction**, **network analysis**, DNS snooping, proxies search, **fingerprinting**, open directories search, etc.

Metagoofil

ExifTool



## Document Metadata

- Most document formats include a significant amount of metadata (that is, data about data)
  - This metadata is often associated with how the document is formatted for display
  - But, some of the metadata goes further, providing highly useful tidbits for the penetration tester to gather in the reconnaissance phase
- Information sometimes included in metadata:
  - Usernames
  - File system paths
  - E-mail addresses
  - Client-side software in use (Office suite, PDF-generating tool, operating system type, and such)
  - Other information not displayed on the screen from within the application associated with the document ("undo" data, previous revisions, hidden or obscured fields, and more)

## Document Types That Are Rich in Metadata

- Most types of documents have some metadata in them, but the following types are often especially interesting:
  - pdf
  - doc, dot, and docx
  - xls, xlt, and xlsx
  - ppt, pot, and pptx
  - jpg and jpeg
  - html and htm (for example, comments and hidden form elements)
  - Numerous others
- This isn't an exhaustive list, but it is a good start



# FOCA (Fingerprinting Organizations with Collected Archives)

FOCA is a tool used mainly to find metadata and hidden information in the documents it scans. These documents may be on web pages, and can be downloaded and analysed with FOCA.

It is capable of analysing a wide variety of documents, with the most common being **Microsoft Office**, **Open Office**, or **PDF** files, although it also analyses Adobe InDesign or SVG files, for instance.

These documents are searched for using three possible search engines: **Google**, **Bing**, and **DuckDuckGo**. The sum of the results from the three engines amounts to a lot of documents. It is also possible to add local files to extract the EXIF information from graphic files, and a complete analysis of the information discovered through the URL is conducted even before downloading the file.

## 👥 Releases

Check [here](#) our latest releases.

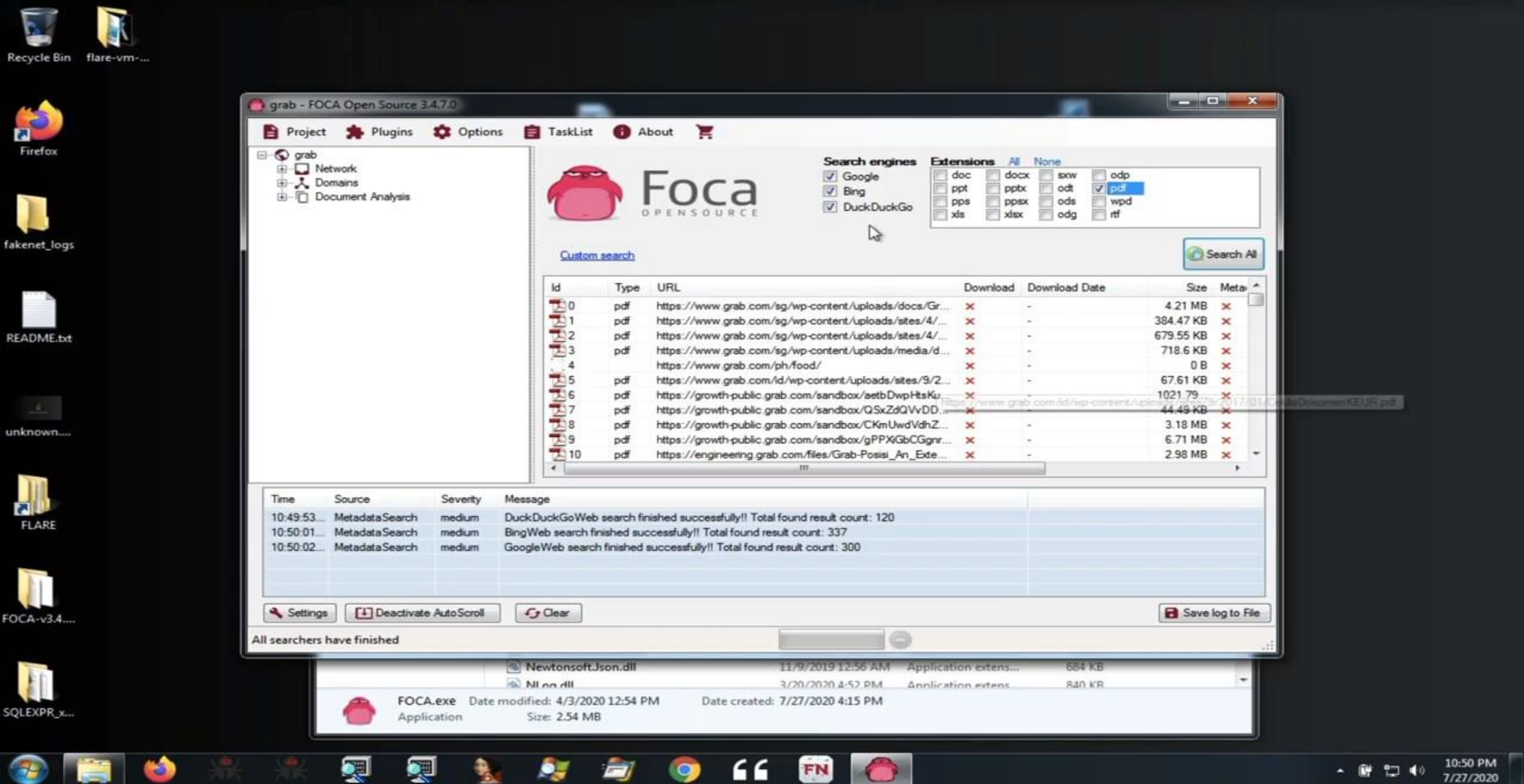
## ✓ Requisites

To run the solution locally the system will need:

- **Microsoft Windows** (64 bits). Versions 7, 8, 8.1 and 10.
- **Microsoft .NET Framework 4.7.1**.
- **Microsoft Visual C++ 2010 x64** or greater.
- An instance of **SQL Server 2014** or greater.

## 📝 Notes

- When starting the app the system will check if there is a **SQL Server** instance available. If none is found, the system will prompt a window for introducing a connection string.





```
kali@kali: ~  
File Actions Edit View Help  
(kali@kali)-[~]  
$ metagoofil --help  
usage: metagoofil.py [-h] -d DOMAIN [-e DELAY] [-f] [-i URL_TIMEOUT] [-l SEARCH_MAX] [-n DOWNLOAD_FILE_LIMIT] [-o SAVE_DIRECTORY] [-r NUMBER_OF_THREADS] -t FILE_TYPES [-u [USER_AGENT]] [-w]  
  
Metagoofil - Search and download specific filetypes  
  
optional arguments:  
-h, --help            show this help message and exit  
-d DOMAIN              Domain to search.  
-e DELAY               Delay (in seconds) between searches. If it's too small Google may block your IP, too big and your search may take a while. Default: 30.0  
-f                    Save the html links to html_links_<TIMESTAMP>.txt file.  
-i URL_TIMEOUT         Number of seconds to wait before timeout for unreachable/stale pages. Default: 15  
-l SEARCH_MAX          Maximum results to search. Default: 100  
-n DOWNLOAD_FILE_LIMIT Maximum number of files to download per filetype. Default: 100  
-o SAVE_DIRECTORY      Directory to save downloaded files. Default is current working directory, "."  
-r NUMBER_OF_THREADS   Number of downloader threads. Default: 8  
-t FILE_TYPES          file_types to download (pdf,doc,xls,ppt,odp,ods,docx,xlsx,pptx). To search all 17,576 three-letter file extensions, type "ALL"  
-u [USER_AGENT]        User-Agent for file retrieval against -d domain.  
                        no -u = "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"  
                        -u = Randomize User-Agent  
                        -u "My custom user agent 2.0" = Your customized User-Agent  
-w                    Download the files, instead of just viewing search results.
```

EX:\$ Metagoofil -d apple.com -t doc,pdf -l 200 -n 20 -o applefiles -f results.html

- ExifTool: Reads, writes, and changes metadata
  - Freely distributed, written by Phil Harvey
    - <http://www.sno.phy.queensu.ca/~phil/exiftool/>
  - Runs on Windows, Linux, and Mac OS X
  - Supports more than 100 file types and many metadata formats
    - Original focus was on image and audio files
    - Many different image types, pulling out camera type, editing tools, and geotags if they are present
    - Now it has been expanded to include many file types, including various document file types (doc, docx, xls, xlsx, ppt, pptx, pdf, and so on)
    - Parses out specific fields, and is handy for determining usernames and software versions used to create or edit files
    - Processes entire directories, with recursion supported

```
Administrator: C:\Windows\system32\cmd.exe
F:\exiftool-10.67>"exiftool(-k).exe" "ECC ilabs.docx"
ExifTool Version Number      : 10.67
File Name                    : ECC ilabs.docx
Directory                   : .
File Size                   : 16 kB
File Modification Date/Time  : 2018:10:08 08:34:10+04:00
File Access Date/Time       : 2018:10:15 22:57:19+04:00
File Creation Date/Time     : 2018:10:15 22:57:05+04:00
File Permissions             : rw-rw-rw-
File Type                   : DOCX
File Type Extension         : docx
MIME Type                   : application/vnd.openxmlformats-officedocument.
wordprocessingml.document
Zip Required Version        : 20
Zip Bit Flag                : 0x0006
Zip Compression             : Deflated
Zip Modify Date              : 1980:01:01 00:00:00
Zip CRC                     : 0x7dec21f0
Zip Compressed Size         : 398
Zip Uncompressed Size      : 1555
Zip File Name               : [Content_Types].xml
Template                    : Normal.dotm
Total Edit Time              : 1 minute
Pages                       : 1
Words                       : 377
Characters                  : 2150
Application                 : Microsoft Office Word
Doc Security                 : None
Lines                       : 17
Paragraphs                  : 5
Scale Crop                  : No
Heading Pairs               : Title, 1
Titles Of Parts             :
Company                     : HP
Links Up To Date            : No
Characters With Spaces      : 2522
Shared Doc                  : No
Hyperlinks Changed          : No
App Version                 : 14.0000
Creator                     : Kanesan
Last Modified By            : Kanesan
Revision Number             : 2
Last Printed                 : 2018:06:27 09:36:00Z
Create Date                 : 2018:10:08 02:36:00Z
```



**NAME**

exiftool - Read and write meta information in files

**SYNOPSIS****Reading**

exiftool [OPTIONS] [-TAG ...] [--TAG ...] FILE ...

**Writing**

exiftool [OPTIONS] -TAG[+<]=[VALUE] ... FILE ...

**Copying**

exiftool [OPTIONS] -tagsFromFile SRCFILE [-[DSTTAG<]SRCTAG ...] FILE ...

**Other**

exiftool [ -ver | -list[w|f|r|wf|g[NUM]|d|x] ]

For specific examples, see the EXAMPLES sections below.

This documentation is displayed if exiftool is run without an input **FILE** when one is expected.

**DESCRIPTION**

A command-line interface to Image::ExifTool, used for reading and writing meta information in a variety of file types. **FILE** is one or more source file names, directory names, or "-" for the standard input. Metadata is read from source files and printed in readable form to the console (or written to output text files with **-w**).

To write or delete metadata, tag values are assigned using **-TAG=[VALUE]**, and/or the **-geotag**, **-csv=** or **-json=** options. To copy or move metadata, the **-tagsFromFile** feature is used. By default the original files are preserved with "\_original" appended to their names -- be sure to verify that the new files are OK before erasing the originals. Once in write mode, exiftool will ignore any read-specific options.

Note: If **FILE** is a directory name then only supported file types in the directory are processed (in write mode only writable types are processed). However, files may be specified by name, or the **-ext** option may be used to force processing of files with any extension. Hidden files in the directory are also processed. Adding the **-r** option causes subdirectories to be processed recursively, but subdirectories with names beginning with "." are skipped unless **-r.** is used.

Below is a list of file types and meta information formats currently supported by ExifTool (r = read, w = write, c = create):

**File Types**

360	r/w		DR4	r/w/c		JNG	r/w		O	r		RAW	r/w
3FR	r		DSS	r		JP2	r/w		ODP	r		RIFF	r
3G2	r/w		DV	r		JPEG	r/w		ODS	r		RSRC	r
3GP	r/w		DVB	r/w		JSON	r		ODT	r		RTF	r
A	r		DVR-MS	r		JXL	r		OFR	r		RW2	r/w
AA	r		DYLIB	r		K25	r		OGG	r		RWL	r/w
AAE	r		EIP	r		KDC	r		OGV	r		RWZ	r
AAX	r/w		EPS	r/w		KEY	r		ONP	r		RM	r
ACR	r		EPUB	r		LA	r		OPUS	r		SEQ	r
AFM	r		ERF	r/w		LFP	r		ORF	r/w		SKETCH	r

Manual page exiftool(1p) line 1 (press h for help or q to quit)