GitHub - decalage2/oletools: oletools - python tools to analyze MS OLE2 files (Structured Storage, Compound File Binary Format) and MS Office documents, for malware analysis, forensics and debugging.

By decalage2

Archived: 2026-04-02 12:00:28 UTC

oletools is a package of python tools to analyze Microsoft OLE2 files (also called Structured Storage, Compound File Binary Format or Compound Document File Format), such as Microsoft Office 97-2003 documents, MSI files or Outlook messages, mainly for malware analysis, forensics and debugging. It is based on the olefile parser.

It also provides tools to analyze RTF files and files based on the OpenXML format (aka OOXML) such as MS Office 2007+ documents, XPS or MSIX files.

For example, oletools can detect, extract and analyse VBA macros, OLE objects, Excel 4 macros (XLM) and DDE links.

See http://www.decalage.info/python/oletools for more info.

Quick links: Home page - Download/Install - Documentation - Report Issues/Suggestions/Questions - Contact the Author - Repository - Updates on Twitter Cheatsheet

Note: python-oletools is not related to OLETools published by BeCubed Software.

News

- **2025-05-22 v0.60.3**:
  - olevba:
    - fixed a security issue in the CLI display when ANSI escape codes are present (PR #873)
    - encrypted files: password is now reported in the logs, added --decrypted_dir option (PR #842)
- **2024-07-02 v0.60.2**:
  - olevba:
    - fixed a bug in open_slk (issue #797, PR #769)
    - fixed a bug due to new PROJECTCOMPATVERSION record in dir stream (PR #723, issues #700, #701, #725, #791, #808, #811, #833)
  - oleobj: fixed SyntaxError with Python 3.12 (PR #855), SyntaxWarning (PR #774)
  - rtfobj: fixed SyntaxError with Python 3.12 (PR #854)

    - clsid: added CLSIDs for MSI, Zed
    - ftguess: added MSI, PNG and OneNote formats
    - pyxswf: fixed python 3.12 compatibility (PR #841, issue #813)
    - setup/requirements: allow pyparsing 3 to solve install issues (PR #812, issue #762)
- **2022-05-09 v0.60.1**:
    - olevba:
        - fixed a bug when calling XLMMacroDeobfuscator (PR #737)
        - removed keyword "sample" causing false positives
    - oleid: fixed OleID init issue (issue #695, PR #696)
    - oleobj:
        - added simple detection of CVE-2021-40444 initial stage
        - added detection for customUI onLoad
        - improved handling of incorrect filenames in OLE package (PR #451)
    - rtfobj: fixed code to find URLs in OLE2Link objects for Py3 (issue #692)
    - ftguess:
        - added PowerPoint and XPS formats (PR #716)
        - fixed issue with XPS and malformed documents (issue #711)
        - added XLSB format (issue #758)
    - improved logging with common module log_helper (PR #449)
- **2021-06-02 v0.60**:
    - ftguess: new tool to identify file formats and containers (issue #680)
    - oleid: (issue #679)
        - each indicator now has a risk level
        - calls ftguess to identify file formats
        - calls olevba+mraptor to detect and analyse VBA+XLM macros
    - olevba:
        - when XLMMacroDeobfuscator is available, use it to extract and deobfuscate XLM macros
    - rtfobj:
        - use ftguess to identify file type of OLE Package (issue #682)
        - fixed bug in re_executable_extensions
    - crypto: added PowerPoint transparent password '/01Hannes Ruescher/01' (issue #627)
    - setup: XLMMacroDeobfuscator, xlrd2 and pyxlsb2 added as optional dependencies

See the full changelog for more information.

Tools:

Tools to analyze malicious documents

- oleid: to analyze OLE files to detect specific characteristics usually found in malicious files.
- olevba: to extract and analyze VBA Macro source code from MS Office documents (OLE and OpenXML).
- MacroRaptor: to detect malicious VBA Macros
- msodde: to detect and extract DDE/DDEAUTO links from MS Office documents, RTF and CSV

- pyxswf: to detect, extract and analyze Flash objects (SWF) that may be embedded in files such as MS Office documents (e.g. Word, Excel) and RTF, which is especially useful for malware analysis.
- oleobj: to extract embedded objects from OLE files.
- rtfobj: to extract embedded objects from RTF files.

Tools to analyze the structure of OLE files

- olebrowse: A simple GUI to browse OLE files (e.g. MS Word, Excel, Powerpoint documents), to view and extract individual data streams.
- olemeta: to extract all standard properties (metadata) from OLE files.
- oletimes: to extract creation and modification timestamps of all streams and storages.
- oledir: to display all the directory entries of an OLE file, including free and orphaned entries.
- olemap: to display a map of all the sectors in an OLE file.

Projects using oletools:

oletools are used by a number of projects and online malware analysis services, including ACE, ADAPT, Anlyz.io, AssemblyLine, Binary Refinery, CAPE, CinCan, Cortex XSOAR (Palo Alto), Cuckoo Sandbox, DARKSURGEON, Deepviz, DIARIO, dridex.malwareconfig.com, EML Analyzer, EXPMON, FAME, FLARE-VM, GLIMPS Malware, Hybrid-analysis.com, InQuest Labs, IntelOwl, Joe Sandbox, Laika BOSS, MacroMilter, mailcow, malshare.io, malware-repo, Malware Repository Framework (MRF), MalwareBazaar, olefy, Pandora, PeekabooAV, pcodedmp, PyCIRCLean, QFlow, Qu1cksc0pe, Tylabs QuickSand, REMnux, Snake, SNDBOX, Splunk add-on for MS O365 Email, SpuriousEmu, Strelka, stoQ, Sublime Platform/MQL, Subparse, TheHive/Cortex, ThreatBoook, TSUGURI Linux, Vba2Graph, Viper, ViperMonkey, YOMI, and probably VirusTotal, FileScan.IO. And quite a few other projects on GitHub. (Please contact me if you have or know a project using oletools)

Download and Install:

The recommended way to download and install/update the **latest stable release** of oletools is to use pip:

- On Linux/Mac: `sudo -H pip install -U oletools[full]`
- On Windows: `pip install -U oletools[full]`

This should automatically create command-line scripts to run each tool from any directory: `olevba`, `mraptor`, `rtfobj`, etc.

The keyword `[full]` means that all optional dependencies will be installed, such as XLMMacroDeobfuscator. If you prefer a lighter version without optional dependencies, just remove `[full]` from the command line.

To get the **latest development version** instead:

- On Linux/Mac: `sudo -H pip install -U https://github.com/decalage2/oletools/archive/master.zip`
- On Windows: `pip install -U https://github.com/decalage2/oletools/archive/master.zip`

See the documentation for other installation options.

Documentation:

The latest version of the documentation can be found online, otherwise a copy is provided in the doc subfolder of the package.

How to Suggest Improvements, Report Issues or Contribute:

This is a personal open-source project, developed on my spare time. Any contribution, suggestion, feedback or bug report is welcome.

To suggest improvements, report a bug or any issue, please use the issue reporting page, providing all the information and files to reproduce the problem.

You may also contact the author directly to provide feedback.

The code is available in a GitHub repository. You may use it to submit enhancements using forks and pull requests.

License

---

olevba contains modified source code from the officeparser project, published under the following MIT License (MIT):

officeparser is copyright (c) 2014 John William Davison

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.
