

KOST Koordinationsstelle für die dauerhafte Archivierung
elektronischer Unterlagen

Ein Gemeinschaftsunternehmen von Schweizer Archiven

Formaterkennung: FIDO

nestor Workshop: "Formaterkennung,
Formatvalidierung und Tools"
ZBW Kiel,
29.06.2017

Claire Röthlisberger-Jourdan

Inhaltsverzeichnis

- Was ist FIDO
- Eingabe
- Ausgabe
- Anwendung

Was ist FIDO



- FIDO (Format Identification for Digital Objects)
- Erstmals veröffentlicht: 2010
- Entwicklung: Open Preservation Foundation
- Lizenz: Apache 2.0 Open Source License
- Nutzungsmöglichkeiten: CMD
- Ausgabe: Konsole, CSV

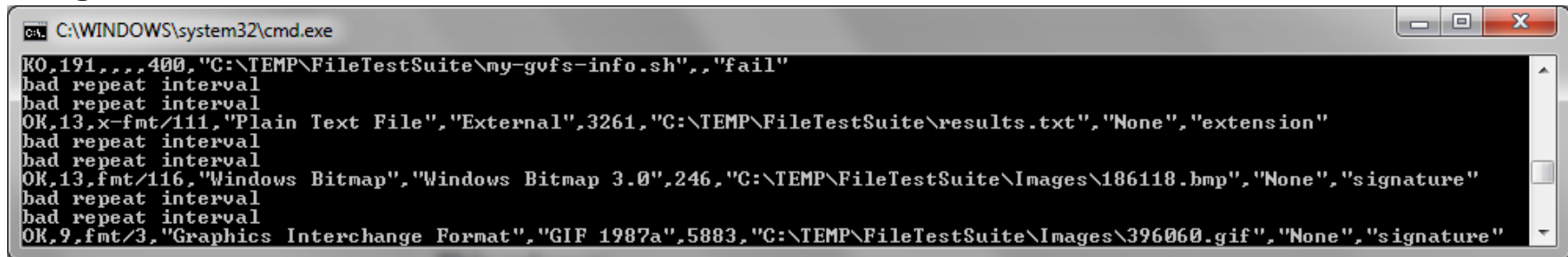
Eingabe & Ausgabe: CMD

fido --help (zeigt die Hilfe / alle Optionen)

Aufruf:

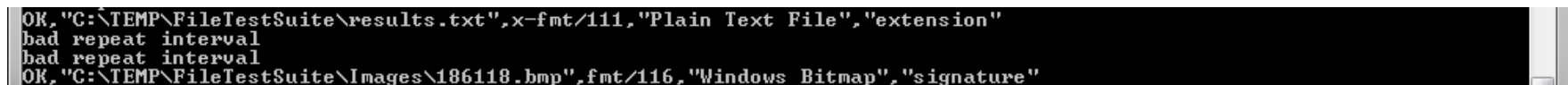
fido.exe -recurse C:\TEMP\FileTestSuite

Ausgabe: (OK, time, puid, formatname, signaturename, filesize, filename, mimetype, matchtype)



```
C:\WINDOWS\system32\cmd.exe
KO,191,,,,400,"C:\TEMP\FileTestSuite\my-gvfs-info.sh",,"fail"
bad repeat interval
bad repeat interval
OK,13,x-fmt/111,"Plain Text File","External",3261,"C:\TEMP\FileTestSuite\results.txt","None","extension"
bad repeat interval
bad repeat interval
OK,13,fmt/116,"Windows Bitmap","Windows Bitmap 3.0",246,"C:\TEMP\FileTestSuite\Images\186118.bmp","None","signature"
bad repeat interval
bad repeat interval
OK,9,fmt/3,"Graphics Interchange Format","GIF 1987a",5883,"C:\TEMP\FileTestSuite\Images\396060.gif","None","signature"
```

-matchprintf "OK,%(info.filename)s\","%(info.puid)s,
\"%(info.formatname)s\","\"%(info.matchtype)s\"\\n"



```
OK,"C:\TEMP\FileTestSuite\results.txt",x-fmt/111,"Plain Text File","extension"
bad repeat interval
bad repeat interval
OK,"C:\TEMP\FileTestSuite\Images\186118.bmp",fmt/116,"Windows Bitmap","signature"
```

> Log.txt (leitet die Konsole in die Log.txt.Datei um)

Anwendung

- Anwendungsmöglichkeiten
 - Übersicht aller Formate eines Ordners
 - Informationen zu einer «unbekannten» Datei
 - Erkennung gegenüber eines bestimmten Sets möglich
- Erfahrungswerte
 - Einbindung in einen automatischen Prozess ist möglich
 - Ausgabe kann angepasst werden

Anwendung

- Stärken:
 - + Flexibel einsetzbar, auch in einem Prozess
 - + Entwickler ist eine Archiv-Community
 - + Open Source
- Schwächen:
 - Kein GUI
 - Python-Programm (Installation)
 - Sehr feine Granularität, bedingt durch Pronom
 - Keine Library

KOST Koordinationsstelle für die dauerhafte Archivierung
elektronischer Unterlagen

Ein Gemeinschaftsunternehmen von Schweizer Archiven

Claire Röthlisberger-Jourdan

Koordinationsstelle für die Archivierung elektronischer Unterlagen

KOST

c/o Schweizerisches Bundesarchiv

Archivstrasse 24

CH-3003 Bern

T +41 79 464 07 41

E claire.roethlisberger@kost.admin.ch

W www.kost-ceco.ch