

Eigene Formatsignaturen für DROID und Zuarbeit zu PRONOM

nestor-Workshop Formaterkennung
Kiel, 29.06.2017
Martin Hoppenheit



Die Basis von PRONOM und DROID

Interne Signaturen (magic numbers)

- **Bitmuster, die charakteristisch für ein Dateiformat sind, z.B. 49492A00 am Anfang einer jeden TIFF-Datei.**
- **PRONOM beschreibt solche Bitmuster.**
- **DROID sucht in Dateien nach bekannten Bitmustern.**

Das perfekte Bitmuster ...

1. ... kommt in *allen* Dateien des Dateiformats vor, bestenfalls immer an der selben Stelle. Einfacher mit kurzen Mustern.
2. ... kommt *nur* in Dateien des Dateiformats vor. Einfacher mit langen Mustern, eigentlich aber unrealistisch.

Woher nehmen?

Zwei Möglichkeiten:

1. Dateiformatspezifikation
2. Analyse von Beispieldateien

Bitmuster in einer Spezifikation finden

Every valid SQLite database file begins with the following 16 bytes (in hex): 53 51 4c 69 74 65 20 66 6f 72 6d 61 74 20 33 00.

This byte sequence corresponds to the UTF-8 string "SQLite format 3" including the nul terminator character at the end.

Formatspezifikation SQLite 3

Bitmuster in einer Spezifikation finden

Auch schön:


- **PDF 1.4**
- **XML 1.0**
- **TIFF 6.0**

Bitmuster in einer Dateisammlung finden

Suche nach gemeinsamen Mustern in einer großen Dateisammlung aus unterschiedlichen Quellen.

- **Manuell via Hex-Editor.**
- **Automatisiert mit Tools wie tridscan.**
- **Nicht auf Dinge wie Computername oder Autor reinfallen!**

Signature File erstellen

 The National Archives

Prototype


PRONOM: Signature Development Utility

Name:

Version: Extension:



PUID: Mimetype:

Signature:

Anchor: 

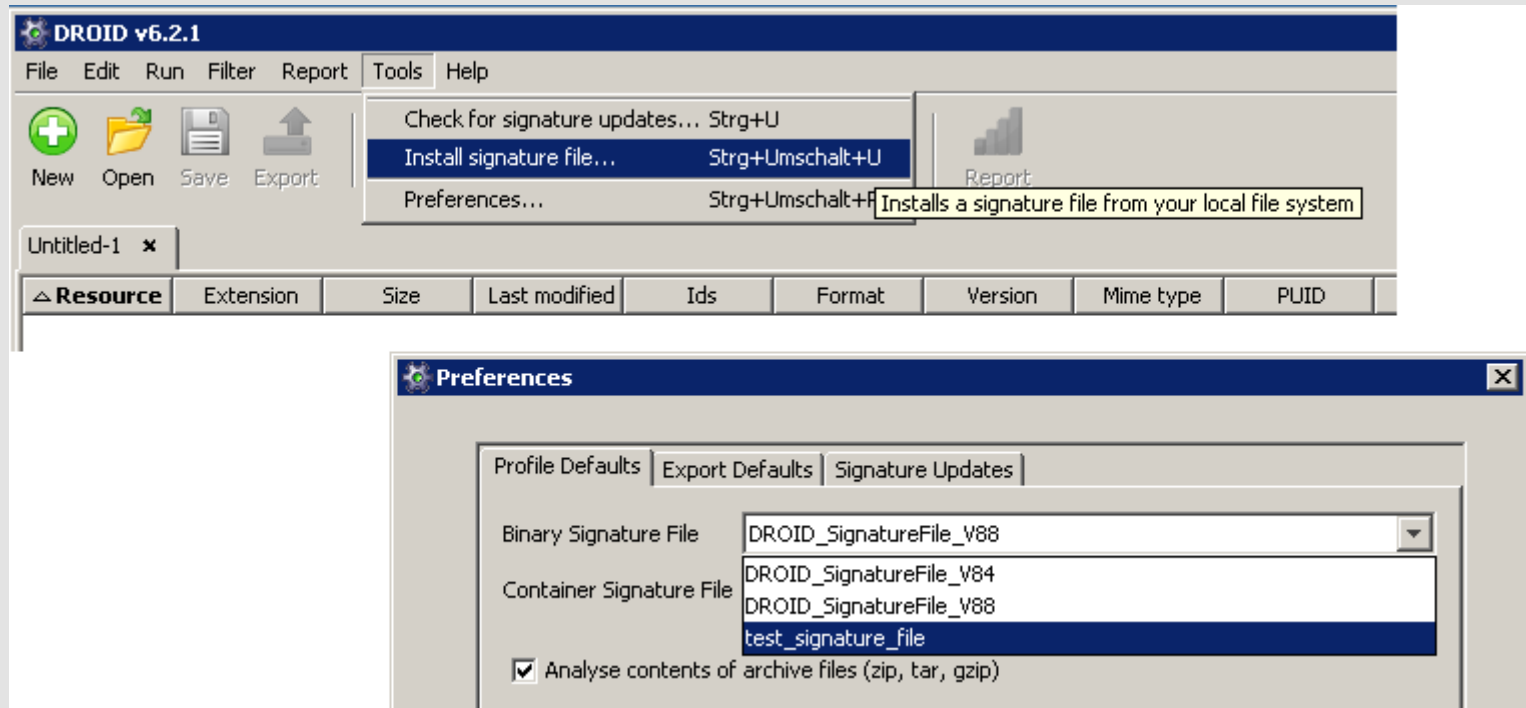
Offset:

Max Offset:

 Export to RDF  [Help](#)

<http://www.nationalarchives.gov.uk/pronom/sigdev/>

Mit DROID testen



Oder: `droid -R -Ns sigfile.xml -Nr testfiles`

Bei PRONOM einreichen

Your submission

Additional information

If you have specific technical information about the file format concerned, please provide it here.

PUID (optional)

References (optional)

To provide validity / authenticity to the entry

<https://www.nationalarchives.gov.uk/contact-us/submit-information-for-pronom/>

Weiterführende Links

Ausführliche Anleitung (deutsch) mit weiteren Links und Hintergrundinformationen:

<https://wiki.dnb.de/pages/viewpage.action?pageId=115213928>

Vielen Dank für Ihr Interesse!

martin.hoppenheit@lav.nrw.de

