



Notwendige technische Metadaten

Stefan E. Funk – SUB Göttingen
funk@sub.uni-goettingen.de

Selbstversorger oder All-Inclusive?

DIE ABLIEFERUNG ELEKTRONISCHER DOKUMENTE ORGANISIEREN

21. Oktober 2010, Frankfurt am Main

Technische Metadaten

- Beschreiben technische Eigenschaften digitaler Objekte
 - Dateiformat, Version des Formats, Dateigröße, Kompressionsart, Orientierung, Erstellungssoftware, ...
- Können mit Hilfsmitteln automatisch extrahiert werden
 - Die Interpretation eines Objektes beruht auf diesen Daten

Technische Metadaten? Wozu?

- Für die Dokumentation technischer Eigenschaften digitaler Objekte
- Für die Verwaltung digitaler Objekte in einem Archivsystem
- Für den Zugriff auf digitale Objekte anhand technischer Eigenschaften
- Für die Erhaltung digitaler Objekte
 - Möglichst bevor Gefahr besteht, dass digitale Objekte nicht mehr nutzbar sind
 - Zur späteren Auswahl von Migrations-Tools
 - Zur Bereitstellung von Emulations-Umgebungen
 - Für eine spätere Interpretation durch andere Tools



JHOVE –

JSTOR/Harvard Object Validation Environment

- Aufgaben von JHOVE bei der Untersuchung von digitalen Objekten
 - Identifikation: Um welchen Objekttyp handelt es sich?
 - Validierung: Entspricht das Objekt der Format-Spezifikation?
 - Charakterisierung: Extraktion technischer Metadaten
 - Bewertung im Sinne der LZA, in JHOVE2 mit DROID



JHOVE –

JSTOR/Harvard Object Validation Environment

- Das Ergebnis der Untersuchung hängt ab
 - Von der Qualität der digitalen Objekte: Wie genau entsprechen sie der Format-Spezifikation?
 - Von der Qualität der Validations- und Extraktions-Module
 - Wie genau wird das Objekt untersucht?
 - Welche technischen Metadaten können extrahiert werden?
 - Welche Formate werden unterstützt?



JHOVE –

JSTOR/Harvard Object Validation Environment

- Kurze Vorführung von JHOVE:

- Nicht valides TIFF Image

`Invalides_TIFF-Image_als_Bytestream_erkannt.xml`

- Nicht valides TIFF Image mit TIFF Modul

`Valides_TIFF-Image.xml`

- Valides TIFF Image (mit vielen Metadaten)

`Valides_TIFF-Image_mit_vielen_Metadaten.xml`

- Bitte umschalten!

Notwendige technische Metadaten

- Auswahl der technischen Metadaten
 - Können wir alle speichern?
 - Welche werden davon benötigt?
 - Für welche Zwecke?
 - Können wir ausschließen, dass wir später Metadaten benötigen, die wir heute als nicht notwendig erachten?
 - Spätere Nutzung der Objekte ist unter Umständen eine andere, als wir heute vorhersehen

Notwendige technische Metadaten

- Automatisierte Verarbeitbarkeit
 - Suche nach Objekten anhand technischer Kriterien
 - Auswahl von Metadaten für spätere Migration und/oder Emulation
 - Fehleridentifikation (für spätere Korrektur von fälschlich als valide erkannten Objekten)
 - Vergleich der Qualität der migrierten Objekte mit dem Original



Vielen Dank!

Gibt es Fragen?

Stefan E. Funk
SUB Göttingen
funk@sub.uni-goettingen.de

DP4lib



Referenzen

JHOVE

<http://hul.harvard.edu/jhove/>

JHOVE2

<https://confluence.ucop.edu/display/JHOVE2Info>