

Metadaten-Grundlagen

Datenformate

Jakob Voß

KIM Workshop, Mannheim, 2023-05-08

Was gibt es für Datenformat?

Welche (Meta)datenformate kennt ihr?

1. Frage: Welche Datenformate sind das oder könnten das sein?

- ▶ Wieder in Kleingruppen
- ▶ Mehrfachantworten sind möglich
- ▶ Bitte leserliche Karteikarten
- ▶ Umhängen ist erlaubt
- ▶ Bitte nicht einfach im Internet nachschlagen!

2. Frage: Welche relevanten Datenformate fehlen?

Beispiele von Datenformaten

1: Tab-Separated Values

- ▶ Tabellarisches Format
- ▶ Zeilen und Spalten
- ▶ Zeilen getrennt durch Zeilenumbruch
- ▶ Spalten getrennt durch Tabulator (Bytecode 09)

Vor- und Nachteile?

2: JavaScript Object Notation (JSON)

- ▶ Inspiriert von JavaScript
- ▶ Inzwischen wichtigstes allgemeines Strukturierungsformat
- ▶ Definiert unter <http://json.org>
- ▶ Elemente: Objekt, Array, String, Zahl, Boolean, Null
- ▶ Grundlage zahlreicher Datenformate und Kodierungen
 - ▶ JSKOS
 - ▶ BibJSON
 - ▶ PICA/JSON
 - ▶ ...

Einige JSON Werkzeuge

Webanwendungen

- ▶ <https://jsoneditoronline.org/>
- ▶ <https://jqplay.org/>

Lokal

- ▶ Texteditor mit Syntax-Highlighting
- ▶ jq (Kommandozeile)

Datenformate rund um JSON

- ▶ Programmiersprache: jq
- ▶ Abfragesprachen: JSONPath, SQL/JSON path, JSON Pointer...
- ▶ Schemasprachen: JSON Schema...
- ▶ Zahlreichen Kodierungen *in* JSON: JSON-LD, PICA/JSON...

3: PICA-Format

- ▶ Internes Datenformat der Bibliothekssysteme von OCLC PICA
- ▶ Feld-Unterfeld-Struktur (oder: Ebenen)
- ▶ Pica3/PICA+
- ▶ Verschiedene konkrete Anwendungsprofile
 - ▶ K10Plus Katalogisierungsformat
 - ▶ GND-Internformat
 - ▶ ...
- ▶ Beispiele: siehe OPAC oder WinIBW-Client

PICA-Datenmodell

- ▶ field, tag, occurrence
- ▶ subfield, code, value

und/oder

- ▶ level 0
- ▶ level 1
- ▶ level 2

Datenformate rund um PICA

Serialisierungen/Kodierungen

- ▶ PICA Plain
- ▶ PICA/XML
- ▶ PICA/JSON
- ▶ ...

Sprachen für PICA

- ▶ Abfragesprache *PICA Path*
- ▶ Änderungsformat *PICA Patch*
- ▶ Schemasprache *Avram*

Weitere konkreten Formate?