

Metadaten austauschen via XML

Metadaten austauschformate in Bibliotheken

Friedrich Summann
Universitätsbibliothek Bielefeld

Bibliothekarische Austauschformate

Übersicht

- CSV
- Textformate
- Proprietäre Formate

- Literaturverwaltungsformate (Endnote, RM, Bibtex)

- MAB2
- MARC
- DC/QDC
- METS/MODS

Use Cases für Metadatenaustausch

- Datenlieferung **von** aussen
- Datenlieferung **nach** aussen
- On-line-Schnittstellen
 - SRU
 - Z39.50
 - OAI-PMH
- Datenbank-Export
 - OPAC
 - JADE
 - PUB

Herausforderungen

- Analyse der vorliegenden Datenstruktur
- Modellierung der (Ziel)-Datenstruktur
- Zeichensatzbehandlung
- Verteilte Informationen
- Optimierung der Laufzeiten
- Reaktion auf Strukturänderungen

Austauschprotokolle

- FTP
- Z39.50
- SRU
- OAI-PMH

CSV-Format (Comma-separated Values)

Beispiel:

```
"Brecht, Bertolt" ; "" ; "Brecht, Bertolt : . - : , um 1932 " ;  
"Abweichend zu dem Druck in Brecht, St<FC>cke Bd 5, S. 116-117 (Gorki,  
Die Mutter), Frankfur  
t 1962 und einer anderen Version in Bd 3, Gedichte, Frankfurt 1961 unter  
dem Titel "Lob der Dialektik"." ;  
"http://www.ub.uni-bielefeld.de/cgi-bin/nav_autogr.c  
gi?autogr=brecht-1" ;
```

Textformate

Beispiel:

*** BRS DOCUMENT BOUNDARY ***

..VERF: Brecht, Bertolt

..TYPE: Eigenhändiges Gedicht-Manuskript auf einem China-Papier

..ORTE:

..DATE: um 1932

..UMFG: 10 Zeilen

..FORM: quer -8<97>

..PICS: netahtml/brechtframe.htm

..INFO: Abweichend zu dem Druck in Brecht, *Stücke* Bd 5, S. 116-117

(Gorki, *Die Mutter*), Frankfurt 1962 und einer anderen Version in Bd 3, *Gedichte*, Frankfur

t 1961 unter dem Titel "Lob der Dialektik".

..PLAZ: Rara-Magazin

..AKZE: 4/75511

MAB

MAB = Maschinelles Austauschformat für
Bibliotheken

Beispiel:

```
00590nM2.01200024          h001 BT000018466^^002a19971104^^003
20080725^^00420090111^^026 HBZBT000018466^^030 a|1uc|||||
17^^036aDE^^037bger^^050a|||||||||||||^
051 s|||||^070 61/NWBib^^070b361^^078n^_aNWBib^^080
60^^100R<C9>uhmann, Heinz^^102a118603949^^331 <88>Das<89>
war's^^335 Erinnerungen^^359 Heinz R<C9>uhmann^^403
Ungek<C9>urzte Ausg., 13. Aufl.^^410 Frankfurt/M. [u.a.]^^412
Ullstein^^425a1994^^433 311 S. : Ill.^^451 Ullstein-Buch ;
20521^^453 HT002139257^^455 20521^^456 20521^^540aISBN 3-548-
20521-6^^574 95,A01,1712^^700n|20^^700n|802080 ; R<C9>uhmann,
Heinz^^902p 4050884-5 R<C9>uhmann, Heinz^^902f |Autobio
graphie^^^]
```


Frühes „XML“ (BL)

```
<HEAD><ID>RN291432139</ID><SHM>9458.600000</SHM><TITLE>ZE  
ITSCHRIFT FUR ENTWICKLUNGSPSYCHOLOGIE UND PADAGOGISCHE  
PSYCHOLOGIE</TITLE><ISSN>0049-  
8637</ISSN><NL>231111</NL><CLASS1>L31</CLASS1><CLASS2>370  
.15;150</CLASS2><RR>GBP010.50</RR><CS>1</CS><ISSUE>2011;  
JAHR 43; HEFT 2</ISSUE><AUTH>Dopfner, M.; Kinnen, C.;  
Weber,  
K.N.; Pluck, J.</AUTH><TEXT>Behavioral problems of  
primary  
school-students: Results from the German Teacher's Report  
Form  
(TRF)</TEXT><PAGE>99-107</PAGE></HEAD  
>
```

BIBTEX

Beispiel:

```
@article{lin1973,  
  author = {Shen Lin and Brian W. Kernighan},  
  title  = {An Effective Heuristic Algorithm for the  
Travelling-Salesman Problem},  
  journal = {Operations Research},  
  volume = {21},  
  year   = {1973},  
  pages  = {498--516},  
}
```

Dublin Core

Felder

Title

Creator

Contributor

Description

Language

Type

Identifier

Relation

Source

Subject

Date

format

Provenance

Coverage

Rights

MABXML

```
<datei  
xsi:schemaLocation="http://www.ddb.de/professionell/mabxml/mabxml-1.xsd  
http://www.ddb.de/professionell/mabxml/mabxml-1.xsd">
```

```
-  
<datensatz id="ID946451494" typ="h" status="n" mabVersion="M2.0">
```

```
<feld nr="001" ind=" " >946451494</feld>
```

```
<feld nr="004" ind=" " >19980605</feld>
```

```
<feld nr="030" ind=" " >a|1iar|z|||37</feld>
```

```
<feld nr="036" ind="a">DE</feld>
```

```
<feld nr="050" ind=" " >a|a|||||||||</feld>
```

```
<feld nr="051" ind=" " >m|||z||</feld>
```

```
<feld nr="070" ind="a">292</feld>
```

```
<feld nr="100" ind="b">Ginsbach, Julia</feld>
```

```
<feld nr="102" ind="a">11547577X</feld>
```

```
<feld nr="104" ind="b">Liebers, Andrea</feld>
```

```
<feld nr="106" ind="a">111570204</feld>
```

```
-  
<feld nr="331" ind=" " >
```

```
<ns>Die</ns>
```

```
schÄ¶ne Lau
```

```
</feld>
```

JSON (JavaScript Object Notation)

Eigenschaften

- gültiges JavaScript
- Kompakt
- Lesbar
- Direkt umsetzbar in Datenstruktur
- Parser in allen Sprachen

Anwendungsbeispiel:

Bibliogr. Datensatz in JSON → CSL-Engine → Zitation (JSON)

```
'{"0001":{"URL":"http://pub.uni-bielefeld.de/pub?
func=drec&id=1706545","author":[{"family":"Pieper","given":"Dirk"},
{"family":"Wolf","given":"Sebastian"}],"container-title":"Information,
Wissenschaft & Praxis
(IWP)","id":"1706545","issue":"3","issued":{"date-parts":[["2007"]]},"page":"179
- 182","publisher":null,"publisher-place":null,"title":"BASE - Eine
Suchmaschine für OAI-Quellen und wissenschaftliche
Webseiten","type":"article","volume":"58"}}'
```

```
'["Pieper, D., & S. Wolf. 2007. \u201cBASE - Eine Suchmaschine  
f\u00fcr OAI-Quellen und wissenschaftliche Webseiten\u201d.  
<em>Information, Wissenschaft & Praxis (IWP)</em>, 58(3), 179 -  
182.http:\\pub.uni-bielefeld.de\\pub?func=drec&id=1706545"]'
```

```
use JSON ;
```

```
my $json = new JSON ;
```

```
$json = $json->canonical(1) ;
```

```
my $json_text = $json->encode($record_ref) ;
```

```
my $json = new JSON ;
```

```
my $citation_ref = $json->decode( $my_response->content) ;
```

XML kompakt

„Die Extensible Markup Language (engl. für „erweiterbare Auszeichnungssprache“), abgekürzt XML, ist eine Auszeichnungssprache zur Darstellung hierarchisch strukturierter Daten in Form von Textdaten.“ (Wikipedia)

- hoch-strukturiert
- lesbar
- Komplexe Strukturen abbildbar
- Bearbeitung zeitintensiv
- Processing wird umfänglich unterstützt

XML-Bestandteile

- * Elemente, deren physische Auszeichnung mittels
 - o einem passenden Paar aus Start-Tag (<Tag-Name>)
 - und End-Tag (</Tag-Name>) oder
 - o einem leerem Element-Tag (<Tag-Name />) erfolgen kann,
- * Attribute als bei einem Start-Tag oder leerem Tag geschriebene Schlüsselwort-Werte-Paare (Attribut-Name="Attribut-Wert") für Zusatz-Informationen über Elemente (eine Art Meta-Information),
- * Verarbeitungsanweisungen (<?Ziel-Name Parameter ?>, engl. Processing Instruction),
- * Kommentare (<!-- Kommentar-Text -->)
- * Text, der als normaler Text oder in Form eines CDATA-Abschnittes (<![CDATA[beliebiger Text]]>) auftreten kann.

XML-Processing

- Validieren
(well-formed, valid)
- Parsen
- Transformieren (XSLT)
- Erzeugen

OAI-PMH

Protokoll zum Metadata Harvesting

- OAI-DC Pflichtformat für Metadatenbereitstellung
- Alle anderen Formate sind optional
- Breite Mischung aus Formaten

OAI-PMH-Formatnutzung

METS

RFC1807

DIDL

XMETADISS/PLUS

QDC

MARC

ORE

EPICUR

OAI-DC

Beispielaufruf:

[http://eldorado.uni-dortmund.de:8080/dspace-oai/request?
verb=ListRecords&metadataPrefix=xmetadissplus](http://eldorado.uni-dortmund.de:8080/dspace-oai/request?verb=ListRecords&metadataPrefix=xmetadissplus)

The End