

Erschließungsdaten in einer heterogenen und verlinkten Welt

Überlegungen zur Datenqualität

Agenda

1. Bibliotheksdaten
2. Metadaten anderer Produzenten
3. Datenqualität und Metadatenmanagement
4. Provenance
5. Wo ist Qualität besonders wichtig?

Agenda

1. Bibliotheksdaten
2. Metadaten anderer Produzenten
3. Datenqualität und Metadatenmanagement
4. Provenance
5. Wo ist Qualität besonders wichtig?

Bibliotheksdaten

Traditionelle Stärken:

- **Basieren auf Regeln**
 - *klar definiert*
 - *langfristig gültig*
 - *überregional angewendet*
 - *transparent dokumentiert*
- **Hohe Verlässlichkeit**
 - *Erschließung durch geschultes Fachpersonal (i.d.R.)*
 - *zu berücksichtigende Datenelemente sind vorgegeben*
 - *genaue Anweisungen für die Erfassung der Daten (z.B. Anwendung von normiertem Vokabular)*
 - *es gibt Kontrollmechanismen (Plausibilitätskontrollen, kooperatives Arbeiten in Verbundumgebungen)*

Optimiert für Zusammenführen und Unterscheiden:

- **Ziel: Nicht nur etwas, sondern alles Passende finden**
Prinzip der „collocation“, z.B.:
 - *alles, woran eine bestimmte Person beteiligt war*
 - *alle Ausgaben eines Werks*
 - *alle Bände einer monografischen Reihe*
 - *alles zu einem Thema (jedoch verschiedene Systeme)*
 - **Auseinanderhalten von Dingen, die gleich aussehen**
Prinzip der Differenzierung, z.B.:
 - *unterschiedliche Manifestationen und Expressionen*
 - *Werke mit demselben Titel*
 - *gleichnamige Personen*
 - *Themen mit homonymen Bezeichnungen*
- ➔ *spezielle Mechanismen (Verknüpfungen, Normdaten, Individualisierung, identifizierende Merkmale)*

„Gehaltvolle“ Daten:

- **Aussagekräftig und informativ**
 - *detaillierte Beschreibung von Ressourcen*
 - *Anreicherungen (z.B. Inhaltsverzeichnisse)*
 - *viele Informationen bei Normdaten*
 - **Netzwerk von Beziehungen**
 - *Verknüpfungen zwischen Titeldatensätzen*
 - *Verknüpfungen zu Normdatensätzen*
 - *Verknüpfungen zwischen Normdatensätzen*
 - **Sinnvoll präsentiert**
 - *standardisierte Anordnung*
 - *Normalisierung (z.B. Groß-/Kleinschreibung)*
- ➔ *Fazit: großer und relativ homogener Bestand an bibliothekarischen Metadaten von hoher Qualität*

Risse in der „heilen Welt“


Einschränkungen bzw. Gefährdungen:

- **Einsparungen aus ökonomischen Gründen**
 - *teilweise Aufgabe von Serienverknüpfungen*
 - *Beschränkung bei der Erfassung auf das Standard-elemente-Set (starke Tendenz bei DNB)*

➔ *verringert Suchmöglichkeiten und Vernetzung*
- **Einsatz maschineller Methoden**

derzeit insbesondere bei DNB ein Thema

 - *Netzpublikationen: rein maschinelle Verbesserungen der Produzenten-Daten (z.B. GND-Verknüpfungen)*
 - *maschinelle Inhaltserschließung in unterschiedlichen Formen*

- **Regelwerksentwicklung und -anwendung**
 - *Bruch durch den Wechsel von RAK zu RDA*
 - *„cataloger’s judgment“: größere Uneinheitlichkeit (abgemildert durch D-A-CH und Standardelemente)*
 - *teilweise fehlende Sicherheit der Katalogisierenden*
 - *dynamische Weiterentwicklung von RDA*
 - *strategische Entscheidungen werden kurzfristiger (z.B. DDC-Anwendung bei DNB)*
 - **Verstärkte Mischung mit anderen Metadaten**
 - *Verlagsdaten als Ausgangsbasis verwendet*
 - *Einspielung von Verlagsdaten für E-Books*
 - *in RDS überwiegen nicht-bibliothekarische Daten*
-  *Heterogenität gab es schon immer, sie hat aber eine neue Dimension erreicht und wird noch weiter zunehmen*

Agenda

1. Bibliotheksdaten
2. Metadaten anderer Produzenten
3. Datenqualität und Metadatenmanagement
4. Provenance
5. Wo ist Qualität besonders wichtig?

Typische Effekte in RDS (hier Summon, Konstanz)

Verwendung unterschiedlicher Formbegriffe für Rezensionen

KonSearch

essential classification broughton

Buchkapitel (157)

Zeitu... (1487)

meh...

FACH

Medi...

Gesc...

Biolo...

Spra...

Psyc...

mehr...

ERSCHEINUNGSJAHR

von bis

letzte 12 Monate

letzte 3 Jahre

letzte 5 Jahre

THEMA

history (1.418)

electronic books (760)

aufsatzsammlung (558)

multi-user (438)

united states (422)

mehr...

SPRACHE

nach oben

Library Management, 08/2005, Band 26, Ausga

Zeitschriftenartikel: **Volltext online**

Details ▾

3 online

Essential Classification20061Vanda **Broughton**. **Essential Classification**. London: Facet Publishing 2004. 324 pp. £24.95 Paperback, ISBN: 1-85604-514-5 von [Macgregor, George](#)

Library Review, 01/2006, Band 55, Ausgabe 1

Zeitschriftenartikel: **Volltext online**

Details ▾

4 online

Essential Classification20051Vanda **Broughton**. **Essential Classification**. London: Facet Publishing 2004. x+324 pp., ISBN: 1-85604-514-5 £24.95 von [Trickey, Keith V](#)

New Library World, 05/2005, Band 106, Ausgabe 5/6

Zeitschriftenartikel: **Volltext online**

Details ▾

5 online

Broughton, Vanda. **Essential Classification** von [Ketcham, Susan E](#)

Library Journal, 02/2005

...**Broughton, Vanda**. **Essential Classification**. Neal-Schuman. 2004. 313p. illus. bibliog. index. ISBN 1-55570-507-3. pap. \$55. If you are...

Buchbesprechung: **Volltext online**

Details ▾

6 online

Essential Classification: by Vanda **Broughton**. 2nd ed. London: Facet Publishing, 2015. x, 421 p. ISBN 978-1-78330-031-0. \$75.00 von [Bothmann, Bobby](#)

Cataloging & Classification Quarterly, 11/2016

Buchbesprechung: **Volltext online**

Details ▾

Materialart ▾ anwenden (0) ✕

Liste filtern

- Buch / eBook (39.725)
- Buchbesprechung (324)
- Buchkapitel (950) ✓ ✕
- Fachzeitschriftenartikel (13) ✓ ✕
- Hochschulschrift (6) ✓ ✕
- Konferenzbericht (19) ✓ ✕
- Magazinartikel (82) ✓ ✕
- Publikation (4) ✓ ✕
- Referenz (11) ✓ ✕
- Schriftstück (2) ✓ ✕
- Zeitschriftenartikel (4.110) ✓ ✕
- Zeitungsartikel (212) ✓ ✕

*Formfacette beim selben Beispiel:
unterschiedliche Bezeichnungen für
dieselbe Sache, teils unklare Begriffe*

THEMA

- history (1.418)
- electronic books (760)
- aufsatzsammlung (558)
- multi-user (438)
- united states (422)
- mehr...

*Formbegriffe in der
Themenfacette*

Beispiel für eine nicht wirklich hilfreiche Facette im Angesicht von 243.030 Treffern...

Ergebnisse verbessern

Auch Treffer ohne Volltextzugriff anzeigen

Sortieren nach Relevanz ▾

Nur zeigen: ▾

Typ der Ressource ▾

Erscheinungsjahr ▾

Sprache ▾

Bibliothek ▾

Neuerwerbungen ▾

Zeitschrift / Serie ▾

Schlagwort ▲

- Introduction (32.880)
- Economics (28.809)
- Governments (25.336)
- Engineering (22.164)
- Computer Science (20.199)

Mehr anzeigen

Fachliche Einschränkung ▲

- Sprachen, Literatur (1)
- Medizin, Technik, Naturwissenschaft (1)
- Informatik (1)

SEITE 1 243.030 Ergebnisse

- ARTIKEL**
The archaeological services
Heinrich, Helen
The Bottom Line
[Volltext verfügbar >](#)
- ARTIKEL**
Vermont: Library and information technology
Klinck, Patricia E
Library Hi Tech, 01 February 1996, Vol.14(2/3), pp.309-316 [Peer Reviewed Journal]
[Volltext verfügbar >](#)
- ARTIKEL**
Research needs for meeting the challenge of decentralized energy supply in developing countries
Sulistyono, Medina, Kabin, Ningsih, Nurwaning, Kristanto
[Peer Reviewed Journal]
res **technical** upgrades and rapid expansion...
[Volltext verfügbar >](#)
- ARTIKEL**
Co-production of ICT and children at risk: The introduction of the Child Index in Dutch child welfare
Lecluijze, Inge ; Penders, Bart ; Feron, Frans J.M. ; Horstman, Klasien
Children and Youth Services Review, September 2015, Vol.56, pp.161-168 [Peer Reviewed Journal]
“ introduction of the Child Index in practice and analysing it as a socio-technical... introduction of the Child Index a socio-technical trajectory: a complex ”
[Volltext verfügbar >](#)

Fachliche Einschränkung ▲

- Sprachen, Literatur (1)
- Medizin, Technik, Naturwissenschaft (1)
- Informatik (1)

Typische Ursachen

(nach Zeng/Qin 2016, S. 321)

1. **„There is no such an element.“**
nicht im Metadatenschema vorhanden
2. **„It is not a required element.“**
vorhanden, aber nicht verpflichtend
3. **„Whoops, wrong place...“**
beim falschen Element erfasst
4. **„I did not use a controlled vocabulary.“**
keine Konsistenz, entsprechend schwer zu mappen
5. **„I did not follow any rule.“**
keine Erfassungsregeln, z.B. zur Syntax
6. **„Somehow it got lost.“**
Fehler bei Transport oder Mapping führt zu Verlust


Verlagsdaten für E-Books

Definition Anforderungen:

<http://www.dnb.de/SharedDocs/Downloads/DE/DNB/wir/anforderungenMetadatenEbooksMarcOnixRDA2015.html>

- **Allgemeine Anforderungen, u.a.**
 - *Standardformat (MARC oder ONIX), UTF-8*
 - *Regelungen für Updates und Löschungen*
 - *Bearbeitung/Korrektur und Weitergabe ist erlaubt*
 - *Datenkernset mit Detailregelungen*

Typische Probleme:

- **Unsauberes Format, Zeichensatzfehler**
z.B. Fehler im Header, falsche Felder, HTML-Tags
- **Erfassungsfehler**
z.B. „Vorname Nachname“ oder beim Umfang nur Seiten statt „1 Online-Ressource (278 Seiten)“
 *Konverter muss pro Anbieter angepasst werden*

=LDR 03065nam 2200277 45e0
=001 9783428493432
=005 20171222085800.0
=007 cr
=008 171222s1998\000\0\ger\d
=020 \\\\$a9783428493432 (electronic)
=020 \\\\$a9783428093434 (print)
=024 3\\\$a9783428493432
=024 7\\\$a10.3790/978-3-428-49343-2\$2doi
=041 0\\\$ager
=100 1\\\$aSchröck, Jörg
=245 10\\\$aSteuersystem und Formwechsel.\$bEine Systemanalyse des ertragsteuerneutralen Wechsels zwischen Kapital- und Personenhandelsgesellschaftsformen auf der Grundlage des **Äbertragungsmodells nach §§ 14, 25 UmwStG.**\$cJörg Schröck\$h[Elektronische Ressource].
=250 \\\\$a1. Aufl.
=260 3\\\$aBerlin :\$bDuncker & Humblot\$c1998
=300 \\\\$aOnline-Ressource **(256 S.)** :
=490 0\\\$aSchriften zum Steuerrecht\$v60
=520 \\\\$aDas mit 1. Januar 1995 in Kraft getretene Umwandlungsbereinigungsgesetz ist wirtschaftspolitisch interessant, weil es Gesellschaften Umwandlungserleichterungen in noch weiterem Umfang als bisher **ermöglicht.** Aber ist es dem Gesetzgeber auch gelungen, die Umwandlungserleichterungen in rechtssystematisch stimmiger Weise zuzulassen? Wurden rechtssystematische Schranken bei der Verfolgung der wirtschaftspolitischen Ziele **berücksichtigt?** Immer wieder führten bisher Mißachtungen des Systemgedankens im Steuerrecht zu einem Steuer-Chaos, das allseits beklagt wird. -- Nach dem neuen UmwG ist es nun möglich, Kapital- in Personengesellschaften und umgekehrt durch bloße Rechtsformänderung - d.h. ohne Äbertragungsakt - umzuwandeln (sog. Formwechsel). Damit diese Umwandlungsvarianten nicht von negativen Steuerfolgen begleitet sind, sollen die **§§ 25 S. 1 und 14 S. 1** UmwStG diesen Formwechsel steuerneutral ermöglichen, sondern als einen Äbertragungsvorgang. Der Verfasser nach Handelsrecht. Dabei will er klären, ob es für die und Darlegung der systematischen Grundlagen hat nicht Gesetzgeber geschaffene UmwStG ist gleichzeitig Präzisierung in induktiv gewonnenen, allgemeinen Ertragsteuerprinzipien kann die gesetzgeberische Konzeption sich entweder als erweisen.
=689 \\\\$aErtragsteuer
=689 \\\\$aGesellschaftsumwandlung /Steuer
=689 \\\\$aHandelsgesellschaft /Steuer
=776 1\\\$z9783428493432

Beispiel für „Verlags-MARC“ (noch in RAK) mit Zeichensatzfehlern; 689 falsch verwendet (Feld für Schlagwortfolgen); Fehler bei Deskriptionszeichen

Mit manchem kann man leben, mit anderem nicht. Typische Verbesserungen beim Einspielen außerdem: Ergänzen von IMD-Typen, teilweise auch Beziehungskennzeichnungen

0500 Oan
 0501 Text\$btxt
 0502 Computermedien\$bc
 0503 Online-Ressource\$bcr
 0601 ZDB-2-LCM\$b2018
 1100 2018\$n2018
 1130 cofz
 1140 text
 1500 eng
 1505 \$erda
 1700 XA-GB
 2000 978-1-137-40658-3
 2003 978-1-137-40657-6\$\$Sp
 2051 10.1057/978-1-137-40658-3
 2113 Springer: 978-1-137-40658-3
 3000 Shingler, Martin
 4000 When Warners Brought Broadway to Hollywood, 1923-1939\$hby Martin Shingler
 4030 London\$nPalgrave Macmillan
 4060 Online-Ressource (X, 237 p. 21 illus. in color, online resource)
 4085 \$uhttp://dx.doi.org/10.1057/978-1-137-40658-3\$xH
 4190 Palgrave Studies in Screen Industries and Performance
 4190 SpringerLink : Bücher
 4243 Erscheint auch als\$nDruck-Ausgabe\$7978-1-137-40657-6
 5010 791.430973
 5020 APF\$2bicssc
 5020 PER004000\$2bisacsh
 5030 PN1993.5.U6
 5500 |s|Culture; Study and teaching
 5500 |s|United States; Study and teaching
 5500 |s|Motion pictures; United States
 5500 |s|Motion pictures; History
 5500 |s|Motion picture acting
 5500 |s|Theater
 5500 |s|Cultural and Media Studies

Beispiel für Verlagsdatensatz nach Einspielung in SWB: Verfasser ohne Verknüpfung und Beziehungskennzeichnung, Verwirrung in 4060 (Umfang), Groß-/Kleinschreibung nicht normalisiert

Agenda

1. Bibliotheksdaten
2. Metadaten anderer Produzenten
3. Datenqualität und Metadatenmanagement
4. Provenance
5. Wo ist Qualität besonders wichtig?

Datenqualität
=
Qualität der Ausgangsdaten
+
Metadatenmanagement

Beurteilung/Messung von Qualität

Kriterien/Indikatoren nach Bruce/Hillmann 2004:

- **„Completeness“**
 - ausführliche Beschreibung (jedoch noch wirtschaftlich)
 - umfassende Anwendung der Elemente auf alle Objekte
- **„Accuracy“**
 - sachlich korrekte Beschreibung der Objekte
 - „high-quality editing“ (z.B. Tippfehler, korrekte Namensformen)
- **„Provenance“ (Herkunft)**
 - von wem und mit welchen Methoden wurden die Daten erstellt?
 - welche Transformationen wurden seit der Erstellung vorgenommen?

- **„Conformance to expectations“**
 - *von der Community erwartete Elemente sind vorhanden*
 - *normiertes Vokabular auf Nutzer zugeschnitten*
 - *nötige Kompromisse sind durchdacht und dokumentiert*
- **„Logical consistency and coherence“**
 - *vor allem für Recherche nötig*
- **„Timeliness“ (Aktualität)**
 - *„currency“: z.B. keine veralteten URLs, aktueller Stand der Erschließungsstandards*
 - *„lag“: Objekt steht früher zur Verfügung als Metadaten*
- **„Accessibility“ (Zugänglichkeit)**
 - *fehlende Verlinkung zwischen Metadaten und Objekt*
 - *technische Barrieren (z.B. proprietäres Datenformat)*
 - *wirtschaftliche Barrieren (kostenpflichtig)*
 - *intellektuelle Barrieren (z.B. zu kompliziert)*

Kriterien/Indikatoren nach Zeng/Qin 2016:

- **„Completeness“**
Menge und Verteilung von Datenelementen in einzelnen Datensätzen und gesamter Datenbank
- **„Correctness“**
zutreffende Beschreibung, richtige Felder, korrekte Syntax, korrekte Eingabe (z.B. Tippfehler, Leerzeichen), korrekte Darstellung im Browser (z.B. Sonderzeichen), korrektes Mapping bei Datenintegration
- **„Consistency“**
u.a. Konsistenz bei Erfassung der Daten, gleichmäßige Angabe von URLs (was genau wird referenziert?)
- **„Duplication“**
Dublettenkontrolle innerhalb einer Sammlung oder bei Datenintegration

Metadatenmanagement

In umfassendem Sinn zu verstehen:

- **Prüfung und Auswahl von Metadaten**
 - *Gibt es überhaupt Metadaten?*
 - *Sind sie grundsätzlich verwertbar?*
 - *Gibt es evtl. mehrere Anbieter?*
- **Detailanalyse und Monitoring**
 - *Welche Probleme und Fehler gibt es?*
 - *Regelmäßige Kontrolle auf mögliche Veränderungen*
- **Kontakt zu Datenproduzenten**
 - *Hinweisen auf Fehler, Einfordern von Standards*
 - *Ziel: Daten an der Quelle verbessern*
- **Bewerten und Planen der Verfahren**
 - *Welche Datenqualität wird benötigt?*
 - *Wahl der Methoden (maschinell und/oder intellektuell)*

- **Vorgaben für Mapping**
 - *nicht nur Felder, sondern auch Feldinhalte (z.B. Mapping von normiertem Vokabular)*
 - **Maschinelle Verbesserungen, z.B.**
 - *regelbasierte Anreicherungen*
 - *Verlinkung mit Normdaten*
- ➔ *zwingende Voraussetzung für derartige Eingriffe:
Die Daten müssen in der eigenen Hand sein!*

Kommerzielle Indizes und Produkte

Auf Seite der kommerziellen Anbieter:

- **Fehlendes Verständnis und Know-how**
Die Profis in den relevanten Bereichen sind wir!
- **Fehlendes Interesse an Datenverbesserung**
Investitionen in Metadatenmanagement werden nur bei absehbaren wirtschaftlichen Vorteilen getätigt
- **Persönlicher Eindruck: kaum Verbesserungen**
z.B. bei Summon stößt man immer noch auf dieselben banalen Fehler wie schon vor Jahren

➔ *nötig ist eine Emanzipation der Bibliotheken!*

Emanzipation ist möglich!

Zwei Beispiele:

- **finc-Artikelindex der UB Leipzig**
 - *freie und lizenzierte Datenquellen*
(2016: über 70 Quellen, über 50 Mio. Aufsätze)
 - *finanziert über Umlage (derzeit ca. 20 Bibliotheken)*
- **K10plus Zentral (ehemals GBV Zentral)**
 - *freie und lizenzierte Datenquellen, alle Medienarten*
(aktuell über 196 Mio. Datensätze)
 - *Kombination mit verschiedenen Oberflächen möglich*
 - *derzeit ca. 75 Bibliotheken*
 - *Methode „Mitmachen“ (jede Bibliothek kann neue Datenquellen melden)*
 - *bundesweit kostenfrei nutzbar*

Tools für Metadatenmanagement:

- **In Systeme integrierte Werkzeuge**
z.B. Pica: File Format Converter (FCV)
- **Freie bibliothekarische Eigenentwicklungen, z.B.**
 - *Metafactory („A versatile tool suite for metadata processing“, entstanden im Culturegraph-Kontext)*
 - *d:swarm („eine skalierbare und nachnutzbare, intelligente Datenmanagement-Plattform“, SLUB Dresden)*
- **Freie Software im Bereich Data Science, z.B.**
KNIME Analytics Platform (grafische Benutzeroberfläche)

Metadatenmanagement in der Ausbildung:

- **Beispiel HdM Stuttgart**
derzeit als Wahlangebot, ab WS 2018/19 Pflicht im neuen Schwerpunkt „Daten- und Informationsmanagement“

Agenda

1. Bibliotheksdaten
2. Metadaten anderer Produzenten
3. Datenqualität und Metadatenmanagement
4. Provenance
5. Wo ist Qualität besonders wichtig?

Mischdatensätze

Angaben unterschiedlicher Provenance im selben Datensatz sind längst Realität:

- **Verschiedene Kombinationen von**
 - *nicht-bibliothekarischen Daten (z.B. Verlagsdaten)*
 - *intellektuellen bibliothekarische Erschließung*
 - *maschinellen Veränderungen/Anreicherungen*
- **Kennzeichnung der Provenance**
 - *nur teilweise explizit in den Daten sichtbar*
 - *u.U. nur im eigenen System erkennbar, kann beim Transport verloren gehen*

 *im Idealfall müsste jede Information eine Provenance-Information besitzen*

0500 Aau

0501 Text\$btxt

0502 ohne Hilfsmittel zu benutzen\$bn

0503 Band\$bnc

0575 imda

1100 2003\$n2003

1130 druck

1131 !208917802!Festschrift

1140 fe

1500 eng\$sengl.

2113 OCLC: 53090269

3010 !186905939!Roberto, Dominique\$BHrsg.

3040 !186905920!Ugo, Renato*1938-*

4000 In honour of Prof. Renato Ugo on the occasion of his 65th birthday\$hguest eds.:
Dominique Roberto ...

4030 Amsterdam [u.a.]\$nElsevier

4060 831 S.

4061 Ill., graph. Darst.

4170 Journal of molecular catalysis : A, Chemical ; 204/205 : Special issue

4180 !045296774!Journal of molecular catalysis / A\$I204/205 : Special issue

4201 Festschrift Renato Ugo

4700 IMD-Felder und 1131 maschinell ergänzt

*Maschinelle Ergänzungen von
RDA-Elementen an RAK-Auf-
nahmen im SWB: menschen-
lesbar gekennzeichnet*

3000 !13267775X!*Fingerle, Birgit*\$B\$VerfasserIn\$4aut

3010 !16854069X!*Mumenthaler, Rudolf**1962-*\$B\$VerfasserIn\$4aut

4000 Innovationsmanagement in Bibliotheken\$h\$Birgit Inken Fingerle, Rudolf Mumenthaler

4030 Berlin\$n\$De Gruyter Saur

4060 1 Online-Ressource (VI, 175 Seiten)

4085 \$uhttp://www.degruyter.com/viewbooktoc/product/2

4085 \$uhttp://dx.doi.org/10.1515/9783110338850\$xH

4085 \$uhttps://doi.org/10.1515/9783110338850\$xH

4190 Praxiswissen

4207 Innovation is essential for libraries to keep up with c

and ensure their future existence. The authors draw on their extensive experience in directing academic libraries throughout the German-speaking world. Building upon a rigorous theoretical foundation, they develop practical guidelines for introducing innovation management practices to libraries.

4207 Innovation ist für Bibliotheken ein Muss, um mit den aktuellen technischen Entwicklungen mithalten zu können und die Zukunft zu sichern. Die Autoren greifen auf einen reichhaltigen Erfahrungsschatz in führenden wissenschaftlichen Bibli

4243 Erscheint auch als\$n\$Druck-Ausgabe!**393787362!***Innovationsmanagement in Bibliotheken / Fingerle, Birgit*

4700 ZRED: Sacherschließungsdaten maschinell aus paralleler Ausg. übernommen

5010 025.1

5010 020

5010 650

5090 !200653733!*AN 70300 [Allgemeines # Buch- und Bibliothekswesen, Informationswissenschaft # Bibliothekswesen # Bibliotheksverwaltung # Bibliotheksorganisation und -betrieb, Rationalisierung, Zentrale Dienstleistungen]*

5100 !208865586!*Bibliothek*

5101 !208969136!*Innovation*

5102 !209868775!*Innovationsmanagement*

*Maschinelle Übernahme
von SE aus gedruckter
Parallelausgabe: dito*

Titel	Der Legionär als Leistungssportler : die Leistung römischer Soldaten auf dem Prüfstand / herausgegeben von Carl Pause ; Clemens-Sels-Museum Neuss
Person(en)	Pause, Carl (Herausgeber) Mersch, Norbert (Mitwirkender) Severin, Sascha (Mitwirkender) Heinen, Andreas (Mitwirkender)
Verlag	Neuss : Stadt Neuss
Zeitliche Einordnung	Erscheinungsdatum: 2014
Umfang/Format	47 Seiten : Illustrationen ; 30 cm
ISBN/Einband/Preis	978-3-936542-68-4 Broschur : EUR 5.50 (DE), EUR 5.50 (AT), ca. sfr 4.50 (freier Pr.)
EAN	9783936542684
Sprache(n)	Deutsch (ger)
Schlagwörter	Leistung <Recht>* ; Soldat* ; Leistungssportler* ; Lastganglinie* ; Prüfstand* ; Mersch* (*maschinell ermittelt)
Sachgruppe(n)	796 Sport ; 930 Alte Geschichte, Archäologie ; 610 Medizin, Gesundheit
Weiterführende Informationen	Inhaltsverzeichnis

Carl Pause	Arma et impedimenta: Die Ausrüstung der römischen Soldaten	5
Norbert Mersch	Überlegungen zur Rekonstruktion einer Militärtunika nach römischen Soldatengrabsteinen des 1. und 2. Jahrhunderts nach Christus	19
Andreas Heinen und Sascha Severin	Belastungsprofil eines Legionärs: Eine sportwissenschaftliche Annäherung	27
Norbert Mersch	Märsche als Selbsterfahrung	33

2100 14,N15
 2105 17,B48
 2150 15685ef2741848e4b66578db2b6a40fa
 2240 DNB:1049547829
 2242 |o|876382932
 3010 !113788347!Pause, Carl [Tp1]\$BHerausgeber\$4edt
 3019 |m|!1145225225!Mersch, Norbert [Tn7]\$BMitwirkender\$4ctb
 3019 |m|!1145225233!Severin, Sascha [Tn7]\$BMitwirkender\$4ctb
 3019 |m|!1011537516!Heinen, Andreas [Tn3]\$BMitwirkender\$4ctb
 4000 Der @Legionär als Leistungssportler : die Leistung römischer Soldaten auf dem Prüfstand /
 herausgegeben von Carl Pause ; Clemens-Sels-Museum Neuss
 4030 Neuss : Stadt Neuss ***94825
 4060 47 Seiten
 4061 Illustrationen
 4062 30 cm
 4700 |ERW-L|Le
 4700 |FE-L|fi
 4715 \$">=u \$=c 04=d DNB=e 1
 5050 796;930;610\$Ei\$Hie-sg\$D2017-11-23
 5050 930\$Em\$Haep-sg\$K0,98199\$D2017-11-22
 5050 930\$Ea\$Honx\$D2014-04-03
 5051 \$KK_A8_03_20160930_de\$LS_TOC3_WB42_20170726_de
 5052 \$f340\$F0,90279\$g620\$G0,78891\$D2017-11-22
 5540 [GND]!041672666!Leistung\$gRecht [Ts1]\$Em\$Haep-gnd\$K0,00139\$D2017-11-22
 5540 [GND]!040554090!Soldat [Ts1]\$Em\$Haep-gnd\$K0,00125\$D2017-11-22
 5540 [GND]!040741338!Leistungssportler [Ts1]\$Em\$Haep-gnd\$K0,00039\$D2017-11-22
 5540 [GND]!94001081X!Lastganglinie [Ts1]\$Em\$Haep-gnd\$K0,00036\$D2017-11-22
 5540 [GND]!042090377!Prüfstand [Ts1]\$Em\$Haep-gnd\$K0,00024\$D2017-11-22
 5540 [GND]!041008502!Mersch [Tg1]\$Em\$Haep-gnd\$K0,00013\$D2017-11-22
 5560 [Produktform]Book
 5560 [VLB-WN]1553: Hardcover, Softcover / Geschichte/Altertum

*Elaborierte, für Maschinen auswertbare
Kennzeichnung von maschinell erstellten
Schlagwörtern bei DNB*

3010 !113788347!Pause, Carl [Tp1]\$BHerausgeber\$4edt
 3019 |m|!1145225225!Mersch, Norbert [Tn7]\$BMitwirkender\$4ctb
 3019 |m|!1145225233!Severin, Sascha [Tn7]\$BMitwirkender\$4ctb
 3019 |m|!1011537516!Heinen, Andreas [Tn3]\$BMitwirkender\$4ctb

bei DNB

- **Spezialfeld 3019 bei DNB**

Hier bleiben Personen aus Verlagsdaten stehen, die nicht in korrekte RDA-Beziehungen umgesetzt werden; erhalten pauschal die BZK „Mitwirkender“ (ist in diesem Fall sachlich falsch; es sind Verfasser von Aufsätzen)

3010 Pause, Carl\$7113788347\$BHerausgeberIn\$4edt
 3010 Mersch, Norbert\$71145225225\$BMitwirkendeR\$4ctb
 3010 Severin, Sascha\$71145225233\$BMitwirkendeR\$4ctb
 3010 Heinen, Andreas\$71011537516\$BMitwirkendeR\$4ctb

im Fremddaten-
pool des SWB

- **Im SWB nicht mehr zu erkennen**

wird womöglich von KatalogisiererInnen so belassen, weil es ja „von DNB“ kommt und deshalb richtig sein sollte

0500 Oau
 0501 Text\$btxt
 0502 Computermedien
 0503 Online-Ressourcen
 0601 ZDB-41-UTB
 0601 ZDB-41-UTBM
 0601 ZDB-41-UTBMS
 1100 2018\$n[2018]
 1130 cofz
 1140 text
 1500 ger
 1505 \$erda
 1700 XA-DE-BW
 2000 978-3-8252-4840
 2113 EBP: 014401460
 2113 UTB: 20004391
 2113 OCLC: 1015670524
 2240 BSZ: 496233890
 3010 !352278056!Gerick, Julia*1986-*\$B\$HerausgeberIn\$4edt
 3010 !392428733!Sommer, Angela\$B\$HerausgeberIn\$4edt
 3010 !420002669!Zimmermann, Germa*1983-*\$B\$HerausgeberIn\$4edt
 4000 Kompetent Prüfungen gestalten\$53 Prüfungsformate für die Hochschullehre\$hJulia Gerick, Angela Sommer, Germa Zimmermann (Hrsg.)
 4030 Münster ; New York\$nWaxmann
 4060 1 Online-Ressource (267 Seiten)
 4085 \$mX:UTB\$uhttp://www.utb-studi-e-book.de/9783838548401\$xH
 4170 UTB ; 4840
 4180 #4840#!373318286!UTB\$I4840
 4207 Wichtige und innovative Prüfungsformate im ÜberblickDas Buch stellt 53 unterschiedliche, neue Prüfungsformen vor. Dazu werden die Formen in ihrer Anlage und Funktion aufgezeigt und ihre Verwendung in der Hochschullehre erörtert. Der Blick „über den Tellerrand“ klassischer Prüfungsformate hin zu Formen wie Pecha Kucha, Blogbeitrag oder 24-Stunden-Hausarbeit helfen, das eigene Prüfungsportfolio zu erweitern.Die Methoden sind systematisch und übersichtlich dargestellt und lassen sich auf eigene Lehr- und Lernkontexte in der Praxis übertragen.
 4243 Erscheint auch als\$nDruck-Ausgabe!489593542!Kompetent Prüfungen gestalten
 5010 378.1661

Datensatz für ein E-Book im SWB mit drei „Schichten“ (ohne jede Dokumentation):

1. ursprüngliche Verlagsdaten
2. maschinelle Verbesserungen beim Laden (u.a. Verknüpfung von zwei Personen)
3. intellektuelle Verbesserungen (u.a. Datum und Verknüpfung der noch fehlenden Person)

„Provenance“

- **Mehrere Dimensionen**
 - *Datenersteller*
 - *Methode, Standard etc.*
 - *Zeitpunkt der Datenerstellung*
- **Künftig wichtiges Konzept in RDA**
aus einem Entwurf im 3R-Projekt:
„Record provenance for every element if applicable.“
- **Information muss transportiert werden können**
 - *in der Linked-Data-Welt gibt es dafür Lösungen (Stichwörter: Named graphs, PROV-O)*
 - *auch in Bibframe vorgesehen*


 *eröffnet Möglichkeit, nicht mehr in „records“ zu denken, sondern in „statements“*

Von „records“ zu „statements“

- **Ressourcen-Beschreibung als Set von „statements“**
das Set kann sich im Lauf der Zeit ändern, z.B.
 - *Ergänzung eines bisher noch fehlenden „statements“*
 - *Ersetzen eines vorhandenen „statements“ durch ein gleichartiges mit besserer Qualität*
- ➔ *Beispiel: maschinelle DNB-Schlagwörter automatisch ersetzen, sobald irgendwo eine intellektuelle Verschlagwortung vorliegt*
- ➔ *in der derzeitigen Systemumgebung noch „Zukunftsmusik“, erscheint dennoch mittelfristig sinnvoll*

Weiterentwicklung von RDA

3R-Projekt berücksichtigt diese Entwicklungen:

- **Einbezug (halb-)maschineller Methoden**
neue Möglichkeit, Informationen von den Titelseiten automatisch einzulesen
 - *zusätzlich neue, weniger differenzierte Elemente (z.B. für Veröffentlichungsangabe als Ganzes)*
 - *Möglichkeit für Textübernahme ohne Normalisierung*
 - **Grundsätzlich mehrere Erfassungsmethoden**
vier Optionen (von unstrukturiertem Text über normiertes Vokabular und Identifier bis hin zu URIs)
-  *zielt ab auf eine Welt, in der Informationen kombiniert und ggf. durch bessere ersetzt werden können*

Agenda

1. Bibliotheksdaten
2. Metadaten anderer Produzenten
3. Datenqualität und Metadatenmanagement
4. Provenance
5. Wo ist Qualität besonders wichtig?

Welche Qualität für welche Objekte?

Haltung bei DNB:

- **Bewusst keine Abwägung nach Wert oder Nutzung**
Zielstandard hängt vom Vorliegen von Metadaten ab

Aus dem Erschließungskonzept von 2015 (derzeit ausgesetzt):

„Weder kann und will die DNB eine Sortierung nach Qualität der Werke vornehmen, noch kann die DNB im Vorfeld der Katalogisierung abschätzen, ob eine Titelaufnahme häufig nachgenutzt wird oder nicht. Daher verteilt sie ihre Aufwände gleichermaßen: immer dann, wenn ihr Grundelemente der bibliografischen Beschreibung als Vorleistung von Verlagen, Autorinnen und Autoren, anderen Bibliotheken u.a. vorliegen, katalogisiert sie diese auf dem Niveau des Standardelemente-Sets. Liegen ihr keine Grundelemente aus Fremdleistungen vor, katalogisiert sie die Werke auf dem Niveau des Kernelemente-Sets.“

<http://www.dnb.de/SharedDocs/Downloads/DE/DNB/standardisierung/rdaErschlie%C3%9FungDnbOktober2015.pdf>

- **Wahl der Methoden (derzeit)**
primär nach technischer Machbarkeit

Aus den „Strategischen Prioritäten 2017-2020“:

„Umstellung geeigneter Objektgruppen auf eine vollständig maschinelle und/oder auf Fremddaten basierende Erschließung“

<http://nbn-resolving.de/urn:nbn:de:101-2017021403>

Aus „Grundzüge und erste Schritte der künftigen inhaltlichen Erschließung von Publikationen in der Deutschen Nationalbibliothek“:

„Eine intellektuelle inhaltliche Erschließung erfolgt nur noch dann, wenn maschinelle Verfahren entweder nicht zur Verfügung stehen, keine ausreichenden Ergebnisse liefern oder intellektuell erstellte Daten für die Weiterentwicklung der maschinellen Verfahren benötigt werden.“

<http://www.dnb.de/DE/Erwerbung/Inhaltserschliessung/grundzuegeInhaltserschliessungMai2017.html>

Zur Diskussion

- **Ist eine solche Haltung ausreichend?**
z.B. für eine UB?
- **Falls nein: Welche Ressourcen brauchen eine besonders gute Qualität?**
mit welchen Kriterien könnten solche Ressourcen identifiziert werden?

Welche Qualität an welchen Stellen?

Relativ unproblematisch sind z.B.:

- **Abweichungen vom (bibliothekarischen) Übertragen**
*z.B. keine Normalisierung der Groß-/Kleinschreibung,
Angabe des Verlagsnamens nicht gemäß Vorlage*

Besonders wichtig sind hingegen:

- **Korrekte Angabe zentraler Informationen**
z.B. Titelfassung, Erscheinungsdatum
- **Sacherschließung**
nicht nur für Recherche, sondern auch bei Anzeige
- **Personen**
Bedeutung für Recherche, Funktion „Zusammenführen“
- **Elemente, die zum Filtern eingesetzt werden**
z.B. IMD-Typen

Beziehungen zu Personen

- **Automatische Normdatenverknüpfung**
bisher m.W. nur als reiner String-Abgleich üblich
➔ *nötig wäre ein komplexeres Verfahren,
insbes. unter Einbezug von Werk-Clustering*
- **Hier evtl. künftig Verbesserungen bei Fremddaten**
 - *stärkere Verbreitung von ORCID u.a. im Wissenschaftsbereich*
 - *Verlage erkennen die Bedeutung von eindeutigen Personen-Identifikationen für ihr eigenes Marketing*

Normdaten | 22. Januar 2018

Über den Autor zum Produkt

Verlage investieren zunehmend in qualitativ hochwertige Produktbeschreibungen. Denn die Annahme, dass gute Metadaten Bücher verkaufen, hat sich längst bestätigt. Steigende Verkaufszahlen beweisen den Mehrwert von optimal gepflegten Produktinformationen. Welche etablierten Standards für die Erschließung von Inhalten und den Austausch von Metadaten gibt es und welche sind gerade im Kommen? Wie erreicht man eine bessere Auffindbarkeit und Präsentation von Autoren durch den Einsatz von Normdaten?

[drucken](#)
[teilen](#)
[E-Mail](#)


bookbytes ist ein Angebot von

boersenblatt.net 

Mit Unterstützung von:


LITERATURTEST

Agentur für Buch-PR und Buchmarketing



Auszüge aus dem Börsenblatt-Blogbeitrag:

„(...) gilt es, vor allem den Autor als Marke bzw. als oft gewählten Sucheinstieg mehr in den Mittelpunkt des Metadatenmanagements zu stellen.

Um einen Autor in Online-Auftritten adäquat präsentieren zu können, reicht die reine Nennung des Namens längst nicht aus. Neben Fotos der Person sollten vor allem weiterführende Angaben wie Biografien, aber auch Informationen zu aktuellen Events wie Lesungen in Form von Metadaten verfügbar sein. Da man dies als Verlag nicht mit jedem Titel mitsenden und vor allem nicht an jedem Backlist-Titel aktualisieren möchte, ist es erstrebenswert, Personendaten zentral zu pflegen.

(...) Kunden, die sich für diese Person interessieren, möchten natürlich keine Produkte namensgleicher Urheber empfohlen bekommen. (...)

Um Kunden einen spezifischen Autor als funktionierenden Sucheinstieg anzubieten, muss ein Shop ihn als unverwechselbare Person identifizieren können. Dies ist nur möglich, wenn sich in den Produktmetadaten nicht nur der Urhebername findet, sondern auch für jeden Urheber eine Urheber-ID (GND, ORCID, ISNI) hinterlegt ist. Die Eintragung einer Urheber-ID muss folglich fester Bestandteil des Metadatenmanagements in jedem Verlag werden.“

<https://www.boersenblatt.net/bookbytes/artikel-normdaten.1423703.html>

*Soweit mein Input – ich
freue mich auf unsere
Diskussion!*



Zum Nach- und Weiterlesen:

- *Bruce, Th. / Hillmann, D.I.: The continuum of metadata quality – defining, expressing, exploiting. In: Hillmann, D.I. / Westbrook, E.L. (Hrsg.): Metadata in practice. Chicago: ALA, 2004, S. 238-256*
- *Eckert, K. / Pfeffer, M.: (Meta-)Datamanagement with KNIME. Folien zum Workshop auf der SWIB 2017.*
http://swib.org/swib17/slides/eckert_pfeffer_knime.pdf
- *Eckert, K.: Provenance and annotations for linked data. Proceedings of the DC-2013, Lissabon 2013.*
<http://dcpapers.dublincore.org/pubs/article/view/3669>
- *Goebel, R.: K10plus Zentral. Vortrag auf dem BSZ-Kolloquium 2017.*
<https://swop.bsz-bw.de/frontdoor/index/index/docId/1261>

- *Lazarus, J.: Das machen wir selbst – der Aufbau eines eigenen Artikelindex als Alternative zu proprietären Angeboten. Vortrag auf dem 105. Bibliothekartag 2016. <http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:0290-opus4-23591>*
- *Pfeffer, M. / Wiesenmüller, H.: Resource Discovery Systeme. In: Sühl-Strohmenger, W. (Hrsg.): Handbuch Informationskompetenz. 2. Auflage. Berlin: De Gruyter Saur, 2016. S. 105-114*
- *Wiesenmüller, H.: Den Kern erhalten – Qualität an der richtigen Stelle. In: Hohoff, U. (Hrsg.): Bibliotheken für die Zukunft – Zukunft der Bibliotheken. 100. Deutscher Bibliothekartag. Hildesheim: Olms, 2011, S. 327-337*
- *Zeng, M.L. / Qin, J.: Metadata. Second edition. London: Facet Publishing, 2016*