

31st Annual Conference of the German Classification Society on
Data Analysis, Machine Learning, and Applications

- Librarian Workshop:

Subject Indexing and Library Science -

Automatic Analysis of
Dewey Decimal Classification Notations

Automatische Analyse von Notationen der
Dewey-Dezimalklassifikation

Ulrike Reiner

Verbundzentrale des GBV (VZG), Göttingen

31. Jahrestagung der GfKI (ul, 8. März 2007)

Inhalt

- ◆ VZG-Projekt Colibri/DDC
- ◆ DDC-Notationen und DDC-System
- ◆ Analyse von DDC-Notationen
- ◆ Ergebnisse und Anwendungen

DDC: Dewey Decimal Classification
(dt. Dewey Dezimalklassifikation)



VZG-Projekt Colibri/DDC (1)

*C*ontext generation and

*L*inguistic tools for

*B*ibliographic

*R*etrieval

*I*nterfaces

000 – 020 – 510 – 511.3

VZG-Projekt Colibri/DDC (2)



Ausgewählte Colibri-Fragestellungen

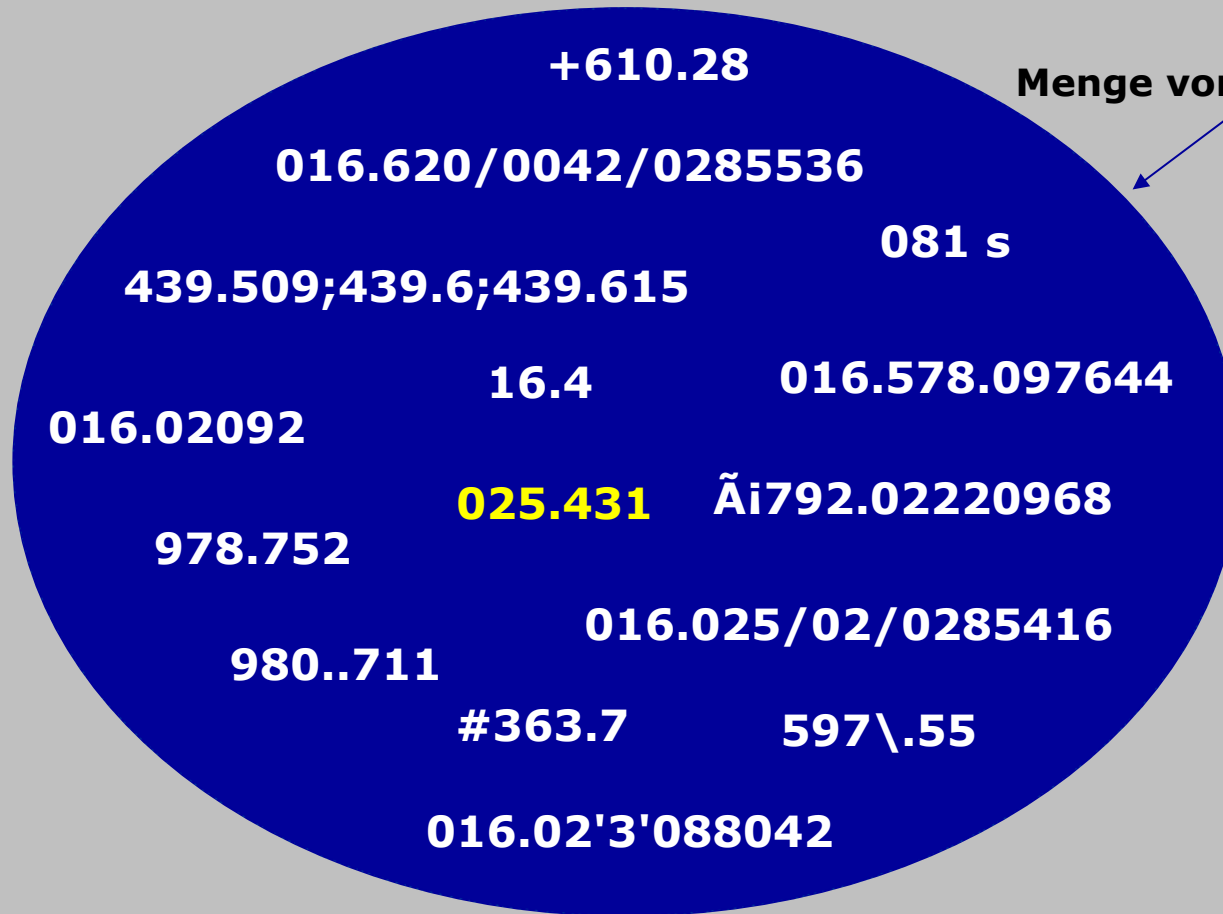
- F1:** Ist es möglich, molekulare DDC-Notationen automatisch (eindeutig) zu zerlegen?⁽¹⁾
- F2:** Ist es möglich, mit Hilfe atomarer DDC-Notationen die Klassifikation und Suche zu verbessern? ⁽¹⁾

[(1) Songqiao Liu: The automatic decomposition of DDC synthesized numbers. Los Angeles, Calif., Univ. of California, Los Angeles, Diss., 1993, p. 18]

DDC-Notationen (1)



3.339.717 DDC-Notationen in
3.073.423 bibliographischen Titeldatensätzen
 des **Gemeinsamen VerbundKataloges (GVK)**



Menge von DDC-Notationen
 „Deep Lib“

DDC-Notationen (2)



Aus **3.073.423 GVK-Titeldatensätzen** (Stand: Juli 2004) können **466.134** unterschiedliche **DDC-Notationen** gewonnen werden. Diese entstehen aus den in den GVK-Titeldatensätzen vorkommenden **3.339.717** DDC-Notationen durch **Elimination von**

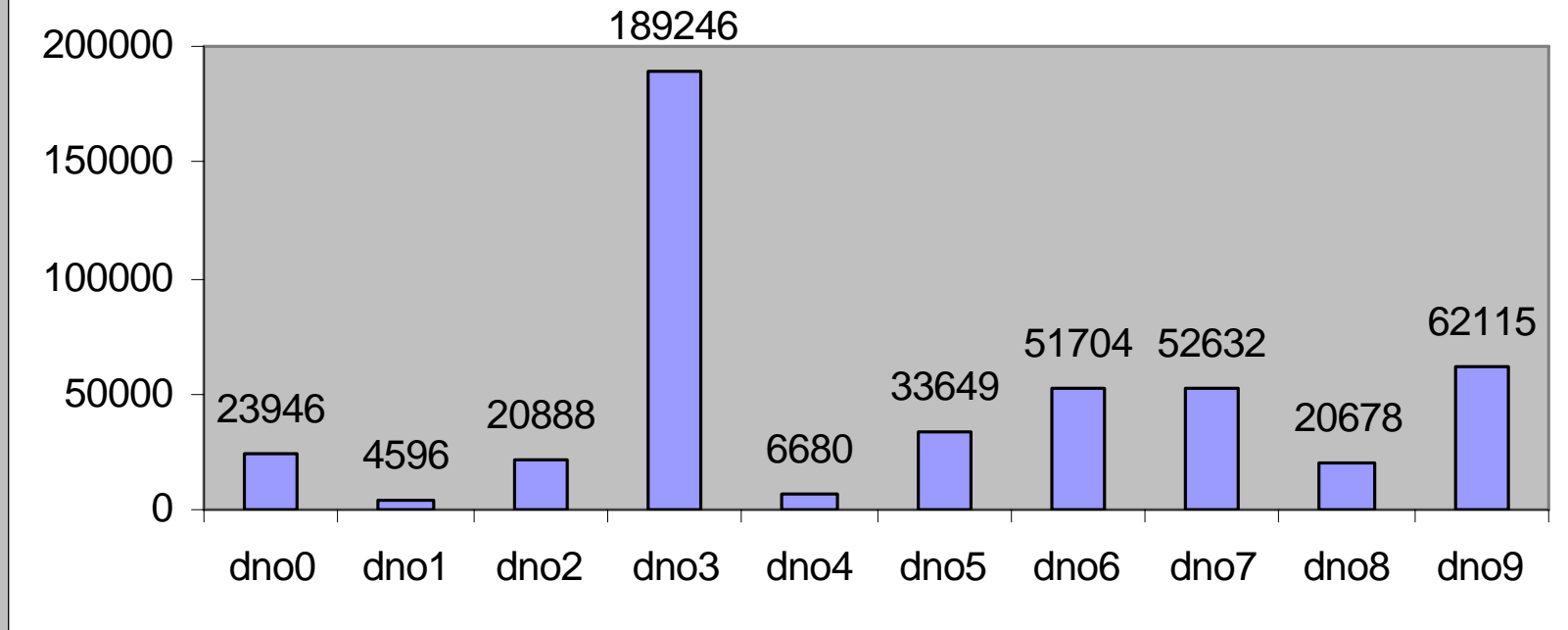
- Segmentierungszeichen
- inkorrekten DDC-Notationen
- gleichen DDC-Notationen

Häufigkeitsverteilungen der DDC-Notationen?

DDC-Notationen (3)



Häufigkeitsverteilung der GVK-DDC-Notationen (nach Zugehörigkeit zur DDC-Hauptklasse)

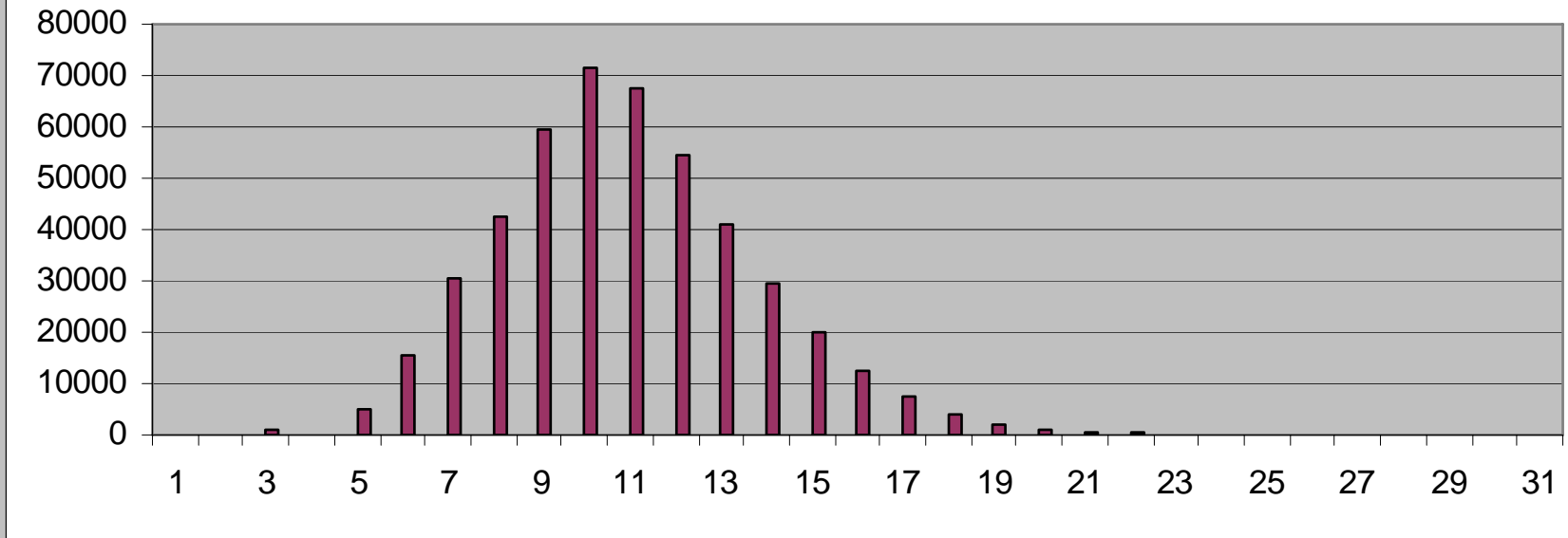


0: Informatik&Informationswissenschaft & Allgemeines; **1** Philosophie & Psychologie; **2:** Religion;
3: Sozialwissenschaften; **4:** Sprache; **5:** Naturwissenschaften & Mathematik; **6:** Technik &
Technologie; **7:** Bildende Kunst & Kunsthandwerk, **8:** Literatur & Rhetorik, **9:** Geschichte & Geografie

31. Jahrestagung der GfKI (ul, 8. März 2007)



Häufigkeitsverteilung der GVK-DDC-Notationen (nach Stelligkeit)

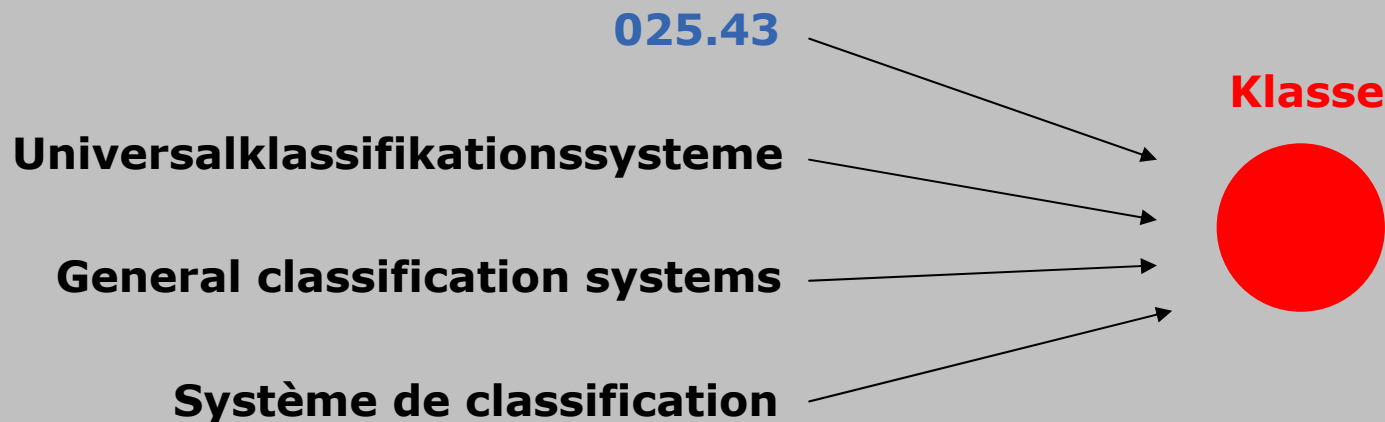


DDC-Notationen (5)



„**Notation** is the system of symbols used to represent the classes in a classification system. ... The notation provides a universal language to identify the class and related classes, regardless of the fact that different words or languages may be used to describe the class. ”

[Online WebDewey Introduction: <http://connexion.oclc.org/>, Zitinfärbung von ul]



DDC-Notationen (6)

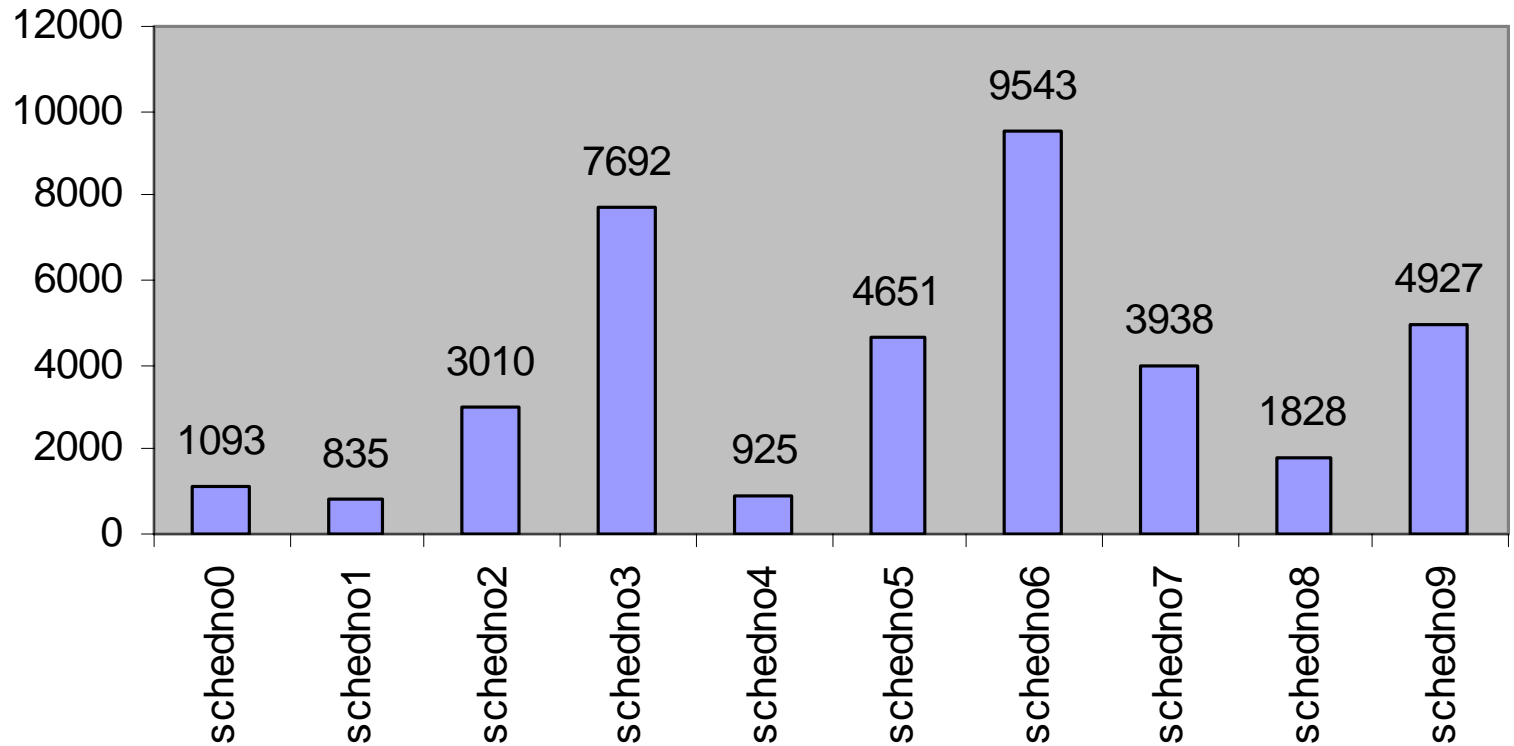


Notationen (Notationsteile) im DDC-System (1)

- 3 Übersichten (summaries)
 - * Hauptklassen (main classes or tens)
 - * Klassen der Ebene 1 (divisions or hundreds)
 - * Klassen der Ebene 2 (sections or thousands)
- Haupttafeln (schedules)
 - * 000, 001, 001.01, 001.012, ... , 997.3, 998, 998.2, 999
- Hilfstafeln (tables)
 - * T1--, T2--, T3A--, T3B--, T3C--, T4--, T5-- und T6--
- Interne Tafeln (internal tables)
 - * Anhängetafeln (add tables), Zeittafeln (build tables), Übersichten der Vorzugsreihenfolge (tables of preference), Praxishilfe (manual tables), ? (history tables), ? (generic tables), ? (option tables)
- Register (relative index)



Häufigkeitsverteilung der Haupttafel-Notationen



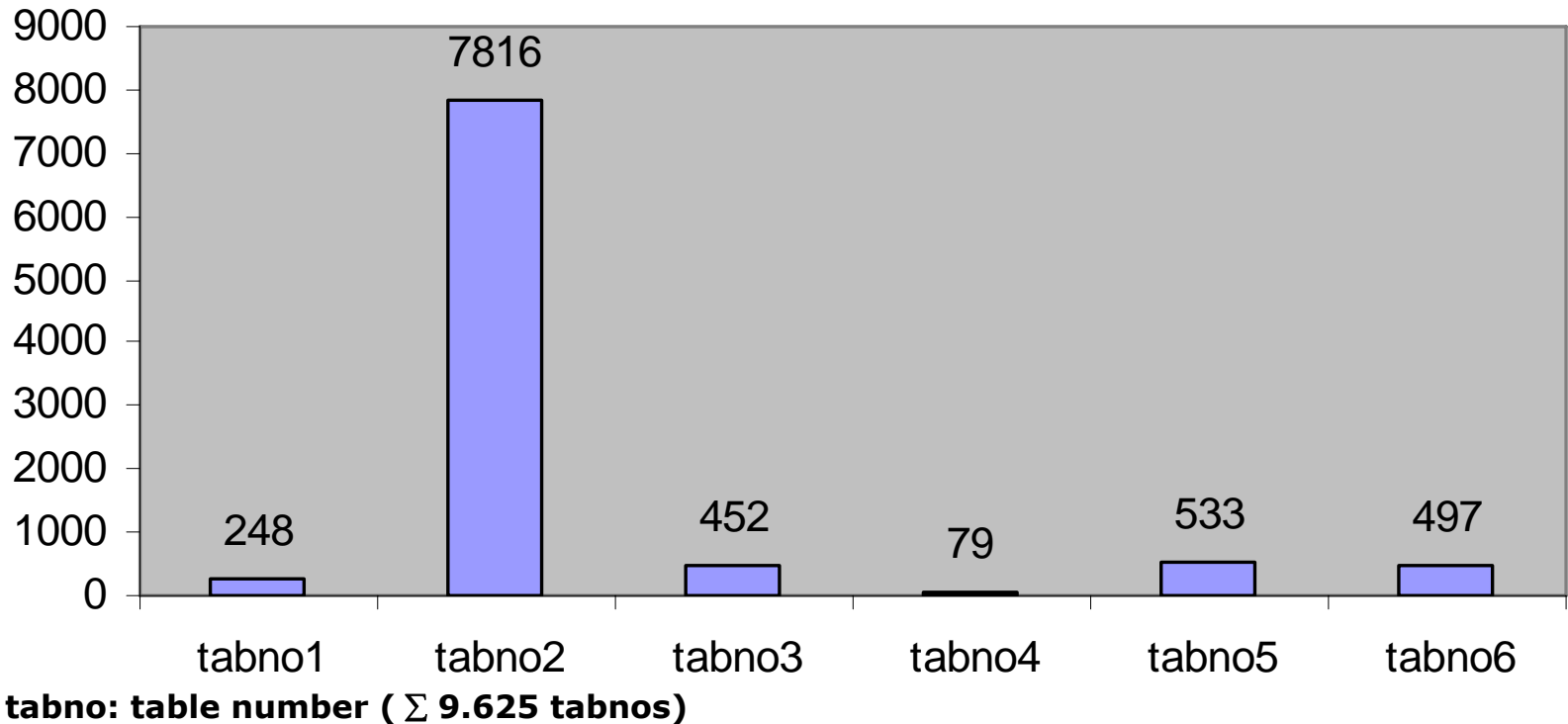
schedno: schedule number (Σ 38.442 schednos)

0: Informatik&Informationswissenschaft & Allgemeines; **1** Philosophie & Psychologie; **2:** Religion;
3: Sozialwissenschaften; **4:** Sprache; **5:** Naturwissenschaften & Mathematik; **6:** Technik &
Technologie; **7:** Bildende Kunst & Kunsthandwerk, **8:** Literatur & Rhetorik, **9:** Geschichte & Geografie

31. Jahrestagung der GfKI (ul, 8. März 2007)



Häufigkeitsverteilung der Hilfstafel-Notationen



tabno1: Standard Subdivisions; tabno2: Geographic Areas, Historical Periods, Persons; tabno3: Subdivisions for the Arts, for Individual Literatures, for Specific Literary Forms; tabno4: Subdivisions of Individual Languages and Language Families; tabno5: Ethnic and National Groups; tabno6: Languages

31. Jahrestagung der GfKI (ul, 8. März 2007)

DDC-Notationen (9)



Notationen (Notationsteile) im DDC-System (2)

- Menge von DDC-Notationen

(Elemente einer aufzählenden Klassifikation)

- 000 Computer science, information & general works
- 020 Library & information sciences
- 025 Operations of libraries, archives, information centers
- 025.6 Circulation services
- 025.62 Interlibrary loans

- Menge von DDC-Notationsteilen

(Elemente einer Facettenklassifikation)

Aus Anhängetafel T1--093-T1-099:

- 01 Forecasts
- 021 Statistics
- ...
- 075 Collecting objects
- 09 Historical and geographic treatment

31. Jahrestagung der GfKI (ul, 8. März 2007)

DDC-Notationen (10)



Auszug aus WebDewey, Haupttafel (WebDewey Edition 22)

020	Library & information sciences
020	Library and information sciences
021	Relationships of libraries, archives, information centers
022	Administration of physical plant
023	Personnel management (Human resource management)
→ [024]	[Unassigned]
025	Operations of libraries, archives, information centers
→ 025.001-025.009	Standard subdivisions
→ 025.02	Technical services
→ 025.04	Information storage and retrieval systems
→ 025.06	Information storage and retrieval systems devoted to specific disciplines and subjects
→ 025.1	Administration
→ 025.8	Maintenance and preservation of collections
030	Encyclopedias & books of facts

=> Dewey Dezimalklassifikation: nicht vollständiger dekadischer Baum

Synthese von DDC-Notationen (1)



Anzahl erzeugbarer DDC-Notationen des DDC-Systems

19th ed. of the DDC

„ ... a formula for calculating ... potential classes in the system ...

$$21.504 \times [116 + (5 \times 6.933) + (2 \times 6.500) + 192 + 440] = \mathbf{1.041.073.100} \quad (*)$$

22nd ed. of the DDC (ul berechnet nach Formel aus (*))

$$38.442 \times [248 + (5 \times 7.816) + (2 \times 7.269) + 533 + 497] = \mathbf{2.110.312.032}$$

Gegenüber der Vorgängerversionen enthält die 22. Auflage wesentlich mehr DDC-Notationen und Möglichkeiten zur Notationsbildung

=> höhere Komplexität des heutigen DDC-Systems als früher

[(*) Francis L. Miksa: „The DDC, the Universe of Knowledge, and the Post-Modern Library“. Forest Press, Albany, New York, 1998, S. 24]



Erzeugen synthetischer DDC-Notationen

Es ist erlaubt, DDC-Notationen (-Notationsteile)

A1. an eine DDC-Grundnotation aufgrund von Synthese-Anweisungen(instructions)

A2. an eine DDC-Notation aus Hilfstafel 1 (Table 1: Standard Subdivisions), falls nicht ausdrücklich ausgeschlossen,

anzufügen

Synthese von DDC-Notationen (3)



„Die Notationssynthese beginnt mit einer **Grundnotation** (die **immer** in der Anweisung angegeben ist), an die eine andere Notation angehängt wird. ...“

[glossar_EG_OCLC_exp.pdf, Zitateinfärbung von ul]

„Table 3A. ... The base number may be identified in an add note, e.g., at 820.1-828 (“add to base number **82**”) or another note, e.g., at 896 (...); otherwise, it is the number given for the literature, e.g., Dutch-language literature **839.31**. ...

...

896 (African literatures) ... Add to **896** ...“

[WebDewey Edition 22]

=> **es steht nicht immer** „ ... Add to base number ...“ !

In der elektronischen 22. DDC-Ausgabe (kurz: e22) sind mind. **1580** DDC-Grundnotationen (automatisch ermittelt) enthalten

Synthese von DDC-Notationen (4)



Auszug aus WebDewey, Hilfstafel 1 (WebDewey Edition 22)

T1--0	Table 1. Standard Subdivisions
T1--01	Philosophy and theory
T1--02	Miscellany
T1--0202	Synopses and outlines
T1--0207	Humorous treatment
T1--0208	Audiovisual treatment
T1--021	Tabulated and related materials
...	
T1--03	Dictionaries, encyclopedias, concordances
T1--04	Special topics
T1--05	Serial publications
T1--06	Organizations and management
T1--07	Education, research, related topics
T1--08	History and description with respect to kinds of persons
T1--09	Historical, geographic, persons treatment

31. Jahrestagung der GfKI (ul, 8. März 2007)

Synthese von DDC-Notationen (5)



Beispiel zur Synthese einer DDC-Notation (1)

Die DDC-Notation (DDC number)

025.17

hat die Klassenbenennung (caption)

Administration of collections of special materials

Nach A1. kann mit der Syntheseanweisung (instruction)

Add to base number **025.17** the numbers following **025.34** in
025.341-025.349

die neue (synthetische) Notation

025.174 (Electronic resources)

gebildet werden.

Synthese von DDC-Notationen (6)



Beispiel zur Synthese einer DDC-Notation (2)

Nach A2. kann an die DDC-Notation

025.174

der DDC-Notationsteil

09

angefügt werden, die neue (synthetische) DDC-Notation lautet

025.17409

Nach A1. ist es mit der Syntheseanweisung

Add to base number T1--09 notation T2--3-T2—9 from Table 2

erlaubt, den DDC-Notationsteil

73

hinzuzufügen, die synthetische DDC-Notation lautet nun

025.1740973

(begonnen mit 025.17)

Analyse von DDC-Notationen (1)



DDC-Notationsanalyse durch vc_day

```

025.1740973 <dno_gvk_all_17389_to_analyze; length: 11>
0----- Computer science, information & general works <hatzen>
02----- Library & information sciences <hatzen>
025----- Operations of libraries, archives, information centers <hat>
025.1----- Administration <hat>
025.17----- Administration of collections of special materials <hat>
025.174---- CD-ROMs--library treatment,... <hatien>
---.---4---- Electronic resources <nalr1:025.344>
---.---09-- Historical, geographic, persons treatment <T1--09>
---.---097- Treatment by specific continents, countries, localities;
extraterrestrial worlds <T1--097>
---.-----7- North America <ba4r2span:T1--093-T1--099:T2--7>
---.-----73 United States <ba4r2span:T1--093-T1--099:T2--73>

```

<i>vc_day</i>	vzg colibri_ddd number <i>analyzer</i>
<i>vc_daygram</i>	vzg colibri_ddd <i>analysis diagram</i>
<hat>	hierarchy at class (e22 xml tag)
<hatzen>	gebildet aus <hat> und <zen>: zen built entry (e22 xml tag)
<hatien>	gebildet aus <hat> und <ien>: built schedule entry (e22 xml tag)



DDC-Notationssynthese

(notational synthesis, number building)

Erzeugung einer molekularen DDC-Notation mit DDC-Regeln

DDC-Notationsanalyse

(analysis of notations, decomposition of notations)

Analyse einer molekularen DDC-Notation zur Gewinnung aller in ihr enthaltenen atomaren DDC-Notationen und ihren Bedeutungen



Atomare DDC-Notation (*dno_atom*)

Semantisch kleinste Einheit, die eine DDC-Klasse repräsentiert

dno_atom: 0 (Computer science, information & general works), ...

dno_atom: 025.174 (CD-ROMs--library treatment,...), ...

dno_atom: 025.344 (Electronic resources)

dno_atom: T2—73 (United States)

Molekulare DDC-Notation (*dno_mol*) (Synthetische DDC-Notation)

Nach Regeln (Synthese-Anweisung) zusammengesetzte DDC-Notation

dno_mol: 025.174

025.17-*<na1r1>*-rule Add to base number 025.17 the numbers following 025.34 in 025.341-025.349

Modell (3)



Ausgezeichnete *dno_atoms*

1. *dno_atom* beginnend mit 0, 1, ..., 9

(Notation aus Haupttafel, 045H/02)

025.344

2. *dno_atom* beginnend mit T1--, ..., T3A--, ..., T3C--,

T4--, ..., T6-- (Notation aus Hilfstafel, 045H/03)

T1-09

3. base number *bano* (Grundnotation, 045H/01)

025.17, 55, T1-09, T3C--382

4. prefix *dno predno* (Ausgangsnotation) *vc_dayset_mab2*

5. wie in 3. oder 4. gefolgt von „+“ Ziffernfolge (*vc_KB-Repräsentation*)

(Notation aus interner Tafel, 045H/04)

T1--093-T1--099+021 (Statistics)

913-919+0025 (Directories of persons and organizations)

Analyse von DDC-Notationen (2)



Stand der Technik (Zerlegung von DDC-/UDK-Notationen)

- [Liu1993]** Songqiao Liu: *The automatic decomposition of DDC synthesized numbers*. Los Angeles, Calif., Univ. of California, Los Angeles, Diss., 1993.
- [Riesthuis1997]** Gerhard J. A. Riesthuis: *Decomposition of complex UDC notations*. In: Knowledge Organization for Information Retrieval. Proceedings of the Sixth International Study Conference on Classification Research. University College London, 16-18 June 1997. International Federation for Information and Documentation (FID), The Hague, Netherlands, 1997, pp. 139-143.
- [Reiner2005]** Ulrike Reiner: *DDC-Notationsanalyse und –synthese*. Sept. 2004 - Feb. 2005. VZG-Colibri-Bericht 2/2004. Verbundzentrale des Gemeinsamen Bibliotheksverbundes (VZG des GBV), Göttingen, 2005.

Analyse von DDC-Notationen (3)



Riesthuis1997 (Decomposition of complex UDC notations)

- * **complex notation: connectors (':', '::', '+')**
simple notation: splitted complex notation

- * **Beispiele von UDC-Notationen (p. 141):**

323.1/.2(47+57):81'01(47+57)

**National and ethnic minorities. Internal political activities.
Origins (of languages). Soviet Union**

327(44)(092)Talleyrand(044.4)

**International affairs. France. Biographical presentation. Talleyrand.
Official correspondence**

611.018.1:616-073.572-091.8:613.221(043)

**Cytology. Diagnosis by methods adopted from physics:
electrostatics. Cytopathology. Nutrition of infants. Dissertations**

- * **In natürlicher Sprache skizzierte Algorithmen (pp.141-143):**

Analyse von DDC-Notationen (4)



Arbeitsschritt 1 (Syntheseanweisungen)

Reiner2005

R1: Studium und Analyse der Syntheseanweisungen, Auszug aus e22:

```

...
<hat><lnk refC="ddc: :000" refR="ddc:025.62"
  span1="025.62"><txt>025.62</txt></lnk><txt>Interlibrary loans</txt></hat>
...
<hat><lnk refC="ddc: :T1" refR="ddc:T1--093-099" tbl="1" span1="093"
  span2="099"><txt>T1â€"093-T1â€"099</txt></lnk><txt>Treatment by specific continents,
  countries, localities; extraterrestrial worlds</txt></hat>
...
<ba4><txt>Add to base number T1â€"09 notation </txt><lnk refC="ddc: :T2" refR="ddc:T2--3-9"
  tbl="2" span1="3" span2="9"><txt>T2â€"3-T2â€"9</txt></lnk><txt> from TableÂ 2, e.g.,
  the subject in North America </txt><lnk refC="ddc: :T1" refR="ddc:T1--093-099" tbl="1"
  span1="097"><txt>T1â€"097</txt></lnk><txt>, in Brazil </txt><lnk refC="ddc: :T1"
  refR="ddc:T1--093-099" tbl="1" span1="0981"><txt>T1â€"0981</txt></lnk><txt>;
  then add further as follows: </txt></ba4>
<hn><txt>T2--7</txt></hn>
<hn><txt>T2--81</txt></hn>
<set><txt>Segment as shown in TableÂ 2; if the TableÂ 2 number does not have a segmentation
  mark, segment after the TableÂ 2 number</txt></set>
<ren><txt>01</txt></ren>
<reh><txt>Forecasts</txt></reh>
<ren><txt>02</txt></ren>
<reh><txt>Statistics, illustrations; conservation, preservation, restoration</txt></reh>
...

```

31. Jahrestagung der GfKI (ul, 8. März 2007)

Analyse von DDC-Notationen (5)



Arbeitsschritt 1 (Syntheseanweisungen)

Reiner2005

R1: Klassifizierung der e22-DDC-Syntheseanweisungen, Beispiele:

<na1> Add to base number ... the numbers following ... in ...

<na5> Add to each subdivision identified by ... as instructed under ...

<na6> Add notation ... from table under

<nad> Add to ... the historical period numbers following ... that appear in subdivisions of ...

<nfa> Add [to base number|historical periods] as instructed under ...

<nfu> Use notation ... from Table 1 as modified at ...

<ba4> Add to ... notation from the period table ... then add further as follows, ...

R1: Bestimmung der Anzahl der e22-DDC-Syntheseanweisungen (7.705)

```
colibri/ul-test> egrep -i "add to base number.*(notation|the numbers following.*[notation]*)|
  add to each subdivision identified by.*(as follows|notation|the numbers following|as instructed
  (at|under))|(add [to]*|add to base number).*as instructed (at|under)|(add.*notation.* from.*
  table.*[under]*)|add [to]*.*(the [historical period|period division]*numbers following|notation)
  " e22 | wc -l
7705
```

R1: 25 DDC-Regeln für *vc_day* (vzg colibri_ddc number analyzer)

L1: 17 rules for decomposition (class 700, tables and internal tables)

Analyse von DDC-Notationen (5)



Arbeitsschritt 1 (Syntheseanweisungen)

R1: 25 DDC-Regeln für *vc_day* (*vzg colibri_ddc number analyzer*)

r0 = "[Aa]dd to base number.*three-digit numbers"	# <na2>
r1 = "[Aa]dd to base number.*the numbers following.*in"	# <ba[13]>,<na[134]>
r2 = "[Aa]dd to.*notation"	# <ba4>,<na[246]>
r3 = "[Aa]dd to each subdivision identified by"	# <ba5>,<na5>
r4 = "told to add as instructed under"	# <ba5>
r5 = "[Aa]dd to.*the.*numbers following.*in"	# <na[13]>,<nad>
r6 = "[Aa]dd the numbers following.*in"	# <na1>
r7 = "[Aa]dd to base number.*notation"	# <na2>
r8 = "[Aa]dd to.*as instructed under"	# <na6>
r9 = "[Aa]dd as instructed under"	# <na6>,<nfa>
r10 = "[Aa]s modified under"	# <nst>
r11 = "before adding"	# <ba1>,<mt>
r12 = "[Aa]dd notation.*from table under"	# <na6>
r13 = "[Aa]dd to base number.*as instructed at beginning of"	# <na6>,<nfa>
r14 = "then.*add further as instructed"	# nasi16, nasi26, nasi34, nasi36, nasi46
r15 = "then.*add.*0.*and.*add notation"	# nasi24, nasi34, nasi44
r16 = "then.*add.*[01].*and to the result add.*the numbers following.*in"	# nasi31, nasi41, na1nad
r17 = "then to the result add the numbers following.*in"	# nasi31, nasi41
r18 = "then add.*0 and to the result add further as instructed"	# nasi46
r19 = "[Aa]dd to base number as instructed under"	# <nfa>
r20 = "Add historical periods as instructed under"	# <nfa>
r21 = "then.*add notation"	# <nad>
r22 = "then.*add standard sub"	# <nad>
r23 = "then.*add historical period numbers"	# <nad>
r24 = "[Aa]dd.*0.*to base number"	# <nad>
r25 = "[Uu]se notation.*under each subdivision identified"	# <nsi>

31. Jahrestagung der GfKI (ul, 8. März 2007)

Analyse von DDC-Notationen (7)



Arbeitsschritt 2 (Stichprobe)

Liu1993

L2: Erstellung der Stichprobe (Umfang: 600)

Reiner2005

R2: Verwendung von Liu's Stichprobe.

Erstellung als Eingabedatei *in_liu_t* für Analysekomponente *vc_day* :

```
-----  
in_liu_1  
700.23  
in_liu_2  
700.90440747471  
...  
in_liu_600  
799.26092  
-----
```

**von den 600 DDC-Notationen von Liu's Stichprobe sind (nach 8 Jahren)
36 DDC-Notationen „veraltet“ (stillgelegte oder verlegte DDC-Notationen)**

Analyse von DDC-Notationen (8)



Arbeitsschritt 2 (Stichprobe)

L2: Umfang der Stichprobe (sample size for desired margin of error) (1)

$$n = (2z^*/w)^2 p^* (1-p^*)$$

w : Fehlerbereich (margin of error), auch Genauigkeit

p* : geschätzter Anteil zerlegbarer DDC-Notationen

z* : tabellierte statistische Funktion

Für $p^* = 0.5$

$$n = (z^*/2w)^2$$

Bei Konfidenzniveau 95% ($z^* = 1.96$) und Fehlerbereich $-0.04 \leq w \leq 0.04$ (Genauigkeit 4%) $\Rightarrow n = (1.96/.08)^2 = 600$ (Stichprobenumfang)

[(1) David S. Moore; George P. McCabe: Introduction to the Practice of Statistics. 3rd ed., W. H. Freeman and Company, New York, 1998, p. 593]

Analyse von DDC-Notationen (9)



Arbeitsschritt 2 (Stichprobe)

L2: Recherche im OCLC Union Catalog (EPIC, telecommunication service)

gesucht mit: 700?, 701?, ..., 700.1?, 700.2?, ...

gefunden: 11.662(*) bibliographische OCLC-Datensätze

in Datenbank abgespeichert: 1.743 bibliographische OCLC-Datensätze

(auf vollständiger, 20. DDC-Ausgabe beruhende, einmalig vorkommende, synthetische, von der LoC vergebenene DDC-Notationen)

L2: Grundgesamtheit (1.743-42) = 1.701 DDC-Notationen

(42 inkorrekt gebildete DDC-Notationen)

L2: Stichprobe: per Zufallsgenerator EPISTAT 600 DDC-Notation

bestimmt (EPISTAT: public-domain compute statistical package)

(*) zum Vergleich: 287.704 bibliographische GVK-Datensätze der Klasse 700 (Stand: Juli 2004)

Analyse von DDC-Notationen (10)



Arbeitsschritt 3 (DDC-Wissensbasis)

Liu1993

L3: Erstellung der DDC Datenbank/Wissensbank (OCLC Daten, 20th Ed. of DDC)

Reiner2005

R3: Erstellung der DDC-Wissensbasis *vc_KB*(1) mit Programm *vc_ckb*(2) aus *e22* (22nd Ed. of DDC). Wissensrepräsentation von Fakten und Regeln:

dno (3) Ÿ descr_val bzw. regelspezifische descr_vals

025.344 Ÿ <hat> Ÿ Electronic resources
 025.17 Ÿ <na1r1> Ÿ Ÿ 025.17 Ÿ 025.341-025.349 Ÿ 025.34 Ÿ Ÿ Ÿ Ÿ Ÿ
 T1--093-T1--099+021 Ÿ <ba4r2> Ÿ Statistics

Bedeutung

Fakt: Die Klassenbenennung von „025.344“ ist „Electronic resources“

na1r1-Regel: Wenn es sich um die Grundnotation „025.17“ handelt, dann ist es erlaubt, an die Grundnotation 025.17 die Ziffern anzufügen, die der DDC-Notation 025.34 im Bereich 025.341-025.349 folgen

Fakt: Die Klassenbenennung von T1--093-T1--099+021 ist „Statistics“

(1) *vzg colibri_ddc Knowledge Base*; (2) *vzg colibri_create ddc knowledge base*

(3) Ÿ: florin, zur Trennung von Datenfeldern

Analyse von DDC-Notationen (13)



Arbeitsschritt 3 (DDC-Wissensbasis) R3: Erstellung der DDC-Wissensbasis *vc_KB*

***vc_KB* : 48.067 Fakten, 8.033 Regeln (2 MB)**
***vc_KB_spli* : 121.035 Fakten, 80.324 Regeln (15 MB)**

Ausschnitt aus *vc_KB_spli*

...

025ÿ<hat>ÿOperations of libraries, archives, information centers

025.001-025.009ÿ<hat>ÿStandard subdivisions <--- 1. has to be splitted

025.001ÿ<hatspan:025.001-025.009>ÿStandard subdivisions

025.002ÿ<hatspan:025.001-025.009>ÿStandard subdivisions

...

025.008ÿ<hatspan:025.001-025.009>ÿStandard subdivisions

025.009ÿ<hatspan:025.001-025.009>ÿStandard subdivisions

025.00285ÿ<hatien>ÿLibraries--automation,...

025.00285536ÿ<hatien>ÿLibraries--automation--microcomputer programs,...

025.02ÿ<hat>ÿTechnical services

025.04ÿ<hat>ÿInformation storage and retrieval systems

025.06ÿ<hat>ÿInformation storage and retrieval systems devoted to specific disciplines and subjects

025.060001-025.060009ÿ<hat>ÿ[Standard subdivisions] <--- 1. has to be splitted

025.060001ÿ<hatspan:025.060001-025.060009>ÿ[Standard subdivisions]

...

039ÿ<hatspan:031-039>ÿGeneral encyclopedic works in specific languages and language families

031-039+01ÿ<ba5>ÿPhilosophy and theory <--- 3. has to be splitted

031+01ÿ<ba5span:031-039+01>ÿPhilosophy and theoryÿÿÿÿÿÿÿÿ

032+01ÿ<ba5span:031-039+01>ÿPhilosophy and theoryÿÿÿÿÿÿÿÿ

033+01ÿ<ba5span:031-039+01>ÿPhilosophy and theoryÿÿÿÿÿÿÿÿ

...

31. Jahrestagung der GfKI (ul, 8. März 2007)

Analyse von DDC-Notationen (14)



Arbeitsschritte 4-7

Liu1993

**L4: Erstellung Computersystem DND* (Dewey Number Decomposer)
realisiert in FoxPro** (* unseres Wissens nach nicht verfügbar)

L5: Durchführung der Zerlegung der 600 DDC-Notationen (Hauptklasse 700)

L6: Expertenbewertung der Ergebnisse der Zerlegung

L7: Schlussfolgerungen / Anwendungsbereiche

Reiner2005

R4: Erstellung der Analysekomponente *vc_day* realisiert in gawk

**R5: Durchführung der Analyse der 600 DDC-Notationen,
Durchführung der Analyse aller GVK-DDC-Notationen (alle DDC-Klassen)**

R6: Ergebnisse, Vergleich DND mit *vc_day*

R7: dto. Liu1993

Analyse von DDC-Notationen (15)



R4: Erstellung der Analysekomponente *vc_day* (1)

Syntaktische Betrachtungen (1)

Alphabet (endliche Menge von Symbolen)

- 1. Numerische Symbole (Ziffern): 0,1,2,3,4,5,6,7,8,9**
- 2. Alphabetische Symbole (Buchstaben): A, B, C, ... für Optionen**
- 3. Dewey-Punkt: "."**
- 4. Segmentierungssymbole: "' ' " bzw. "/" (ev. auch andere)**

Beispiele:

000, 016.620/0042/0285536, 016.02'3'088042, 016.02092

05H (Option A: Hindi-language serial publications)

344.73/0125763347.30412576325763347.30412576 (ppn180304127)

720.2224945420903307449494542 (29-stellig)



DDC Appendix: Segmentation in DDC numbers

- „• One aid to reduction of the full DDC number is the segmentation provided in DDC numbers assigned by such centralized cataloging services as the Decimal Classification Division of the Library of Congress and the Library and Archives Canada.
- The segmentation is indicated by a prime mark ('), a slash mark (/), or other comparable indicators.
 - The segmentation mark indicates the end of the abridged number (including, if applicable, the end of the abridged standard subdivision in the abridged number).“

[<http://connexion.oclc.org/WebZ/CorcHelp>]

Analyse von DDC-Notationen (17)



R4: Erstellung der Analysekomponente *vc_day* (2)

Syntaktische Betrachtungen (2)

Korrektheit

mögliche Fehler: zu wenig oder zu viele Ziffern, an falscher Stelle stehende Ziffern(folgen), Ziffernfolge in verdrehter Reihenfolge, fehlender "Deweypunkt", zuviele "Deweypunkte", an falscher Stelle stehende Buchstaben, ...

Entscheidbarkeit

Die Eigenschaft einer Zeichenkette, DDC-Notation zu sein, ist entscheidbar.

(Offene) Fragen:

- Welcher Sprachtyp wird durch die Menge aller DDC-Notationen definiert?
- Welcher Grammtiktyp wird durch die Syntheseanweisungen definiert?
- Ist die Notationssynthese eindeutig?
- Ist die Notationssynthese endlich?
- Komplexität des Algorithmus zur Analyse von DDC-Notationen?

Analyse von DDC-Notationen (18)



R4: Erstellung der Analysekomponente *vc_day* (4)

Syntaktische Betrachtungen (4), Erzeugen synthetischer DDC-Notationen (2)

```

C          ::= ""
DI         ::= „0“ | „1“ | „2“ | „3“ | „4“ | „5“ | „6“ | „7“ | „8“ | „9“
DO         ::= „.“
bano      ::= schedno | tabno
clano     ::= schedno | tabno
dno       ::= dno_atom | dno_mol | dno_head dno_tail
dno_atom  ::= DI | DI C DI | schedno_atom | tabno_atom
dno_head  ::= sm | sd | ss | s
dno_mol   ::= DI C DI | bano dno | schedno_atom t1_atom | schedno_mol | t1_atom |
             schedno_atom t1_mol | schedno_atom t1_atom | dno_mol dno_atom
dno_span  ::= dno_span1-dno_span2
dno_span1 ::= dno
dno_span2 ::= dno
dno_tail  ::= [ DI C DO C ]* DI ([C DI ])*
s         ::= ss C DO C dno_tail
sd        ::= sm C DI
sm        ::= DI
ss        ::= sd C DI
schedno   ::= dno | dno_span | s
tabno     ::= dno | dno_span
...

```

=> reguläre Grammatik zur Erzeugung von DDC-Notationen (Vermutung)

Analyse von DDC-Notationen (19)



R4: Erstellung der Analysekomponente *vc_day* (5)

Semantische Betrachtungen

```
colibri/ul-test> egrep "47-" res/res_liu_t | egrep -v liu_ | sort -u
---.---4347----- Stuttgart and Tübingen districts
      <ba4r2span:T1--093-T1--099:T2--4347>
---.---47-- Eastern Europe  Russia <na4r7:T2--47>
721.0447----- Metals <hat>
739.47-- [Historical, geographic, persons treatment] <hat>
747---- Interior decoration <hat>
---.-747- New York <ba4r2span:T1--093-T1--099:T2--747>
---.-----7947- Southern Coast Range counties <na4r7:T2--7947>
```

„47“ hat je nach Stelle und Kontext unterschiedliche Bedeutung

„74“ hat andere Bedeutung als „47“ (Reihenfolgerelevanz), z. B.

74 Drawing & decorative arts

=> semantische Analyse von DDC-Notationen erforderlich

Analyse von DDC-Notationen (20)



L4: Erstellung Computersystem DND

Arbeitsschritte zur Zerlegung synthetischer DDC-Notationen

- „1. Match the number to be decomposed against the Schedule by dropping digits on the right. If the complete number is found in the Schedule, the number is not a synthesized number and decomposition is complete. If no match is made after all digits are dropped, set the number aside as containing a potential error.**
- 2. When a match is made, search the entry in the Schedule for an Add Note. If one is found, determine which note type it is and apply the rule defined for that type. If none is found, apply appropriate rules for Standard Subdivisions.**
- 3. Repeat the above two steps for the remaining digits, but in repetitions, in the first step the number may be searched against the Tables rather than the Schedule.”**

Analyse von DDC-Notationen (21)



R4: Erstellung der Analysekomponente *vc_day* (6)

in 1	700.23	the subj as poj.			
in 2	700.0440747471	1940-1949	museum	new york	
in 3	700.942	historical	historical	england and wales	
in 4	700.972910904	historical	cuba	20th century	
in 5	700.97305	natural endowment	serial publ.		
in 6	700.97309034	19th century			
in 7	700.97309045	1950-1959			
in 8	700.974710904	historical	new york	20th century	
in 9	700.97949407479494	historical	los ang.	museums	los ang.
in 10	704.0396073	african americans			
in 11	704.03972	athabascan indians			
in 21	709.470744745	historical	russia	museums	tartar republic
in 22	709.4709048	historical	russia	1980-1989	
in 23	709.49209032	historical	netherlands	17th century	1600-1699
in 24	709.598	historical	indonesia and east timor		
in 25	709.730747445	american art	museums	essex	country
in 26	711.4092	city planners			
in 27	711.40977311	local	historical	chicago	
in 28	712.0942	historical	england and wales		
in 29	712.0945632	historical	rome		
in 30	712.0973092	historical	united states	persons	

„United States“ nicht entdeckt

... erste Schritte ...

nicht analysiert

grobe Zerlegung

Analyse von DDC-Notationen (22)



R4: Erstellung der Analysekomponente *vc_day* (7)

Algorithmus zur Analyse von DDC-Notationen (1)

BEGIN

```

1. FUNCTION fu_init_days_variables
2. FUNCTION fu_read_vc_kb_spli
3. WHILE (loop forever until break) DO BEGIN
4. | WHILE (port SRVPORT 7654 is connected) DO BEGIN
| | | a. count infile_lines
| | | b. determine dno, count number of (correct, incorrect) dnos
| | | c. if (vc_daygram_test_pri is 1) then print vc_daygram_test_pri_header
| | | d. phase1
| | |   1. FUNCTION fu_get_facts (from left to right)
| | | e. phase2
| | |   1. initialization
| | |   2. FUNCTION fu_get_facts_via_rules (from left to right)
| | | f. FUNCTION fu_chk_pri (determine output format, print)
| | | g. determine whether dno is totally complete or partially complete analyzed, count no. of dnos
| | | h. print -- FINI -- on SRVPORT
| | | i. FUNCTION fu_init_days_variables
| | END
| | print "client has closed the connection, classifying has finished"
| | close SRVPORT
| | PRINT statistics (number of [totally/partially complete analyzed dnos], ...)
| END
END

```

Analyse von DDC-Notationen (23)



R4: Erstellung der Analysekomponente *vc_day* (8)

Algorithmus zur Analyse von DDC-Notationen (2)

```

FUNCTION fu_get_facts_via_rules
| 1. FUNCTION fu_init_phase2
| 2. WHILE (there is an unanalyzed dno_tail)
|   | a. determine dno_head, dno_tail and their length
|   |   (dno_head and dno_tail serves as data structure of dno,
|   |   analyzing begins with dno_head of length 1 and ends with
|   |   length of dno_head X (X: length of dno)
|   |   length of dno_tail is length of dno minus length of dno_head)
|   | b. FUNCTION fu_execute_phase2
|   | c. adjust dno_head, dno_tail
|   | d. if (try_tbl is 1)
|   |   1. if (bnst is not empty)
|   |     then FUNCTION fu_execute_bnst
|   |     else if (nst is not empty)
|   |       then FUNCTION fu_apply_nst
|   |     else
|   |       FUNCTION fu_execute_tbl
|   |   else (try_tbl is 1 and tried_tbl is 1)
|   |     1. FUNCTION fu_look_fwd
|   |     2. FUNCTION fu_apply_nfu
|   | e. adjust dno_head, dno_tail
|   END
| 3. FUNCTION fu_chk_ready
END

```

Analyse von DDC-Notationen (24)



R4: Erstellung der Analysekomponente *vc_day* (9)

Algorithmus zur Analyse von DDC-Notationen (3)

```
FUNCTION fu_execute_phase2
| 1. determine fact_rule_type
| 2. if (dno_atom_pri_last is not empty)
|     then FUNCTION fu_apply_rule(dno_atom_pri_last)
| else if (dno_head_pri_last is not empty)
|     then FUNCTION fu_apply_rule(dno_head_pri_last)
| else if (dno_head_new is not empty)
|     then FUNCTION fu_apply_rule(dno_head_new)
| else if (try_tbl is 0 && is_hydra is 1)
|     then FUNCTION fu_apply_rule(npreno dno_hy_pri)
| else if (try_tbl is 0 && is_hydra is 0)
|     then FUNCTION fu_apply_rule(dno_head)
END
```

Analyse von DDC-Notationen (25)



R4: Erstellung der Analysekomponente *vc_day* (10)

Algorithmus zur Analyse von DDC-Notationen (4)

```

FUNCTION fu_apply_rule
| 1. determine fact_rule_type
| 2. if (nnotations matches "+:")           then FUNCTION fu_ref_to_add_tb
| 3. if (dno_nfu != "")                   then FUNCTION fu_look_fwd
| 4. determine dno_head, dno_tail
| 5. if (fact_rule_type matches "<hat" and matches not "<hatien>")
|                                           then FUNCTION fu_get_dno_atoms
|
| 6. FUNCTION fu_get_span
| 7. if (descr matches "<nfa")             then FUNCTION fu_apply_nfa
|    else if (descr matches "<ba")        then FUNCTION fu_apply_ba
|    else if (descr matches "<na")        then FUNCTION fu_apply_na
|    else if (descr matches "<nsi")       then FUNCTION fu_apply_nsi
|    else if (descr matches "<nfu")       then FUNCTION fu_apply_nfu
|    else                                  FUNCTION fu_apply_tbl
| 8. if (try_nad is 1)                     then FUNCTION fu_execute_nad
| 9. if (pri_line matches "T5--.*0.*>")   then FUNCTION fu_execute_nad
END

```


Analyse von DDC-Notationen (26)



Liu1993 (pp. 99-100)

„720.7073 has been decomposed as follows:

720: Architecture

0707: Geographical treatment

73: United States

The title of this book is:

#aVoices in architectural education : #bcultural politics and

The subject headings for this book are:

#aArchitecture #xStudy and teaching #zUnited States.

#aArchitecture and state #zUnited States.“

Reiner2006

720.7073 <ul-liu/liu_37_to_analyze; length: 8>

7----- Arts & recreation <hatzen>

72----- Architecture <hatzen>

720----- Architecture <hat>

--0.7--- Education, research, related topics <T1--07>

--0.707- Geographic treatment <T1--0707>

---.--7- North America <na4r7span:T1--0701-T1--0709:T2--7>

---.--73 United States <na4r7span:T1--0701-T1--0709:T2--73>

Analyse von DDC-Notationen (27)



```

780.710749 <liu_127_to_analyze; length: 10>
7----- Arts & recreation <hatzen>
78----- Music <hatzen>
780----- Music <hat>
780.7----- Education, research, performances, related
              topics <hat>
780.71---- Music--education <hatien>
--0.7----- Education, research, related topics <T1--07>
--0.71---- Education <T1--071>
--0.7107-- Geographic treatment <T1--07107>
---.---7-- North America <na4r7span:T1--07101-T1-07109:T2--7>(1)
---.---74- Northeastern United States (New England and
              Middle Atlantic states)
              <na4r7span:T1--07101-T1--07109:T2--74>
---.---749 New Jersey <na4r7span:T1--07101-T1--07109:T2--749>

```

(1) <na4r7span: Add to base number T1--0710 notation T2--1-T2--9>

zum Vergleich Liu1993: **780.7, (780), 07107, 749**

Ergebnisse (1)



1. Semantische Analyse von DDC-Notationen in einzelne Atome, nicht nur (grobe) Zerlegung (wie bei Liu1993, Riesthuis1997)

decompose

„to divide sth into smaller parts“ (1)

analyse

„to examine the nature or structure of sth, especially by separating it into its parts, in order to understand or explain it“ (1)

2. Zweiphasiger Analyseprozess von links nach rechts (Liu1993: von rechts nach links):

1. Phase: Ermittlung der Fakten (Haupttafeln)
2. Phase: Ermittlung der durch Regeln gewonnenen Fakten (aus Haupt- und Hilfstafeln und internen Tafeln)

[(1) Oxford Advanced Learner's Dictionary. Online: <http://www.cornelsen.de/oald/1.c.115282.de>]



3. Erfolgreiche Reproduktion der Ergebnisse von Liu1993

- alle 600 DDC-Notationen der Stichprobe von Liu1993 analysiert (nach ev. notwendiger Korrektur von DDC-Notationen wegen fehlender oder überflüssiger Ziffern, Tippfehlern, ...)
- Unterschiede durch Verwendung unterschiedlicher DDC-Ausgaben (Liu1993: 20th Ed. of DDC, Reiner2005: 22nd Ed. of DDC)
- von 600 DDC-Notationen 27 DDC-Notationen zusätzlich analysiert, (wegen Beschränkung auf Hauptklasse 700 von DND nicht zerlegbar)

4a. Ergebnisse der Analyse aller GVK-DDC-Notationen (gesamt)

466.134 analysierte DDC-Notationen

297.782 total analysierte DDC-Notationen (63.9 %)

168.352 partiell analysierte DDC-Notationen (36.1 %)



4b. Ergebnisse der Analyse aller GVK-DDC-Notationen (einzeln)

```

res_gvk_all: 466134
res_gvk0_all: 23946 analyzed dno(s) | 14887 (62.169%) totally complete analyzed dno(s)
res_gvk1_all: 4596 analyzed dno(s) | 4095 (89.0992%) totally complete analyzed dno(s)
res_gvk2_all: 20888 analyzed dno(s) | 16122 (77.1831%) totally complete analyzed dno(s)
res_gvk3_all:189246 analyzed dno(s) |119378 (63.0809%) totally complete analyzed dno(s)
res_gvk4_all: 6680 analyzed dno(s) | 3124 (46.7665%) totally complete analyzed dno(s)
res_gvk5_all: 33649 analyzed dno(s) | 27706 (82.3383%) totally complete analyzed dno(s)
res_gvk6_all: 51704 analyzed dno(s) | 44879 (86.7988%) totally complete analyzed dno(s)
res_gvk7_all: 52632 analyzed dno(s) | 43378 (82.4175%) totally complete analyzed dno(s)
res_gvk8_all: 20678 analyzed dno(s) | 4205 (20.3356%) totally complete analyzed dno(s)
res_gvk9_all 62115 analyzed dno(s) | 21036 (33.8662%) totally complete analyzed dno(s)

```



5. Analyse von DDC-Notationen der DDB-Datei "ddbt2"

285 analysierte DDC-Notationen

215 total analysierte DDC-Notationen (75.4 %)

70 partiell analysierte DDC-Notationen (24.6 %)

Ergebnisausgabe u.a. im MAB2-Format

6. Analyse von DDC-Notationen der DDB-Datei "0605all.ma2"

2.659 analysierte DDC-Notationen

2.210 total analysierte DDC-Notationen (83 %)

449 partiell analysierte DDC-Notationen (17 %)

Dauer: 105 sec.

7. Unterschiedliche Ausgabeformate

7a. Ausgabe als Analysediagramm: *vc_daygram*

7b. Ausgabe als Menge: *vc_dayset*

Ergebnisse (5)



7a1. Ausgabeformat *vc_daygram*

- zu analysierende DDC-Notation (molekulare DDC-Notation)
- Name und Länge der zu analysierenden DDC-Notation in 1. Zeile
<liu_127_to_analyze ; length: 10>
- Dewey-Punkt an 4. Stelle der DDC-Notation
- Reihenfolge und Stellung der Ziffern (innerhalb) der DDC-Notation
- relevante DDC-Notationsteile pro Analyseschritt (atomare DDC-Notation),
- Kennzeichnung der im Analyseschritt irrelevanten Stellen mit „-“,
- Bedeutung der DDC-Notationsteile (Klassenbenennung) pro Analyseschritt,
- angewendete Regeln mit ermittelten DDC-Notationen

=> *vc_daygram* enthält Analyse- und Syntheseinformation



7a2. Ausgabeformat *vc_daygram*

```
-----
vc_daygram: 025.6209746021 <dno_gvk_all_18388_to_analyze>
-----
```

```
025.6209746021 <dno_gvk_all_18388_to_analyze; length: 14>
0----- Computer science, information & general works <hatzen>
02----- Library & information sciences <hatzen>
025----- Operations of libraries, archives, information centers <hat>
025.6----- Circulation services <hat>
025.62----- Interlibrary loans <hat>
---.--09----- Historical, geographic, persons treatment <T1--09>
---.--097----- Treatment by specific continents, countries, localities;
                  extraterrestrial worlds <T1--097>
---.-----7----- North America <ba4r2span:T1--093-T1--099:T2--7>
---.-----74----- Northeastern United States (New England and Middle Atlantic states)
                  <ba4r2span:T1--093-T1--099:T2--74>
---.-----746---- Connecticut <ba4r2span:T1--093-T1--099:T2--746>
---.-----02- Statistics, illustrations; conservation, preservation, restoration
                  <ba4r2span:T1--093-T1--099+02>
---.-----021 Statistics <ba4r2span:T1--093-T1--099+021>
```


Ergebnisse (7)



7b1. Ausgabeformate *vc_dayset*

vc_dayset_dno_cap: 025.6209746021 <dno_gvk_all_18388_to_analyze>

0;Computer science, information & general works
02;Library & information sciences
025;Operations of libraries, archives, information centers
025.6;Circulation services
025.62;Interlibrary loans
T1--09;Historical, geographic, persons treatment
T1--097;Treatment by specific continents, countries, localities; extraterrestrial worlds
T2--7;North America
T2--74;Northeastern United States (New England and Middle Atlantic states)
T2--746;Connecticut
T1--097+02;Statistics, illustrations; conservation, preservation, restoration
T1--097+021;Statistics

vc_dayset_fine: 025.6209746021 <dno_gvk_all_18388_to_analyze>

dno_gvk_all_18388:025.6209746021;0;02;025;025.6;025.62;T1--09;T1--097;T2--7;
T2--74;T2--746;T1--097+02;T1--097+021

vc_dayset_mab2: 025.6209746021 <dno_gvk_all_18388_to_analyze>

705a^_a025.6209746021^_p025.62^_cT1--09^_f097^_g746^_tT1--097+021

Anwendungen (angelehnt an Liu1993)



- * **Information Retrieval und automatische Indexierung:**
Verbesserung der Retrievalergebnisse (precision, recall) durch beliebige Kombination (facets) der atomaren DDC-Notationen als Suchbegriffe
 - * **Entwurf und Entwicklung von Expertensystemen**
(Semi-)automatische Analyse und Synthese von DDC-Notationen
 - * **Automatische Indexierung:**
Bereicherung der Erschliessung bibliographischer Datensätze (catalog enrichment) durch automatische Hinzufügung von atomaren DDC-Notationen
 - * **Sprachkonvertierung (switching language development[1]):**
Durch die in DDC-Notationen vorliegende künstliche Zwischensprache besteht – unter Nutzung des in ca. 30 Sprachen vorliegenden DDC-Systems - die Möglichkeit, multilinguale Recherchen durchzuführen
 - * **DDC-Lehrsystem:**
Mithilfe der Erklärungskomponente eines Expertensystems kann die Synthese von DDC-Notationen gelernt werden
 - * **Qualitätskontrolle der DDC-Notationsvergabe:**
Zur Unterstützung der intellektuellen Klassifizierung von DDC-Notationen und zur Prüfung der Korrektheit vergebener DDC-Notationen.
- [1] [Liu1993] pp. 80, Fußnote 25: „A switching language is a device that can be used to convert from one vocabulary to another.“ 31. Jahrestagung der GfKI (ul, 8. März 2007)

DDC-Suchsystem (1)

DDC-Suchsystem *vc_ds*

DDC-basierte
Anfragen →

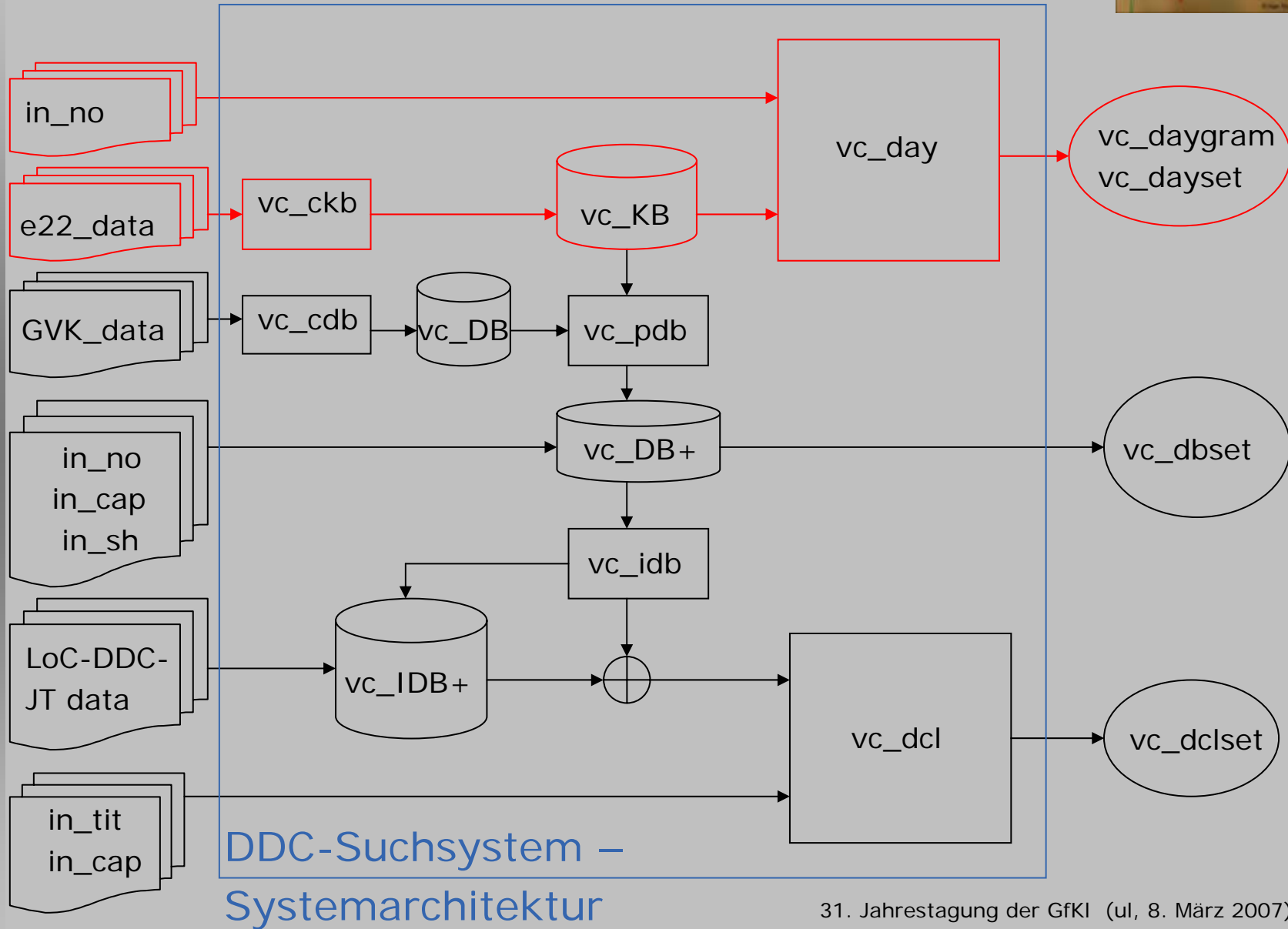


→ DDC-basierte
Antworten

vc_ds: *vzg colibri_dewey search system*



DDC-Suchsystem (2)



Colibri-Weiterflug



- ◆ Erhöhung des Analysegrades (-> 100 %)
- ◆ Deutsche Analyse-Ausgaben (vc_daygram_de):

```

025.1740973 <dno_gvk_all_17389_analysiert; Länge: 11>
0----- Informatik, Informationswissenschaft, allgemeine Werke
          <hatzen>
02----- Bibliotheks- und Informationswissenschaften <hatzen>
025----- Tätigkeiten in Bibliotheken, Archiven, Informationszentren
          <hat>
025.1----- Verwaltung <hat>
025.17----- Verwaltung von Sonderbeständen <hat>
025.174---- CD-ROMs-Behandlung in Bibliotheken, . . . <hatien>
---.---4---- Elektronische Ressourcen <na1r1:025.344>
---.---09-- Historische, geografische, personenbezogene Behandlung
          <T1--09>
---.---097- Behandlung nach einzelnen Kontinenten, Ländern, Ortschaften;
          außerirdischen Welten <T1--097>
---.---7- Nordamerika <ba4r2span:T1--093-T1--099:T2--7>
---.---73 USA (Vereinigte Staaten von Amerika)
          <ba4r2span:T1--093-T1--099:T2--73>

```