Computer-aided Assignment of DDC Numbers
(Computergestützte Zuweisung von DDC-Notationen)
July 17, 2008

Ulrike Reiner
Verbundzentrale des GBV (VZG), Göttingen
Contents

Preface

Assignment of DDC* numbers to
- Bibliographic title records
- Non-DDC terms
- Atomic/molecular DDC numbers

Computer aid: DDC Search System
- State of the Art
- Continuation

[* DDC: Dewey Decimal Classification *]
Preface (1)

AG BIB:
Subject Indexing 2008: Accept Progress!

“14.00 Uhr Werkstattbericht: DDC automatisieren
Dr. Ulrike Reiner, Göttingen (angefragt)”

[http://www.ub.uni-dortmund.de/listen/inetbib/msg36527.html]

Computer-aided Assignment of DDC Numbers

[http://www.ub.uni-dortmund.de/listen/inetbib/msg36838.html]

VZG Project Colibri/DDC (Research and Development)

025.4310285 Dewey Decimal Classification--
Data processing Computer applications

025.431072 Dewey Decimal Classification--
Research; statistical methods
Preface (3)

“World Library and Information Congress: 74th IFLA* General Conference and Council

Libraries without borders: Navigating towards global understanding
10-14 August 2008, Québec, Canada”

[ * International Federation of Library Associations and Institutions,
http://www.ifla.org/IV/ifla74/index.htm,
http://www.ifla.org/IV/ifla74/2008ifla_logo.jpg ]
Montréal, QC, Canada (ul, 15 May 2008)
**Unconventional Computation**


Series: Lecture Notes in Computer Science, Vol. 4618
Sublibrary: Theoretical Computer Science and General Issues

Akl, S.G.; Calude, C.S.; Dinneen, M.J.; Rozenberg, G.; Wareham, H.T. (eds.)
2007, X, 243 p. with online files/update, softcover
ISBN: 978-3-540-73553-3

**Keywords:**
DNA computing, algorithms, ant colony optimization, approximation, authentication, biomolecular computing, cellular automata, chaos, computational models, computing theory, genetic algorithms, heuristic algorithms, local search, mebrane computing, meta-algorithmics, natural computing, neural network, optimization, quantum computing, quantum key distribution, theoretical computer science, theoretical informatics, theory of computation

Bibliographic Title Record of the UC 2007 (1) (MAB2 format)

001 984632514
002a20070614
...
026 DNB984632514 ← Identification number
...
037beng
...
070 1145
070aDNB
070b9999
100bAkl, Selim G. ¹[Hrsg.]¹
102a112655688
200bUC <6, 2007, Kingston, Ontario>
202a6517086-6
331 Unconventional computation
335 6th international conference ; proceedings
410 Berlin ; Heidelberg ; New York
412 Springer
425 2007
425a2007
433 X, 241 S.
...
451 Lecture notes in computer science ; Vol. 4618
...
Verbundzentrale des GBV (VZG) 32nd Annual Conference of the German Classification Society (ul, July 17, 2008, p. 9)
### Bibliographic Title Record of the UC 2007 (2) (MAB2 format)

...  

<table>
<thead>
<tr>
<th>501</th>
<th>Literaturangaben</th>
</tr>
</thead>
<tbody>
<tr>
<td>540a</td>
<td>ISBN 978-3-540-73553-3 kart. : EUR 48.15 (freier Pr.), ...</td>
</tr>
<tr>
<td>540a</td>
<td>ISBN 3-540-73553-4 kart. : EUR 48.15 (freier Pr.), ...</td>
</tr>
<tr>
<td>551a</td>
<td>12088782</td>
</tr>
<tr>
<td>553a</td>
<td>9783540735533</td>
</tr>
<tr>
<td>568</td>
<td>07, N28, 0093</td>
</tr>
<tr>
<td>574</td>
<td>07, A48, 0090</td>
</tr>
<tr>
<td>700</td>
<td>004 İDNB</td>
</tr>
<tr>
<td>705a</td>
<td>a006.3 c006.3 eDDC22ger</td>
</tr>
<tr>
<td>705a</td>
<td>a004.0151 c004.0151 eDDC22ger</td>
</tr>
<tr>
<td>902s</td>
<td>4196735-5 Theoretische Informatik</td>
</tr>
<tr>
<td>902f11</td>
<td>Kongress</td>
</tr>
<tr>
<td>902g11</td>
<td>Kingston &lt;Ontario, 2007&gt;</td>
</tr>
<tr>
<td>907s</td>
<td>4455833-8 Soft Computing</td>
</tr>
<tr>
<td>907f11</td>
<td>Kongress</td>
</tr>
<tr>
<td>907g11</td>
<td>Kingston &lt;Ontario, 2007&gt;</td>
</tr>
<tr>
<td>912s</td>
<td>4611085-9 Bioinformatik</td>
</tr>
<tr>
<td>912f11</td>
<td>Kongress</td>
</tr>
<tr>
<td>912g11</td>
<td>Kingston &lt;Ontario, 2007&gt;</td>
</tr>
<tr>
<td>917s</td>
<td>4190671-8 Zellulärer Automat</td>
</tr>
<tr>
<td>917f11</td>
<td>Kongress</td>
</tr>
<tr>
<td>917g11</td>
<td>Kingston &lt;Ontario, 2007&gt;</td>
</tr>
</tbody>
</table>

**Notation of a classification system**

DDC (Dewey Decimal Classification) analytical (a = Full edition)

Chain link of 1st subject heading chain (s = topical heading, f = form heading, g = geographical / ethnographical heading)

Chain link of 2nd subject heading chain

Chain link of 3rd subject heading chain

Chain link of 4th subject heading chain

ISBN formally (technically) valid

Verbundzentrale des GBV (VZG) 32nd Annual Conference of the German Classification Society (ul, July 17, 2008, p. 10)
Example of

- Bibliographic title record

- Non-DDC terms (notations, numbers or subject headings of other non-DDC classification systems)
  - Atomic/molecular DDC numbers
<table>
<thead>
<tr>
<th>Non-DDC Terms</th>
<th>in PICA+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basisklassifikation (BK)</td>
<td>[045Q]</td>
</tr>
<tr>
<td>Theoretische Informatik (54.10)</td>
<td></td>
</tr>
<tr>
<td>Kettenglied einer RSWK-Kette</td>
<td>[041A]</td>
</tr>
<tr>
<td>Berechnung</td>
<td></td>
</tr>
<tr>
<td>Einzelschlagwörter</td>
<td>[044K]</td>
</tr>
<tr>
<td>Zellularer Automat</td>
<td></td>
</tr>
<tr>
<td>British Library Subject Headings (BLSH)</td>
<td>[044G]</td>
</tr>
<tr>
<td>Heuristic programming</td>
<td></td>
</tr>
<tr>
<td>Library of Congress Subject Headings (LCSH)</td>
<td>[044A]</td>
</tr>
<tr>
<td>Heuristic programming</td>
<td></td>
</tr>
<tr>
<td>Library of Congress Classification (LCC)</td>
<td>[045A]</td>
</tr>
<tr>
<td>QA76.9 (Computer programming)</td>
<td></td>
</tr>
</tbody>
</table>
Objects

Example of

- Bibliographic title record
- Non-DDC terms (notations, numbers, or subject headings of other non-DDC classification systems)
- Atomic/molecular DDC numbers
Atomic/Molecular DDC Numbers

An **atomic DDC number** (dno_atom) is a semantically indecomposable string (of symbols) that represents a DDC class.

\[
511.36 \text{ (Proof theory and constructive mathematics)}
\]

A **molecular DDC number** (dno_mol) is a string that is syntactically decomposable into atomic DDC numbers.

\[
511.36028563 \text{ (Automatic theorem proving)}
\]

[ 511.36 (Proof theory ... mathematics),
T1--028563 (Artificial intelligence) ]
Computer-aided Assignment of DDC Numbers

- bibliographic title records
- non-DDC terms
- atomic/molecular DDC numbers

DDC Search System

DDC number(s)
Assignment of DDC Numbers (1) with the aid of the DDC Search System

```
vc_dcl : vzg colibri_ddc classifier
vc_dqa : vzg colibri_ddc question answerer
vc_day : vzg colibri_ddc number analyzer
vc_dsy : vzg colibri_ddc number synthesizer
```

```
\{005.1[1], 006.3[1]\}
\{006[4]\}
\{004.35, 005.1, 006.31, 006.332\}
\{500, 510, 511, 511.3, 511.36, 511.36028563, T1--02, T1--028, T1--0285, T1--028563\}
\{511.36028563\}
```

---

Verbundzentrale des GBV (VZG)  32nd Annual Conference of the German Classification Society (ul, July 17, 2008, p. 16)
Assignment of DDC Numbers (2)

with the aid of the

DDC Search System

vc_ds

1. vc_dsy
   vzg colibri_ddc number synthesizer

2. vc_day
   vzg colibri_ddc number analyzer

3. vc_dqa
   vzg colibri_ddc question answerer

4. vc_dcl
   vzg colibri_ddc classifier

with the aid of the DDC Search System
DDC Notational Synthesis (1)

“The number of devices for synthesis and instructions for their use are so large that no one knows how many million useful DDC numbers can be composed. From an enumerative scheme of limited scope it has grown to be a sophisticated machine for number synthesis.” [1]

1,041,073,100 Potential classes (19th ed. of the DDC) [2]
7,705 Instructions (22nd ed. of the DDC) [3]

“WebDewey has many additional features, although so far there is no provision for any expert system for the automatic synthesis of numbers following add to instructions.” [1]

State of the Art of vc_dsy (1)

vzg colibri_ddc number synthesizer

First considerations on a computer-aided number synthesis

DDC number synthesis with the aid of components of the DDC search system (vc_ds):

vc_dsy: user interface for DDC synthesis (not implemented yet)

vc_dqa: determining of dno_atoms (slides 26 & 27) and dno_mols (slide 28)

vc_day: testing the suitability or correctness of dno_mol(s) (slide 24)
Automatically DDC notational synthesis?

Main Title: Wisdom and compassion = Śes rab dan śnīṅ rje' i rol pa : the sacred art of Tibet / Marylin M. Rhie, Robert A.F. Thurman; photographs by John Bigelow Taylor.

Subjects: Art, Buddhist--China--Tibet. Buddhist art and symbolism--China--Tibet. Art, Tibetan

Note: ... an exhibition organized by the Asian Art Museum of San Francisco in conjunction with Tibet House, New York

=> Common knowledge and DDC knowledge

<table>
<thead>
<tr>
<th>Tibet =&gt;</th>
<th>dno_atoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Historical, geographic, persons treatment</td>
<td>T1--09</td>
</tr>
<tr>
<td>Tibet Autonomous Region (Xizang Zizhiqu)</td>
<td>T2--515</td>
</tr>
</tbody>
</table>

| exhibition => |
|-------------------|-----------|
| Museums, collections, exhibits | T1--074 |
| San Francisco, New York => |
| United States | T2--73 |

| the sacred art of Tibet ... photographs => |
|-------------------|-----------|
| Arts | 700 |
| Iconography | 704.9 |
| Other religions | 704.9489 |

| Tibetan => |
|-------------------|-----------|
| Tibetan Buddhism (Lamaism) | 294.3923 |

[ red: terms of the bibliographic title record; turquoise: terms of the DDC System ]
State of the Art of vc_dsy (4)

vzg colibri_ddc number synthesizer

dno_atoms := set of atomic DDC numbers
dno_mols  := set of molecular DDC numbers

=> dno_atoms: \{T1--09, T2--515, T1--074, T2--73, 704.9489, 294.3923\}

Tibet

exhibition

vc_dqa

...dno_atoms_without_prefixes

T2--515

T1--074

704.9489439230951507473

vzg colibri_ddc number synthesizer
State of the Art of vc_day (1)
vzg colibri_ddc number analyzer

Analysis Diagram of a DDC number (in_liu_417)

704.9489439230951507473 <liu_417_to_analyze; length: 23>
7------------------------ Arts & recreation <hatzen>
70------------------------ Arts <hatzen>
704---------------------- Special topics in fine and decorative arts <hat>
704.9------------------- Iconography <hat>
704.94------------------ Specific subjects <hat>
704.948---------------- Religion <hat>
704.9489---------- Other religions <hat>
704.94894-------- Indic religions--art representation <hatien>
704.948943--------- Buddhism--art representation <hatien>
---.---4--------- Religions of Indic origin <na1r1:294>
---.---43-------- Buddhism <na1r1:294.3>
---.---439-------- Branches, sects, reform movements <na1r1:294.39>
---.---4392------- Mahayana Buddhism (Northern Buddhism) <na1r1:294.392>
---.---43923------ Tibetan Buddhism (Lamaism) <na1r1:294.3923>
---.-------09------ Historical, geographic, persons treatment <T1--09>
---.-------095----- Treatment by specific continents, countries, localities; extraterrestrial worlds <T1--095>
---.-------5-------- Asia OrientFar East <ba4r2span:T1--093-T1--099:T2--5>
---.-------51------- China and adjacent areas <ba4r2span:T1--093-T1--099:T2--51>
---.-------515------ Tibet Autonomous Region (Xizang Zizhiqu) 
<ba4r2span:T1--093-T1--099:T2--515>
---.-------07------- Museums, collections, exhibits; collecting objects <ba4r2span:T1--093-T1--099+07>
---.-------074------ Museums, collections, exhibits <ba4r2span:T1--093-T1--099+074>
---.-------7-------- North America <ba4r2span:T1--093-T1--099+074:na4r2:T2--7>
---.-------73------- United States <ba4r2span:T1--093-T1--099+074:na4r2:T2--73>
State of the Art of vc_day (2)
vzg colibri_ddc number analyzer

Presentation
GfKI 2007, Librarian Workshop, Freiburg

Publication
Automatic Analysis of Dewey Decimal Classification Notations
[ http://www.springerlink.com/content/l044082243v1l7u6/, pp. 697-704 ]
State of the Art of vc_dqa (1)
vzg colibri_ddc question answerer

Searching for dno_atoms (1)

```sql
mysql> select * from dno_kb where (dno like "t%" and descr_val like "%tibet%")
+
| dno      | descr | descr_val                                  |
|----------+-------+----------------------------------------------------------------|
| t2--515 | <hat> | tibet autonomous region (xizang zizhiqu) |
| t5--954  | <hat> | tibetans                                     |
| t6--9541 | <hat> | tibetan                                      |
| t6--954  | <hat> | tibeto-burman languages                      |
| t6--95   | <hat> | languages of east and southeast asia sino-tibetan languages |
+
5 rows in set (0.07 sec)
```

```sql
mysql> select * from dno_kb where (dno like "t%" and descr_val like "%exhibit%")
+
| dno                   | descr   | descr_val                   |
|-----------------------+---------+-------------------------------------------------------------------------|
| t1--074               | <hat>   | museums, collections, exhibits                                          |
| t1--0753-t1--0755     | <hat>   | [organizing and preparing collections and exhibits, service to patrons] |
| t1--08+074            | <ba5>   | museums, collections, exhibits                                          |
| t1--0901-t1--0905+074 | <ba5>   | museums, collections, exhibits                                          |
| t1--0901-t1--0905+07  | <ba5>   | museums, collections, exhibits                                          |
| t1--093-t1--099+074   | <ba4r2> | museums, collections, exhibits                                          |
| t1--093-t1--099+07    | <ba4r2> | museums, collections, exhibits                                          |
+
7 rows in set (0.07 sec)
```

[ for dno_kb (DDC knowledge base), <hat>, <ba5>, ... see p. 29 and p. 32
State of the Art of vc_dqa (2)
vzg colibri_ddc question answerer

Searching for dno_atoms (2)

dno_atoms : \{T1--09, T2--515, T1--074, T2--73, 704.9489, 294.3923\}  (slide 22-24 & 26)

```sql
mysql> select * from dno_kb where dno="704.9489";
+----------+---------+---------------------------+
| dno      | descr   | descr_val                 |
|----------+---------+---------------------------|
| 704.9489 | <hat>   | other religions           |
| 704.9489 | <na1r1> | |704.9489|292-299|29||||| |
| 704.9489 | <ri>    | hell                      |
+----------+---------+---------------------------+
3 rows in set (0.07 sec)
```

"Add to base number 704.9489 the numbers following 29 in 292-299"

cross-check by

```sql
mysql> select * from dno_kb where dno="294.3923";
=> no rules!
```

[ for the knowledge representation of DDC facts and rules, see p. 32
State of the Art of vc_dqa (3)

vzg colibri_ddc question answerer

Searching for dno_mols*

```sql
mysql> select * from dno_db where (dno like "%73%" and dno like "%515%" and dno like "%43923%" and dno like "704.9489%" and dno like "%09%" and dno like "%074%");
+-------------------------+------------+-------------------------------------+
| dno                     | descr      | descr_val              |
+-------------------------+------------+-------------------------------------+| 704.9489439230951507473 | <001A>     | 02003                  |
| 704.9489439230951507473 | <003@>     | 0113549423             |
| 704.9489439230951507473 | <021A>     | compassion             |
| 704.9489439230951507473 | <021A>     | wisdom                 |
| 704.9489439230951507473 | <022A>     | france                 |
| 704.9489439230951507473 | <022A>     | south                  |
| 704.9489439230951507473 | <028A_da>  | marylin m.#rnie        |
| 704.9489439230951507473 | <028C>     | john bigelow#taylor    |
| 704.9489439230951507473 | <028C>     | robert a. f.#thurman   |
| 704.9489439230951507473 | <033A>     | <033A>-royal academy of arts@london |
| 704.9489439230951507473 | <044A_a>   | art, buddhist          |
| 704.9489439230951507473 | <044A_a>   | art, tibetan           |
| 704.9489439230951507473 | <044A_a>   | buddhist art and symbolism |
| 704.9489439230951507473 | <044A_z>   | china                  |
| 704.9489439230951507473 | <044A_z>   | tibet                  |
| 704.9489439230951507473 | <045A>     | N8193.T5               |
| 704.9489439230951507473 | <145Z_a>   | 1h 65990               |
| 704.9489439230951507473 | liu_417_to | length: 23             |
+-------------------------+------------+-------------------------------------+
18 rows in set (0.02 sec)
```

* unique dno_mols in GVK: 590,120 (January 2008), 466,134 (July 2004); approx. 3,000 dno_mols/month

Verbundzentrale des GBV (VZG) 32nd Annual Conference of the German Classification Society (ul, July 17, 2008, p. 28)
#### State of the Art of vc_dqa (4)

**vzg colibri_ddc question answerer**

**Searching for dno* with Non-DDC terms**

```sql
mysql> select * from dno_db where descr like "<044A%" and descr_val like "genetic programming%computer science";
```

```
+------------+----------+--------------------------------------+
<table>
<thead>
<tr>
<th>dno</th>
<th>descr</th>
<th>descr_val</th>
</tr>
</thead>
<tbody>
<tr>
<td>004.1</td>
<td>&lt;044A_s&gt;</td>
<td>genetic programming computer science</td>
</tr>
<tr>
<td>004.35</td>
<td>&lt;044A_a&gt;</td>
<td>genetic programming computer science</td>
</tr>
<tr>
<td>005.1</td>
<td>&lt;044A_a&gt;</td>
<td>genetic programming computer science</td>
</tr>
<tr>
<td>006.31</td>
<td>&lt;044A_a&gt;</td>
<td>genetic programming computer science</td>
</tr>
<tr>
<td>006.31</td>
<td>&lt;044A_s&gt;</td>
<td>genetic programming computer science</td>
</tr>
<tr>
<td>006.332</td>
<td>&lt;044A_s&gt;</td>
<td>genetic programming computer science</td>
</tr>
<tr>
<td>006.3</td>
<td>&lt;044A_a&gt;</td>
<td>genetic programming computer science</td>
</tr>
<tr>
<td>006.3</td>
<td>&lt;044A_s&gt;</td>
<td>genetic programming computer science</td>
</tr>
<tr>
<td>332.015118</td>
<td>&lt;044A_s&gt;</td>
<td>genetic programming computer science</td>
</tr>
<tr>
<td>621.381</td>
<td>&lt;044A_s&gt;</td>
<td>genetic programming computer science</td>
</tr>
</tbody>
</table>
+------------+----------+--------------------------------------+
10 rows in set (0.00 sec)
```

**LCSH (Library of Congress Subject Headings)**

```sql
mysql> select * from dno_db where descr like "<045A%" and descr_val="qa76.623";
```

```
+---------+--------+-----------+
<table>
<thead>
<tr>
<th>dno</th>
<th>descr</th>
<th>descr_val</th>
</tr>
</thead>
<tbody>
<tr>
<td>004.35</td>
<td>&lt;045A&gt;</td>
<td>QA76.623</td>
</tr>
<tr>
<td>005.1</td>
<td>&lt;045A&gt;</td>
<td>QA76.623</td>
</tr>
<tr>
<td>006.31</td>
<td>&lt;045A&gt;</td>
<td>QA76.623</td>
</tr>
<tr>
<td>006.332</td>
<td>&lt;045A&gt;</td>
<td>QA76.623</td>
</tr>
</tbody>
</table>
+---------+--------+-----------+
4 rows in set (0.00 sec)
```

**LCC (Library of Congress Classification)**

* dno: DDC number
State of the Art of vc_dcl (1)
vzg colibri_ddc number classifier

Automatic classification

- DDC database \( \text{vc\_DB} \)
- DDC knowledge base \( \text{vc\_KB} \) \{ \( \text{vc\_DB\_PLUS} \) (intellectual basis) \}

- vector product* as similarity measure

\[
S_{uc} = \sum_{i=1}^{l} u_i c_i
\]

between the terms of

DDC-unclassified title records \( u \) (elements of in_dnb_ABH**) and DDC-classified title records \( c \) (elements of vc_DB_PLUS)

- two heuristic functions: \text{cutoff\_val\_dyn}, \text{cutoff\_val\_stat}
- two methods to calculate DDC class candidates: \text{calc1}, \text{calc2}

---


** input test data (25,653 bibliographic title records) from the German National Library (Deutsche Nationalbibliothek DNB) ]
Evaluation of the automatic classification

- automatic evaluation
  two (automatic) correlation measures:
  correlation pattern CP, e.g., 110.xxx xxx xxx xxx
  correlation number CN ∈ [0,1], e.g., 0.666667, 1

- (first) intellectual evaluation of vc_dcl (July 3, 2008)
  - 11 DNB experts evaluated Sample1*
  - “Colibri ist derzeit für eine automatische Klassifizierung von
    Publikationen insbesondere Netzpublikationen nicht einsetzbar
    ... Eine Modifizierung des Colibri-Systems vor einem erneuten
    Testlauf wäre wünschenswert.” **

[ * 1,000 automatic DDC-classified title records of in_dnb_ABH by vc_dcl]  
** Yvonne Jahns; Elisabeth Mödden: Colibri-Test Juni 2008. Automatisches
  Klassifizieren_DDC_Colibri. L2/F2 AG Netzpublikationen. 3. Juli 2008 ]
State of the Art of vc_dcl (3)
**vzg colibri_ddc number classifier**

**Accuracy of vc_dcl (cf. slide 28)**

dl1385-122 colibri/ul-test> vc_dcl_cli < vc_DB/in_ppn_113549423

<table>
<thead>
<tr>
<th>number of ddc-classified title:</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>identifier (dno,schedno):</td>
<td>113549423 (704.948943230951507473,704.948943)</td>
</tr>
<tr>
<td>Data of LoC:</td>
<td>704.948943230951507473</td>
</tr>
<tr>
<td>calculated cutoff value:</td>
<td>40</td>
</tr>
</tbody>
</table>

**title:**
Wisdom and compassion

considered descriptor values:
| 12 | {<028A_da>-marylin m.#rhie[6], <028C>-robert f.#thurman[16], <028C>-joh bigelow#taylor[24], <021A>-wisdom[1830], <021A>-compassion[274], <033A>-royal academy arts@london[47], <044A_a>-tibet[786], <044A_a>-buddhist art symbolism[40], <044A_a>-china[4756], <044A_a>-art tibetan[20], <044A_a>-art buddhist[44], <045A>-N8193.T5[2]} |

matched descriptor values:
| 6 | {N8193.T5, marylin m.#rhie, john bigelow#taylor, art tibetan, robert f.#thurman, buddhist art symbolism} |

max. match value of matched descriptor values: | 6 |

| calculated1 ddc classes (subdiv): | 1 | {704.948943} |
| calculated1 ddc classes (sections): | 1 | {704} |
| calculated1 ddc classes (main):     | 1 | {700} |

| calculated2 ddc classes (subdiv):   | {704.948943[1]} |
| calculated2 ddc classes (sections): | {704[1]} |
| calculated2 ddc classes (divisions):| {700[1]} |
| calculated2 ddc classes (main):     | {700[1]} |

**correlation (113549423,704.948943):**
111.111 111 xxx xxx (1)

**Sorted list of frequencies**

111 [FUCUTVAL] S_list: 2-6-16-20-24-40-44-47-274-786-1830-4756

Verbundzentrale des GBV (VZG) 32nd Annual Conference of the German Classification Society (ul, July 17, 2008, p. 32)
Assignment of DDC numbers to non-DDC terms by vc_dcl

dl385-122 colibri/ul-test> vc_dcl_cli < vc_DB/in_ul_gfk108_FICTIVE

number of ddc-classified title: 1

identifier (dno,schedno): FICTIVE (XXX,X)

Data of ????: XXX

calculated cutoff value: 11

title: Unconventional computation

considered descriptor values: |1|{<331>-genetic programming computer science[11]}

matched descriptor values: |1|{genetic programming computer science}

max. match value of matched descriptor values: |1|

calculated1 ddc classes (subdiv): |11| {004.1, 004.35, 005.1, 006.3,
006.31, 006.32, 006.332, 332, 620.00113, 621.3, 621.381}

calculated1 ddc classes (sections): |3| {004, 005, 006}

calculated1 ddc classes (main): |1| {000}

calculated2 ddc classes (subdiv): {004.1[1], 004.35[1], 005.1[1],
006.31[1], 006.32[1], 006.332[1], 006.3[1]}

calculated2 ddc classes (sections): {006[4]}

calculated2 ddc classes (divisions): {000[7]}

calculated2 ddc classes (main): {000[7]}

correlation (FICTIVE,X): xxx.xxx xxx xxx xxx xxx (X)
Classification of “Unconventional computation” by vc_dcl (1)

number of ddc-classified title: 9131
identifier (dno,schedno): DNB0984632514 (004.0151,004.0151)
DNB DDC notation (MAB2 field 700) {004}
calculated cutoff value: 393
title: Unconventional computation
title (remainder): 6th international conference ; proceedings
title (series): Lecture notes in computer science ; Vol. 4618
considered descriptor values: |18| {<100b>-selim#akl[0], <331>-unconventional[393], <331>-computation[306], <335>-proceedings[12003], <335>-6th[801],<335>-international[12911], <335>-conference[6787], <412@410>-springer@berlin[750], <451>-computer[2601], <451>-science[7828], <451>-vol.[3], <451>-lecture[1875], <451>-notes[4125], <540a>-3-540-73553-4[0], <902s>-theoretische informatik[62], <902g1>-902g11|kingston <ontario 2007>[0], <907s>-soft computing[74], <907g>-907g11|kingston <ontario 2007>[0]}
matched descriptor values: |4| {theoretische informatik, unconventional, soft computing, computation}
max. match value of matched descriptor values: |4|
calculated1 ddc classes (subdiv): |3| {004, 005.1, 006.3, 511.3} calculated1 ddc classes (sections): |3| {004, 005, 006}
calculated1 ddc classes (main): |1| {000}
calculated2 ddc classes (subdiv): {005.1[1], 006.3[1]} calculated2 ddc classes (sections): {004[1], 005[1], 006[1]}
calculated2 ddc classes (divisions):{000[3]} calculated2 ddc classes (main): {000[3]}
correlation (DNB0984632514,004.0151: 111.000 0xx xxx xxx (0.428571)

State of the Art of vc_dcl (5)
vzg colibri_ddc number classifier
**Classification of "Unconventional computation" by vc_dcl (2)**

<table>
<thead>
<tr>
<th>number of ddc-classified title</th>
<th>9506</th>
</tr>
</thead>
<tbody>
<tr>
<td>identifier (dnoschedno)</td>
<td>DNB0984632514 (006.3,006.3)</td>
</tr>
<tr>
<td>DNB DDC notation (MAB2 field 700)</td>
<td>{004}</td>
</tr>
<tr>
<td>DDC notation (MAB2 field 705)</td>
<td>{006.3,004.0151}</td>
</tr>
<tr>
<td>calculated cutoff value</td>
<td>260</td>
</tr>
<tr>
<td>title</td>
<td>Unconventional computation</td>
</tr>
<tr>
<td>title (remainder)</td>
<td>6th international conference; proceedings</td>
</tr>
<tr>
<td>title (series)</td>
<td>Lecture notes in computer science; Vol. 4618</td>
</tr>
<tr>
<td>considered descriptor values</td>
<td>{19}</td>
</tr>
<tr>
<td>matched descriptor values</td>
<td>{5}</td>
</tr>
<tr>
<td>max. match value of matched descriptor values</td>
<td>{4}</td>
</tr>
<tr>
<td>calculated1 ddc classes (subdiv)</td>
<td>{3}</td>
</tr>
<tr>
<td>calculated1 ddc classes (sections)</td>
<td>{2}</td>
</tr>
<tr>
<td>calculated1 ddc classes (main)</td>
<td>{1}</td>
</tr>
<tr>
<td>calculated2 ddc classes (subdiv)</td>
<td>{006.3[1]}</td>
</tr>
<tr>
<td>calculated2 ddc classes (sections)</td>
<td>{004[1], 006[1]}</td>
</tr>
<tr>
<td>calculated2 ddc classes (divisions)</td>
<td>{000[2]}</td>
</tr>
<tr>
<td>calculated2 ddc classes (main)</td>
<td>{000[2]}</td>
</tr>
<tr>
<td>correlation (DNB0984632514,006.3)</td>
<td>111.1xx xxx xxx xxx (1)</td>
</tr>
</tbody>
</table>
**State of the Art of vc_dcl (7)**

vzg colibri_ddc number classifier

**Automatic classification obstacles (1)** (extract of *)

- “A subject may occur in almost any discipline. ... Thus there is theoretically no single class number for any subject.”

- “Determination of the specific subject of a given document is an art which machines cannot do.”

- “The specific subject of a document may be determined by ... title, subtitle, blurb, preface and table of contents, and scanning through the text ... The institutional affiliation of the author, the index at the back of the book, the series and cited references ... published reviews or consult some reference tools ... a subject expert.”

State of the Art of vc_dcl (8)

vzg colibri_ddc number classifier

Automatic classification obstacles (2) (following *)

- “...there is too much ambiguity and complexity in the world of publishing and the DDC...”

- a title could be incomplete, fanciful, or vague
- a title could contain redundant words
- it could be a new subject that has not yet been given a place in the DDC
- obscure subjects, e.g., “travels in transoxiana”
- different terminology, e.g., “German Bundestag”, “US Congress”
- the terminology for some concepts differs even between countries with the same language, e.g., gas station (US) - petrol pump (UK)

Fundy National Parc, NB, Canada (ul, 25 May 2008)

Advancement of the DDC Search System

**DDC analysis:** classes 800 und 900

**DDC classification:** improvement

**DDC question answering:** ready to use, easily expandable

**DDC synthesis:** occasionally

Thank you for your interest in the VZG Project Colibri/DDC!