The National Library Bibliographic Classification (BBK) as a Base for Subject Search in the Integrated RSL Digital Library. The Project presentation.

Lavrenova Olga

the Russian State Library

Moscow

The integrated RSL Digital Library contains about **805 000** titles,

including the Dissertation Digital Library (690 000 full texts).

A high quality of subject search is accessible only provided that a system can take into account semantic relationships between search characteristics (hierarchy, synonymy, association).

The project purpose is the ensuring of these conditions on the base of the existing use-proven classification means.

Full version of the Library Bibliographic Classification BBK:

about 120000 classification numbers,

Classification Scheme code *rubbk* in MARC 21

Some libraries install in their OPACs various machine-readable classification schemes in their initial forms.

But it is not very easy for users to construct a search characteristic if a classification structure is complicated.

Many modern users like to get an outcome quickly and “on default”.

2 variants of technologies to use semantic relationships by information retrieval in the RSL:

- **inclusion interpretations (hierarchical strings of captions) of BBK classification numbers in every bibliographic record**
- **installation in open access an «operational» classification variant related with OPAC as a knowledge model.**

Bibliographic record includes:

- BBK classification numbers
- hierarchical strings of captions for numbers
- free (uncontrolled) key words in addition

The «operational» classification will include:

- all complex classification numbers from all bibliographic records generated during a cataloguing process
- caption hierarchy for every complex classification number
- the relevant amount of bibliographic records in the RSL OPAC for every classification number (caption) indicated by search.

The request:

“The Volga language group of Finno-Ugrian languages”

Class. Number **SH 166.32-211 |2 rubbk**

Title The Derivation of Negative in the Mari language

BBK caption hierarchy

Philological sciences. Belles-lettres -- Linguistics – Languages of the world -- Finno-Ugric languages -- Volga languages group -- Mari languages (Mari, Cheremis) -- Grammar -- Morphology -- Word-formation

E-address

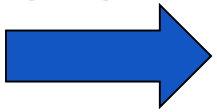
http://dlib.rsl.ru/rsl01002000000/rsl01002732000/rsl01002732180/rsl01002732180.pdf

The classification number structure (9 levels):

SH Philological sciences. Belles-lettres
+ SH1 Linguistics
++ SH16 Languages of the world
+++ SH166 Finno-Ugrian languages
++++ SH166.3 Volga languages group
+++++ SH166.32 Mari languages (Mari, Cheremis)
++++++ SH166.32-2 Grammar
+++++++ SH166.32-21 Morphology
++++++++ SH166.32-211 Word-formation

The aim consists in "decoding" of classification numbers in all remaining records for books = they should get caption hierarchies.

Classification numbers + their captions in the form of catalogue guide cards for data editing by software and experts.



SHCH Art. Study of Art

+SHCH 3 Separate music varieties and musical performance

++ SHCH 31 Music

+++ SHCH 315 Instrumental music

+++ +SHCH 315.3/9 Music instruments. Study of instruments

+++++ SHCH 315.31 Ancient instruments

++++ +SHCH 315.32 Folk instruments

++++ +SHCH 315.4 Keyboard instruments

+++++ +SHCH 315.41 Clavichord. Clavecin. Spinet

+++++ +SHCH 315.42 Piano

BBK classification numbers and there caption hierarchies (caption strings) for

- bibliographic records,
- «operational» classification related with OPAC

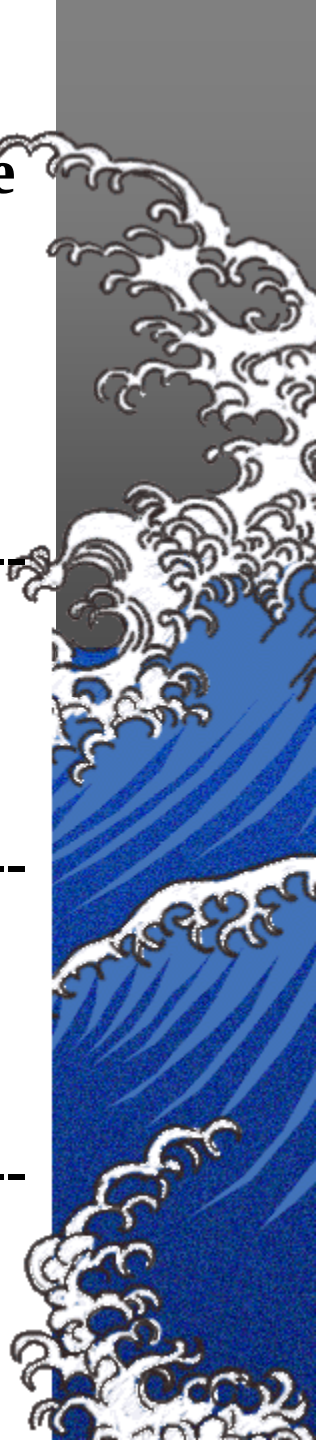
SHCH 315.3/9 Art. Study of Art -- Music -- Separate music varieties and musical performance -- Instrumental music -- Music instruments. Study of instruments

And so on …

SHCH 315.4 Art. Study of Art -- Music -- Separate music varieties and musical performance -- Instrumental music -- Music instruments. Study of instruments -- Keyboard instruments

SHCH 315.41 Art. Study of Art -- Music -- Separate music varieties and musical performance -- Instrumental music -- Music instruments. Study of instruments -- Keyboard instruments – Clavichord. Clavecin. Spinet

SHCH 315.42 Art. Study of Art -- Music -- Separate music varieties and musical performance -- Instrumental music -- Music instruments. Study of instruments -- Keyboard instruments – Piano



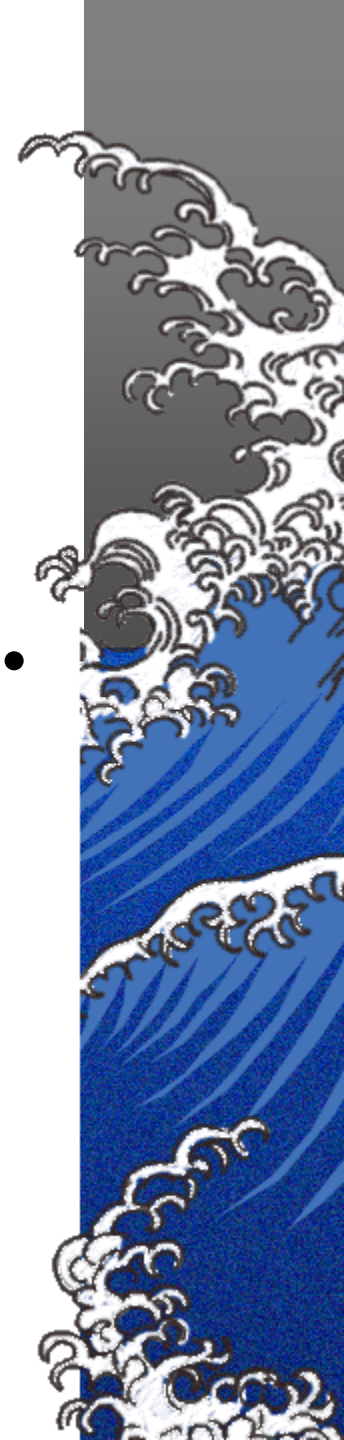
3 new levels from the OPAC for that example:

SHCH 315.42 Art. Study of Art -- Music -- Separate music varieties and musical performance -- Instrumental music -- Music instruments. Study of instruments -- Keyboard instruments – Piano --

Piano music history -- -- 19th century -- Themes and images

Hierarchical strings of captions from the RSL OPAC.

Some examples



084 |a SHCH103(0)320.6l62 |2 rubbk
084 |a T3(0)324.54l62 |2 rubbk
24500|a Art antique de Chypre : ...
260|a [Paris] |b Bibl. nat. de France |c 1994
650 7|a Art. Study of art – Fine art – History of the fine art. Study of art - Ancient fine Art history -- Antique fine art -- Aegean fine art -- Fine art of Cyprus – Exhibitions |2 rubbk
650 7|a History. Historical sciences -- Archaeology -- Antique archaeology -- Archaeology of nations and countries the ancient world – Cyprus -- Exhibitions |2 rubbk

084|a Щ163(2Р=Р)7-8Киселев П.43я61
|2 rubbk
1001|a Kiselev Pavel
24510|a This beautiful and fragile world... Cyprus: [Photograph album]
650 7 |a Art. Study of art -- Art photography – History of the art photography – Russian Federation – Russian art photography – Kiselev Pavel - photographer – Subjects and images -- Album |2 rubbk
653|a Cyprus

084|a Ш143.21-32,0 |2 rubbk
084|a Ш145.3-32,0 |2 rubbk
24500|a The “preposition” category in the Russian, English and Modern Greek |b Dissertation
650 7|a Philological sciences. Belles-lettres -- Linguistics – Languages of the world -- Indo-European languages -- Germanic languages – English -- Lexicology – Semantics (semasiology) |2 rubbk
650 7|a Philological sciences. Belles-lettres -- Linguistics – Languages of the world -- Indo-European languages -- Greek -- Modern Greek -- Lexicology – Semantics (semasiology) |2rubbk
653|a preposition

The embedding of that classification numbers with caption strings into database of the «operational» classification system is a task for the next year.

It is very impotent to allocate the data base in the open access for processing by well-known Internet systems.

the «operational» classification system in the open access

the direct hierarchical search and search on free word combinations

the relevant amount of bibliographic records for every classification number by search

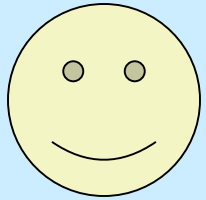
the information retrieval in the OPAC, the DL

A check-out of the technology “decoding classification numbers”

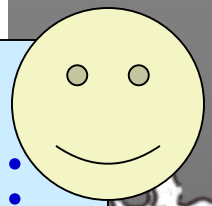
At the input - ***E693.32-739.1,0***
from a bibliographic record.

E693.32-73 **The system take out the string for this part:**

Life sciences-- Zoology -- Animals taxonomy. – Chordata. Chordates – Vertebrata. Vertebrates. Vertebrates zoology
-- Pisces. Fish. Ichthyology –
Physiology, biophysics and biochemistry – Physiology



Examples of full “decoded” numbers:



R410.150.11 (Pericarditis)

SH5(2=P)75-65я44 (Poetry, collection)

E693.363.99 (Dolphins)

V192.18,09 (Mathematics. Iteration method)

Thank you!