

Contents: [Konverzija podataka]

- Konverzija podataka (Konfiguracijske datoteke)
 - ◆ Konverzija podataka (Modify)
 - ◆ Konverzija podataka (Normalize)
 - ◆ Konverzija podataka (Formati)
- Konverzija podataka (Komentari uz pojedina MARC21 polja)
 - ◆ Konverzija podataka (LEADER)
 - ◆ Konverzija podataka (001)
 - ◆ Konverzija podataka (003)
 - ◆ Konverzija podataka (007)
 - ◆ Konverzija podataka (008 - All materials)
 - ◆ Konverzija podataka (008 - Books)
 - ◆ Konverzija podataka (020)
 - ◆ Konverzija podataka (035\$6 ili 035\$9)
 - ◆ Konverzija podataka (040)
 - ◆ Konverzija podataka (041 - indikatori)
 - ◆ Konverzija podataka (041)
 - ◆ Konverzija podataka (044)
 - ◆ Konverzija podataka (080 - UDK)
 - ◆ Konverzija podataka (084 - other classification (R))
 - ◆ Konverzija podataka (100)
 - ◆ Konverzija podataka (111 i 110)
 - ◆ Konverzija podataka (245 indikatori)
 - ◆ Konverzija podataka (245)
 - ◆ Konverzija podataka (246)
 - ◆ Konverzija podataka (250)
 - ◆ Konverzija podataka (260)
 - ◆ Konverzija podataka (300)
 - ◆ Konverzija podataka (362 - brojÄ ani podaci za nakladniÄ ku cjelinu - razjasniti upotrebu)
 - ◆ Konverzija podataka (440 - indikatori)
 - ◆ Konverzija podataka (440)
 - ◆ Konverzija podataka (500)
 - ◆ Konverzija podataka (502 - biljeÅjka o disertacijama)
 - ◆ Konverzija podataka (504)
 - ◆ Konverzija podataka (505)
 - ◆ Konverzija podataka (520)
 - ◆ Konverzija podataka (526 - STUDY PROGRAM INFORMATION NOTE)
 - ◆ Konverzija podataka (538 - zahtjevi sustava)
 - ◆ Konverzija podataka (655)
 - ◆ Konverzija podataka (653)
 - ◆ Konverzija podataka (700)
 - ◆ Konverzija podataka (740 - nadopuniti)
 - ◆ Konverzija podataka (760 w)
 - ◆ Konverzija podataka (852 - signatura)
 - ◆ Konverzija podataka (876 - item information - basic bibliographic unit)
 - ◆ Konverzija podataka (886 - former marc)

Konfiguracijske datoteke

Pravila za konverziju ispisana su u konfiguracijskim datotekama za Webpac2.

Modify

regex mappings

direktorij conf/modify
format:

Omogućava dodatnu pripremu podataka, prije normalizacije.

Koristi se za:

- zamjenu interpunkcije potpoljem i obrnuto
- zamjenu potpolja drugim potpoljem
- dodavanje interpunkcija
- micanje suvišnih znakova

Normalize

Datoteke za normalizaciju nalaze se u direktoriju `conf/normalize/`.
Tu su opisana pravila za konverziju podataka.

Za konverziju FF CDS/ISIS formata u MARC21 koriste se ove datoteke:

- ff-libri.pl
- ff-peri.pl
- ff-arti.pl

nalaze se na adresi: <http://knjiznice.ffzg.hr/isis-konverzija/normalize/> (desni klik, pa Save As, ako želite pogledati datoteku. ako samo kliknete, javit će vam se grozna greška)

Formati

Moguće su različite kombinacije ulaznih i izlaznih formata podataka:

Formati ulaznih podataka:

- CDS/ISIS master file
- MARC (ISO-2709)
- Excel
- DBF
- ISI file format
- Gutenberg file format
- PDF tablice
- ... (potencijalno svaki strukturirani format)

Formati izlaznih podataka

- ISO-2709 (MARC)

- Excel
- JSON
- formati za *search engines*

All variable fields have a maximum length of 9999 characters.

Komentari uz pojedina MARC21 polja

- zapis se konvertira se ako postoji polje 200^a

LEADER

- LDR 05 - Record status
 - ◆ **c** = corrected ; **n** = new - iz 999a ili 999b ili 999 - provjeriti za svaku knjižnicu
 - ◆ 999^aP - SFM
 - ◆ ovo je kandidat za lokalno polje
- LDR 06 - Type of record
 - ◆ **a** = language material - dobivaju ga svi
 - ◆ zbirke drugih tipova (audio snimke i softwer) su male, pa ce se uređivati ručno, nakon koverzije
- LDR 07 - Bibliographic level
 - ◆ **m** = Monograph/item
 - ◆ **s** = Serial - dobivaju ga nakladničke cjeline
- LDR 17 - encoding level
 - ◆ svi dobiju **7** = minimal level
- LDR 18 - Descriptive cataloging form
 - ◆ svi dobivaju **i** = isbd
- LDR 19 - Linked record requirement
 - ◆ dilema: da li oba zapisa trebaju **r** ili samo jedan zapis

001

Voyager

003

- vezan je uz 001
- moglo bi stajati Gradec

-> vidi i polja 035\$9, 035\$a

007

- možda **ta** svima, možda **ne**

008 - All materials

- 008 00 - datum
 - ◆ podatak je u polju 994^c
 - ◆ kad datum nije poznati unjeti se 010101 (2001 je većina knjižnica započela

- s unosom)
 - ◆ ne smiju stajati praznine!
- 008 06 - Type of date/Publication status
 - ◆ **n** - Dates unknown - popunjava se s uuuu ili npr. 19uu
 - ◆ **s** - Single known date/probable date
- 008 07-10 - Date 1
 - ◆ iz 210d
- 008 11-14 - Date 2
 - ◆ to nema za monografije
- 008 15-17 - Place of publication, production, or execution
 - ◆ podatak ne postoji u originalnim bazama. moguće je ga je izvući i iz polja 210^a i tablice s popisom gradova i država
 - ◆ upoznati detalje oko deduplikacije, tj. čemu zapravo služi polje o zemlji izdavanja i da li je bitno ako nismo nacionalni centar
- 008 35-37 - Language
 - ◆ izvlači se iz polja 101

008 - Books

- 008 18-21 - Illustrations
- 008 22 - Target audience
- 008 23 - Form of item
- 008 24-27 - Nature of contents
- 008 28 - Government publication
- 008 29 - Conference publication
- 008 30 - Festschrift
- 008 31 - Index
- 008 32 - Undefined
- 008 33 - Literary form
- 008 34 - Biography

020

- modify za polje 10 -> drugi ISBN počine prefixom "ISBN" koji se miče (pr. u sfb)
- marc-lint javlja krivi checksum

035\$6 ili 035\$9

- Privremeno koristimo 035\$6 umjesto 035\$9, zato što marclint prijavljuje 9 kao grešku.
- marc 035a - System Number - Voyagerov ID (001)
- polje 035\$9 smije sadržavati slova i razmake

040

- za sve je isti
- treba utvrditi točnu oznaku za ustanovu

041 - indikatori

pretpostavke o indikatorima

- Ako postoji polje 300 i ono sadrži /Prijevod/ ili /Izv. stv. nasl./ => i1 = 1
 - ◆ Da li je napomena uvijek pisana?
- Podatak o prevodiocu u 200g. Problem: različiti oblici riječi, različiti jezici
- Ako u UDK oznaci postoji "=", bez zagrada, onda je to prijevod
 - ◆ ako postoje zagrade pr. 94(=411.16) onda su to pomoćne oznake
 - ◆ pr. 821.111-2=163.42 je prijevod engleske drame na HR.
 - ◆ to ima samo za književna djela, eseje....

041

- ponovljivo polje (101) konvertira se u ponovljiva potpolja (041a)

044

zemlja izdavanja

080 - UDK

- 675 potpolja b, c, d ... - što s njima? - provjeriti za svaku bazu

084 - other classification (R)

- NSK ga koristi za stručnu oznaku NSK iz polja 681
- NSK konvertira polje 686 u 084, u skladu s LOC konv. tablicom
- polje 686 koristi FFPS. još netko?

100

- I1 = 0 ako je pseudonim (ima 700a, nema 700b)

111 i 110

- konverzija u 110 ili 111 ovisno o postojanju/nepostojanju nekih polja - provjeriti ispise naslova iz svake baze - moguće su greške.

245 indikatori

- i1 = 0 za anonimne publikacije, i1 = 1 ako postoji 700 ili 710
- i2 = pretpostavlja se na temelju glavnica na početku naslova i jezika
- provjeriti

245

- potpolja i interpunkcije pripremljeni u modify

246

250

- zapisima koji nemaju potpolje, dodaje se ^a u modify - provjeriti

260

- ponovljiva potpolja u originalnom redosljedu - priprema u modify

300

- urediti interpunkcije za sve kombinacije

362 - brojÄni podaci za nakladniÄku cjelinu - razjasniti upotrebu

440 - indikatori

1. 2.ind. prema jeziku i Älanu, provjeriti

440

- marc_compose za 440 radi samo kad polje ni potpolja nisu ponovljiva. ne zapisuje dobro drugo pojavljivanje ^v, tj. ^v iza ^p (iz 225^w)
- 225w treba ici u 440v, na kraj

500

502 - biljeÄka o disertacijama

- odgovarajuÄeg polja nema u originalnim podacima. moÄda se moÄe pretpostaviti?

504

505

- FFPS 327 - sadrzaj
- NSK konvertira 327 (NP) u 501 (R)
- u LOC konv. tablici stoji 327 (NP) u 505 (NR)
- standard i validacije dopuÄtaju 501 (R) i 505 (R)

520

- NSK konvertira 330 (P) u 520 (P) = LOC konv. tablica
- FFPS 330 (NP) - saÄetak

526 - STUDY PROGRAM INFORMATION NOTE

- provjeriti za svaku bazu sta je u 996?

538 - zahtjevi sustava

655

653

700

- ako je u originalu ponovljivo polje treba biti i konvertirano u ponovljivo !!

740 - nadopuniti

760 w

- generiranje zapisa o nakladniÄ koj cjelini i povezivanje podreÄ enog zapisa s nadreÄ enim

852 - signatura

876 - item information - basic bibliographic unit

886 - former marc
