

## Projektas UseMARCON

Brian Phillip HOLT

Britų biblioteka, GB - Boston Spa, Wetherby, West Yorkshire LS23 7BQ, el. paštas: brian.holt@bl.uk

UseMARCON (Vartotojų kontroliuojamas MARC konvertorius) - tai projektas, kurį finansavo Europos Sąjunga. Jo tikslas - pagaminti asmeninio kompiuterio programinę įrangą, kuri galėtų konvertuoti įrašus, parengtus vienu MARC formatu, į kitą MARC formatą per UNIMARC. Rašydamas MARC, aš turiu galvoje „atitinkantis ISO 2709“.

Projektą vykdyti pradėta 1994 m. vasarį, o baigta 1997 m. sausio mėnesį. Rezultatas buvo konvertavimo programa, jos vartojimo vadovas, techninis įvairių programų rašymo vadovas ir kai kurių programų pavyzdžiai. Nyderlandų karališkoji biblioteka įvedė juos viešam naudojimui savo ftp puslapyje - ftp://ftp.konbib.nl/pub/usmarcon/rcadme.txt, kuriame pateikiama turimos medžiagos apžvalga. Europos Sąjungos puslapis <http://www2.echo.lu/libraries/en/software.html> („Programinės įrangos produktai projektams, kuriuos remia Telematics for Libraries“) turėtų užtikrinti ryšį su šiuo puslapiu.

Vartotojo vadovas aiškina, kaip įeiti į taikomąją programą, kaip pasirinkti konvertuojamą failą ir kaip atlikti konversiją. Ši UseMARCON dalis yra patogi vartotojui.

### Programų rengimas

Pagrindinė mintis yra ta, kad kataloguotojas, turintis tam tikrą darbo su formatais patirtį, padedamas sistemų analitiko, turėtų sugebėti parašyti taisykles, kaip konvertuoti įrašus, parengtus šalies formatu, į UNIMARC ir iš UNIMARC. Jo kolega iš kitos šalies turėtų padaryti tą patį, taip sudarydamas sąlygas keistis savo įrašais. Jeigu vienos valstybės formatas būtų UNIMARC variantas, tuomet dalies darbo galima išvengti.

Panagrinėkime Britų bibliotekos ir Kongreso bibliotekos pavyzdį. Jungtinėje Karalystėje parengtos konvertavimo lentelės iš UKMARC į UNIMARC ir iš UNIMARC į UKMARC. JAV tas pats daroma su USMARC. Imami amerikiečių įrašai ir naudojamos UNIMARC į UKMARC taisyklės, kad būtų gauti įrašai, kuriuos galima naudoti. Žinoma, jei procesas tik dvišalis, būtų daug greičiau perrašyti USMARC į UKMARC ir

UKMARC į USMARC konvertavimo programas. Tačiau jeigu reikia dirbti su kitų valstybių įrašais, žymiai prasmingiau yra konvertuoti UNIMARC į UKMARC, nes tai padaro prieinamus, pavyzdžiui, lietuviškus įrašus.

Kai taisyklės jau parengtos, konversija vyksta „Windows“ terpėje įrašas po įrašo arba paketais. Pirmasis metodas įgalina koreguoti įrašus tais atvejais, kai programoje neįmanoma numatyti visų atsitiktinumų. Antrasis, be abejo, yra greitesnis, ir tai vienintelis galimas metodas, kai tenka konvertuoti šimtus įrašų.

### Taisyklių programa

Taisyklių parengti pats vienas bibliotekininkas negali, kadangi net pati paprasčiausia programavimo kalba vis dėlto yra programavimo kalba. Taigi geriausia kreiptis pagalbos į sistemų analitiką.

Pavyzdžiui:

```
65011 | 60611 | "
65012 | 60612 | "
650 | 606 | S; Replace('$y' By '$A'); Replace('$z' By '$y'); Replace('$A' By '$z')
650 | 606$2 | 'lc'
```

Tai reikštų:

Konvertuojant žymės 650 (UKMARC) indikatorių 1 į žymės 650 indikatorių 1 (UNIMARC) palikti indikatoriaus vietą tuščią. Tą patį padaryti indikatoriui 2.

Trečioji eilutė atrodo beprasmiška, tačiau taip nėra. Pakeisti (\$y į \$z) pačioje pradžioje sukuria naują \$z polaukį, kuris po to antruoju nurodymu konvertuojamas atgal į \$y. Tuo būdu išvengiama pavojaus - kaip \$As, kol vyksta \$z į \$y konversija ('S' reiškia „paimk šaltinio įrašo turinį“).

Ketvirtoji eilutė paprasčiausia reiškia „Sukurk \$2 polaukį ir įdėk jį į „lc“, kad būtų matyti, jog dalykinė rubrika yra Kongreso bibliotekos dalykinė rubrika“.

Taigi

### Projektas UseMARCON

```
650 7$a Universities$Lithuania$y1900-
tampa
606 $aUniversities$Lithuania$z1900-$2lc
Tačiau gyvenime ne viskas taip paprasta:
```

```
//Library of Congress Control Number
01011 | 01011 | "
01012 | 01012 | "
010$a | <01A$a | "
010$a | <01B$a | To(2)
010$a | <01C$a | From(4); Val(S); If S>100000
Then S
010$a | <01C$a | From(4); Val(S); If S<100000 And
S>10000 Then'0'+S
010$a | <01C$a | From(4); Val(S); If S<10000 And
S>1000 Then'00'+S
010$a | <01C$a | From(4); Val(S); If S<1000 And
S>100 Then'000'+S
010$a | <01C$a | From(4); Val(S);If S<100 And S>10
Then'0000'+S
010$a | <01C$a | From(4); Val(S); If S<10
Then'00000'+S
01A$a | 010$a |
01B$a | 010$a | +S
01C$a | 010$a | +S
```

Šį pavyzdį suprasti daug sudėtingiau.

Eilutės konvertuoja sutrumpintą Kongreso bibliotekos numerio formą į teisingą formą su dviem tuščiom vietom pradžioje ir nuliais priekyje, pavyzdžiui, 01000\$a73-2284 tampa 010##\$a##73002284. # reiškia tuščia.

Taigi: „/“ eina prieš bet kurį konvertą, skirtą skaityti žmogui, o ne kompiuteriui. Konvertuoti galima tik iš šaltinio įrašo į adresato įrašą, todėl kartais norint atlikti konversiją, būtina į šaltinio įrašą papildomai įvesti laikiną lauką. '<' šiuo atveju yra kairioji strėlė. '<01\$Ca>' reiškia „įvesti šiuos duomenis į laikiną lauką 01C, polaukį \$a, šaltinio įrašė“.

„To (2)“ reiškia „paimk du pirmus ženklus“ (pavyzdyje-73). „From (4)“ reiškia „paimk kiekvieną ženklą pradėdant ketvirtuoju“ (pavyzdyje-2284). „Val“ šiuo atveju, kaip ir kitose programavimo kalbose, reiškia „paimk ženklų eilutę, pvz., 2284 ir elkis su ja kaip su skaičiumi“.

„+S“ reiškia „paimk tai, kas yra šaltinyje ir pridėk prie to, kas ten jau yra“. „00+S“ reiškia, kad 2284 (tai, kas yra šaltinyje) reikia pridėti prie 00, kad gautume 002284.

Visa tai suprantama sistemų analitikui arba programuotojui, bet nesuvokiama niekam kitam. Todėl komentarai turėtų padėti.

Daugeliu atvejų konvertavimas yra gana paprastas:

```
//History note
503 | 30511 | "
503 | 30512 | "
503$a | 305$a | S
```

Failuose skirtuose įvesti į Web tinklą, yra pavyzdinės taisyklių programos, naudojamos kaip modeliai. Tai USMARC į UNIMARC, UNIMARC į USMARC ir UK (AUTHORITIES) į UNIMARC/A. Dauguma modelių dar nėra užbaigti ir naudojami tam, kad parodytų jog tokia konversija yra įmanoma. Išimtį sudaro tik UNIMARC/Authorities konvertavimas, kuris iš tikrųjų naudojamas kitame Europos Sąjungos projekte (Projektas AUTHOR).

UseMARCON turi taisyklių redagavimo funkciją. Nenaudokite jos. Tai silpniausia visos taikomosios programos dalis.

Geriausias būdas kurti taisyklių programą yra naudoti teksto redaktorių. Programą galima išbandyti pasinaudojant pavyzdžių failais, įrašas po įrašo. Šaltinio įrašą ir adresato įrašą galima pamatyti ekrane ir išsiaiškinti klaidas. Tai priedas prie „error log“, kuriame automatiškai ir fiksuojančio klaidas, kurias pastebi UseMARCON taikomoji programa. Po to programą galima ištaisyti naudojant teksto redaktorių.

### Kitos programos

Yra dar kitų smulkesnių elementų, kuriuos reikėtų parašyti. Tai lentelės, padedančios konvertuoti šaltinio ženklų rinkinį į tikslinį ženklų rinkinį ir tikrinančios, ar šaltinio ir adresato įrašas atitinka formatų taisykles (pvz., kad nekartojamas polaukis pasirodytų tik vieną kartą). Visos šios lentelės yra palyginti paprastos, o jose naudojamas simbolius galima išmokyti.

### Pabaigai

Kaip turbūt supratote, UseMARCON yra paprastas. Tačiau jis įgalina konvertuoti įrašus asmeniniu kompiuteriu, naudojant programas, kurias per kelias minutes, o ne per kelias savaites, galima pakeisti taip, kad jos atitiktų naujausius reikalavimus.