

Demonstrate set of features

Added to cataloger's toolkit; works with Voyager

Critical bits also available as code modules for use elsewhere

In toolkit, button is "IA" button

Key bit was figuring out how to display Unicode character set in Visual Basic 6 Thanks, David Bucknum!

1		Convert romanized text to vernacular, and vice versa	
Recycle Bin	Calculator	Script/language:	
		- Chinese	
	4	005 20070108163323.0	
Eudora	scrt419.exe	035: : ‡a (OCoLC)ocm42268073	
		040: : ‡a COO ‡c COO ‡d IEN	
	12	042: : ‡apcc	
VB	wodSFTPdll	043: : ‡a a-kz	
		050: 4: ‡a DK908.867 ‡b .K48 1998	
<u>a</u>		100:1 : ‡a Khli ⁻ u ⁻ pin, V. N. ‡q (Vitaliĭ Nikolaevich), ‡d 1969-	
Production	wodSETPdll.zin	245:10: ‡a "Bol'shai´a` sem'i´a`" Nursultana Nazarbaeva : ‡b politicheskai´a` ėlita	
		sovremennogo Kazakhstana / ‡c V.N. Khli´u`pin.	
<i>.</i>		246:10: ‡a Politicheskai a` ėlita sovremennogo Kazakhstana	
	200	260: : ‡a Moskva : ‡b In-t aktual'nykh politicheskikh issledovaniĭ, ‡c 1998.	
	mm.EXE	300: : ‡a 111 p.;‡c 21 cm.	
		651: 0: ‡a Kazakhstan ‡x Politics and government ‡y 1991-	
		650: 0: ‡a Elite (Social sciences) ‡z Kazakhstan ‡x Political activity.	
notepad.exe	Microsoft Word	600:10: ‡a Nazarbaev, Nursultan ‡x Family.	
	-	650: U: Ta Post-communism TZ Kazakinstan.	
		948: : t b n t c 2000 1009 t a n t e RR	
Explorer	My Computer		
-			
Catalog	ger's toolkit		
		Show diacritics	
		With base characters C As text labels Convert highlighted text Convert whole record	
014	C PP C AR	Replace Ucase 1st Lcase 1st Print Cancel Save	
<u> </u>			
🎒 Start 🛛 🔏	🍯 🕑 💻 Wind	🛍 Snagit 🛛 🖕 Meeti 💭 Voya 🔛 Corre 🝺 SETU 🧞 Eudor 🐁 CTKV 🖏 OCLC 🎕 Catal 🔲 Micro 🛛 🕱 🗞	🧕 🔗 9:14 AM

Click button in toolkit, Voyager record presented in window Not available for direct editing; indirect manipulation via buttons

Convert romanized text to vernacular, and vice versa	
Script/language:	
Chinese ▼	
Greek classical	
Chinese	
Russian	
Chinese Wade-Giles to Pinyin	
Belorussian 🔶 🔶	
Bulgarian	
SerbianMacedonian	ikolaevich), ‡d 1969-
Ukrainian	ultana Nazarbaeva : ‡b politicheskai ົa ` ėlita
sovremennogo Kazakhstana / ‡c V.N. Khl	i Tu Tpin.
246:10: ‡a Politicheskai a` ėlita sovrer	mennogo Kazakhstana
260: : ‡a Moskva : ‡b In-t aktual/nykh	politicheskikh issledovaniĭ, ‡c 1998.
300: : ‡a 111 p. ; ‡c 21 cm .	
651: 0: ‡a Kazakhstan ‡x Politics and	government ‡y 1991-
650: 0: ‡a Elite (Social sciences) ‡z K	azakhstan ‡× Political activity.
600:10: ‡a Nazarbaev, Nursultan ‡x Fa	amily.
650: 0: ‡a Post-communism ‡z Kazak	hstan.
948: : th h tc 20001009 td n te RR	
Show diacritics	
With base characters C As text labels Convert highlighter	téxt <u>Convert w</u> hole record
<u>Replace</u> Ucase 1st Ucase each Lcase 1st P	rint Cancel Save

Just look at this window

Critical bit: select proper conversion table

Can define as many as you want; under local control

Eight defined here

Note: each different script/language combination requires separate table

After that, critical buttons: 'convert highlighted text' and 'convert whole record'



Separate configuration file for each conversion

Principal sections: convert vernacular script to romanized; convert romanized script to vernacular

Also: which fields and subfields participate in whole-record conversion; general conversion method; handling of uppercase; other stuff

RussianRomanization.cfg - Notepad	
File Edit Format View Help	
[RomanToScript]	<u>^</u>
A=U+0410	
a=U+0430	
B=U+0411	
b=U+0431	
V=U+0412	
V=U+0432	
EU+0307=U+042D	
11110304=1110456	
10+0306=0+0439	
iu+FE20eU+EE21=U+0463	
1U+FE20uU+FE21=U+044E	
iU+FE20aU+FE21=U+044F	
i=U+0438	
[ScriptToRoman]	
U+0401=EU+0308	
U+0451=eU+0308	
U+0410=A	
U+0430=a	
U+043L=D	
U+U+32=V	

Extract from Russian table

Top section: convert romanized text to vernacular

Romanized 'A' becomes Unicode character 0410

E with dot above becomes 042D

ligature IE becomes 0462

Bottom section: convert vernacular to romanized

Unicode 0410 becomes 'A'

For Russian, essentially a mirror of the other stanza; not so for other scripts

Sometimes, only goes one way

Convert romanized text to vernacular, and vice versa
Script/language:
Russian -
005 20070108163323.0
035: : ‡a (OCoLC)ocm42268073
040: : ‡a COO ‡c COO ‡d IEN
042: : ‡a pcc
043: : ‡a a-kz
050: 4: ‡a DK908.867 ‡b . K48 1998
100:1 : ‡a Khli ⁻ u ⁻ pin, V. N. ‡q (Vitaliĭ Nikolaevich), ‡d 1969-
245:10: ‡a "Bol'shai´a` sem'i´a`" Nursultana Nazarbaeva : ‡b politicheskai´a` ėlita
sovremennogo Kazakhstana / ‡c V.N. Khli´u`pin.
246:10: ‡a Politicheskai a`ėlita sovremennogo Kazakhstana
260: : ‡a Moskva : ‡b In-t aktual′nykh politicheskikh issledovaniĭ, ‡c 1998.
300: : ‡a 111 p.;‡c 21 cm.
651: 0: ‡a Kazakhstan ‡x Politics and government ‡y 1991-
650: 0: ‡a Elite (Social sciences) ‡z Kazakhstan ‡× Political activity.
600:10: ‡a Nazarbaev, Nursultan ‡x Family.
650: 0: ‡a Post-communism ‡z Kazakhstan.
948: : ‡b h ‡c 20001009 ‡d n ‡e RR
Show diacritics Image: With base characters Image: As text labels Convert highlighted text Convert whole record
Replace Ucase 1st Lcase 1st Print Cancel Save

Cataloger has found a record in romanized Russian

Selected correct conversion table

Feels lucky, clicks 'whole record' button

Convert romanized text to vernacular, and vice versa	
Script/language:	
Russian	
050: 4: ta DK908 867 tb K48 1998	
300: : ‡a 111 p.;‡c 21 cm.	
With base characters C As text labels Convert highlighted text Convert whole record	
Replace Ucase 1st Ucase each Lcase 1st Print Cancel Save	

Collor added to show paired fields

Program converts each piece it's told to convert, and adds a linked 880 field Russian is an easy conversion, so there's nothing major to point out here Exception: 651: maybe this isn't such a good idea; under local control

Convert romanized text to vernacular, and vice versa	
Script/language:	
Greek classical 🚽 🗾	
130:0 : ‡a Bible.‡p N.T.‡l Greek.‡f 1675.	
245 : 14 : ‡a <mark>Hē Kainē Diathēkē</mark> = ‡b Novum Testamentum .	
246:30: ‡a Kainē Diathēkē 👌	
246:31: ‡a Novum Testamentum	
250: : ‡a Editio nova, denuo recusa / ‡b in qua diligentius quàm unquam antea	
variantes lectiones tam ex manuscriptis quàm impressis codicibus collectae, &	
parallela Scripturae loca annotata sunt, studio & labore Stephani Curcellaei.	
260: : ‡a Amstelodami : ‡b Apud Danielem Elsevirium, ‡c 1675.	
300: : ‡a [12], 420, 312 p ; ‡c 15 cm .	
546: : ‡a Greek, with some Latin.	
500: : ‡a "Second edition of Curcellaeus' Testament [of which the 1st ed. was	
published 1658]. The readings previously given in the Appendix are now incorporated	
in the critical apparatus below the text. The Prologus added at the end of the book is	
borrowed from Boecler's edition of 1645" (D. & M.).	
500: : ‡a "Tēs Kainēs Diathēkēs deuteron meros, periechon tas Epistolas kai tēn	
Apokalypsin" P. 1-312 in third group.	
500: : ‡a The text is printed in a single column, divided into verses, with parallel	T
Show diacritics	
With base characters C As text labels Convert highlighted text Convert whole record	
Replace Ucase 1st Lcase 1st Print Cancel Save	

Record contains some bits in romanized Greek

Whole-record conversion is not the way to go; piece by piece

Cataloger selects correct table

Highlights piece of text

Click 'convert highlighted text' button

Convert romanized text to vernacular, and vice versa	
Script/language:	
Greek classical	
130:0 : ‡a Bible.‡p N.T.‡l Greek.‡f 1675.	3
246:37: ‡a Kaine Diatheke	
246:3L: ‡a Novum Testamentum	_
250: *: ‡a Editio nova, denuo recusa / ‡b in qua diligentius quàm unquam antea	
variantes lectiones tam ex manuscriptis quàm impressis codicibus collectae, &	
parallela Scripturae loca annotata sunt, studio & labore Stephani Curcellaei.	L
260: : ±a Amstelodami : ±b Apud Danielem Elsevirium, ±c 1675.	L
300: : ta[12] 420 312 p tc 15 cm	L
546 · · · +a Greek with some Latin	L
540. • ta "Second adition of Curcellague' Testament [of which the 1st ad was	
solo: : 14 Second edition of Curcenzeus restament for which the Isted. was	-
published 1658]. The readings previously given in the Appendix are now incorporated	
in the critical apparatus below the text. The Prologus added at the end of the book is	
borrowed from Boecler's edition of 1645" (D. & M.).	
500: : ‡a "Tēs Kainēs Diathēkēs deuteron meros, periechon tas Epistolas kai tēn	
Apokalypsin" P. 1-312 in third group.	
	1
With base characters C As text labels Convert highlighted text Convert whole record	
Replace Ucase 1st Lcase 1st Print Cancel Save	

Program creates a paired field again

Highlighted text converted

Adjustment of nonfiling characters indicator

Not getting into question of rough breathing mark; under local control

Cataloger has selected next piece, ready to go

Convert romanized text to vernacular, and vice versa	
Script/language:	
Chinese -	
001 4568522	
005 20070109121442.0	
040: : ‡a PIT ‡c PIT	
100 : 1 : ‡a 张海惠.	
245:10: ‡a 二十世纪中国少数民族文献分布及学术研究成果: ‡b 国际性书	目之书目=
An annotated international bibliography of bibliographies on studies of China	's ethnic
minorities in the 20th century / ‡c 编纂者张海惠, 王炬.	
246:30: ‡a 国际性书目之书目	
246:31: ‡a Annotated international bibliography of bibliographies on studie	es of
China's ethnic minorities in the 20th century	
│ 260: : ‡a 北京 : ‡b 商务印书馆, ‡c 2006 .	
300: : ‡a 14, 5, 340 p . ; ‡c 21 cm .	
500: : ‡a Includes indexes.	
546: : ‡a In Chinese and English.	
650: 0: ‡a Ethnology ‡z China ‡v Bibliography ‡v Catalogs.	
651: 0: ‡a China ‡x Ethnic relations ‡v Bibliography ‡v Catalogs	
700:1 : ‡a 王炬.	-
Show diacritics • With base characters C As text labels Convert highlighted text Convert whole record	
Replace Ucase 1st Ucase each Lcase 1st Print Cancel Save	

So far looking at romanized records; this record in vernacular Operator selects conversion table, clicks 'whole record' button



Several things to point out here

- •For Chinese, program normally uppercases first word in each romanized field
- •For personal names, program adds comma, uppercases each word

•In 245, "Zhongguo' is one word instead of two because combination is defined in conversion table; uppercased in conversion table

- •Same with 'Beijing' in 260
- •First word in \$b uppercased because conversion file said to do that, too
- •245 contains 2 unconverted characters: not in conversion table
- •Also: words in \$c beginning "zhang" should be uppercased (authors)
- •Buttons at bottom help operator clean up record (could also clean up in Voyager)
- •'Define': highlight character, supply equivalent
- •'Replace': edit text



Operator has selected 'replace'

Operator now provides complete replacement for this text

Convert romanized text to vernacular, and vice versa
Script/language:
Chinese Wade-Giles to Pinyin 🛛 🚽 🔽
041:0 : ‡a chiengjpn
043: : ‡a a-cc
050:0 : ‡a NK7983 ‡b .K78
066: : ‡c\$1
110:2 : ‡a Guo li gu gong bo wu yuan. 🗡
245:10: ‡a Ku kung t'ung ch'i hsüan ts'ui = ‡b Masterworks of Chinese bronze in the
National Palace Museum.
260: : ‡a Taipei : ‡b Kuo li ku kung po wu yüan = National Palace Museum, ‡c Min
kuo 59 [1970]
300: : ‡a 109 p. : ‡b chiefly ill. (incl. 50 col. plates) ; ‡c 28 cm.
440: 1: ‡a [Ku kung wen wu hsüan ts'ui ; ‡v Ti 1 chi]
500: : ‡a Colophon inserted.
500: : ‡a Chinese, English, and Japanese.
650: 0: ‡a Bronzes, Chinese.
610:20: ‡a Ku kung t'ung ching hsüan ts'ui ‡v Catalogs.
740:01: ‡a Masterworks of Chinese bronze in the National Palace Museum.
948: : ‡a OCLC ‡b a ‡c 20002215 ‡e als
Show diacritics Image: With base characters Image: Convert highlighted text Convert whole record
Replace Ucase 1st Lcase 1st Print Cancel Save

Wade-Giles to Pinyin conversion available with same protocol Only works bit-by-bit; not allowed for 'full record' conversion

Convert romanized text to vernacular, and vice versa
Script/language:
Chinese Wade-Giles to Pinyin
041:0 : ‡a chiengjpn
043: : ‡a a-cc
050:0 : ‡a NK7983 ‡b .K78
066: 📭: ‡c \$1
110 🖌 : ‡a Guo li gu gong bo wu yuan.
260: : ‡a Taipei : ‡b Kuo li ku kung po wu yüan = National Palace Museum, ‡c Min
kuo 59 [1970]
300: : ‡a 109 p. : ‡b chiefly ill. (incl. 50 col. plates) ; ‡c 28 cm.
440: 1: ‡a [Ku kung wen wu hsüan ts'ui ; ‡v Ti 1 chi]
500: : ‡a Colophon inserted.
500: : ‡a Chinese, English, and Japanese.
650: 0: ‡a Bronzes, Chinese.
610:20: ‡a Ku kung t'ung ching hsüan ts'ui ‡v Catalogs.
740:01: ‡a Masterworks of Chinese bronze in the National Palace Museum.
948: : ‡a OCLC ‡b a ‡c 20000215 ‡e als
Show diacritics With base characters As text labels Convert highlighted text Convert whole record
Replace Ucase 1st Ucase each Lcase 1st Print Cancel Save

Difference is that we don't get an 880 field, the text is replaced *in situ* Ready for the next piece



Questions?