

ЭЛЕКТРОННЫЕ ТАБЛИЦЫ: ПРИМЕНЯЕМ С ПОЛЬЗОЙ

Транслитерация «по-индусски»

Для версий:
Microsoft Office 2016
LibreOffice 7.0

Одна из предыдущих заметок этого раздела сайта была посвящена созданию программы для транслитерации*. В общей сложности для её реализации в формулах было использовано 9 разных функций табличного процессора. В двух других публикациях («Имитация игровой кости» и «Новая виртуальная игральная кость» в том же разделе) было продемонстрировано, что одна и та же задача в электронных таблицах может быть решена разными способами.

В свете сказанного можно задаться следующим вопросом: каково минимальное количество функций, с использованием которых в формулах осуществимо успешное выполнение замены в тексте русских букв латинскими? В качестве ответа предлагается описание процесса создания ещё одной версии программы для транслитерации, где будет задействовано всего две функции:

ЕСЛИ (; ;) и ПСТР (; ;).

IF (; ;) и MID (; ;).

Разумеется, с таким арсеналом средств не стоит ожидать от реализации «транслитератора» особой элегантности – он, например, не будет различать строчные и прописные буквы. Изложенное далее мне представляется уместным сравнить с компьютерной программой, написанной «индусским кодом», что и отражено в заголовке настоящей заметки.

Итак, начнём с того, что в листе электронной таблицы объединим диапазоны ячеек “B3:B12” и “D3:D12”:

	A	B	C	D	E
1					
2		Исходный текст:		Обработанный текст:	
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					

Затем в диапазоне “G1:AN2” нужно создать перечень символов-заменителей для каждой буквы русского алфавита:

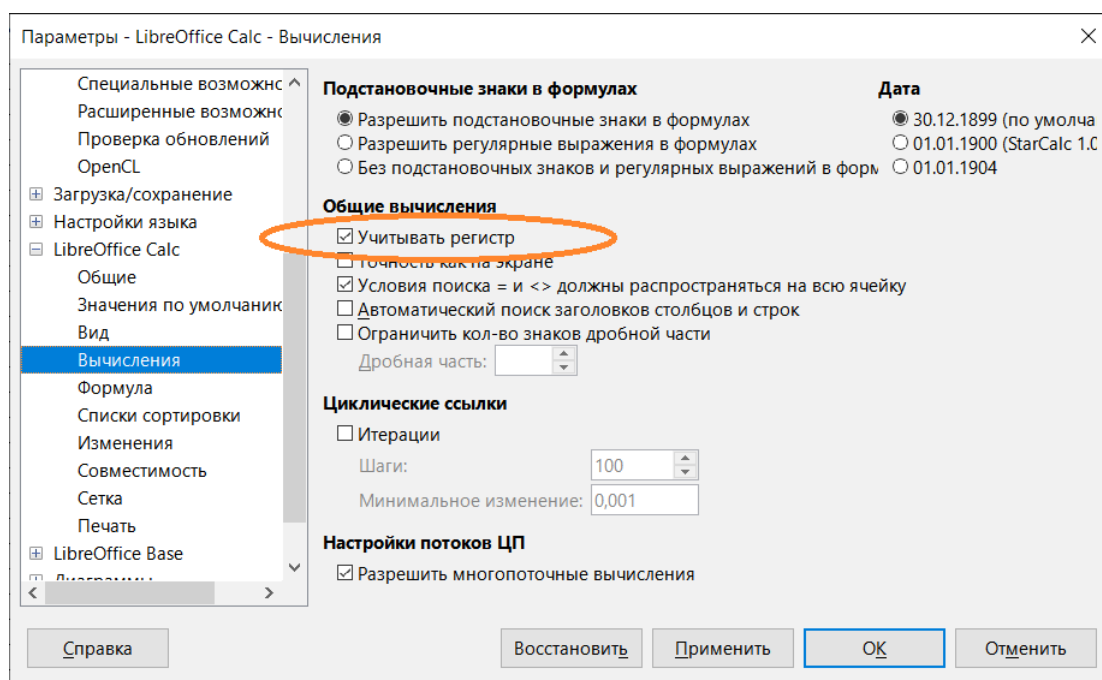
* См. заметку «Транслитерация текста»; ссылка для загрузки pdf-файла с ней находится в разделе «Электронные таблицы: применяем с пользой» сайта (URL: <http://shurichimik.narod.ru/compcreative/10-e-tables.htm>).

F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO
	Буква:	а	б	в	г	д	е	ё	ж	з	и	й	к	л	м	н	о	п	р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я	
	Замена:	а	б	в	г	д	е	yo	zh	z	i	j	k	l	m	n	o	p	r	s	t	u	f	h	ts	ch	sh	shc	"	y	'	e'	yu	ya	

Примечание: на рисунке из-за установленной ширины ячеек не видно, что буква «щ» должна заменяться на комбинацию знаков “shch”:

AF	AG	AH	AI	AJ
ч	ш	щ	ъ	ы
ch	sh	shch	"	y

LibreOffice Calc по умолчанию различает прописные и строчные буквы. Чтобы использовать приведённый выше перечень, следует отключить чувствительность к регистру: в программном меню выбрать команду «Сервис» → «Параметры...» и в появившемся диалоговом окне зайти в ветку “LibreOffice Calc” → «Вычисления», где в секции «Общие вычисления» снять флажок «Учитывать регистр»:



Можно поступить иначе – оставить регистрочувствительность, но указать в перечне замен русские буквы двойными, в виде пар «заглавная-строчная»:

G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN
Буква:	Аа	Бб	Вв	Гг	Дд	Ее	Ёё	Жж	Зз	Ии	Йй	Кк	Лл	Мм	Нн	Оо	Пп	Рр	Сс	Тт	Уу	Фф	Хх	Цц	Чч	Шш	Щщ	Ъъ	Ыы	Ьь	Ээ	Юю	Яя
Замена:	а	б	в	г	д	е	yo	zh	z	i	j	k	l	m	n	o	p	r	s	t	u	f	h	ts	ch	sh	shc	"	y	'	e'	yu	ya

Далее будет показано, как действовать в случае выбора такого варианта.

Теперь можно начинать поэтапную обработку текста, который будет вводиться в ячейку “В3”.

Этап 1. Разбивка фразы на отдельные символы

Заполните числами от 1 до 500 ячейки диапазона “F5:F504”:

E	F	G	H	I	J	K	L	M
		Буква:	а	б	в	г	д	е
		Замена:	а	б	в	г	д	е
		1. Разбивка на символы						
		№ символа	Символ					
		1						
		2						
		3						
		4						
		5						
		6						
		7						
		8						
		9						
		10						
		11						
		12						

В ячейку “G5” введите формулу, вырезающую из исходного текста один символ:

=ПСТР (\$B\$3; F5; 1)

=MID (\$B\$3; F5; 1)

При помощи маркера заполнения её нужно размножить вниз, до ячейки “G504” включительно – из сказанного становится ясно, что программа будет способна проводить транслитерацию текста, длина которого не превышает 500 символов.

Этап 2. Выбор замены для русской буквы

В ячейку “H5” введите формулу:

=ЕСЛИ (\$G5=H\$1; H\$2; "")

Если в настройках приложения чувствительность к регистру была отключена, то в ячейку “H5” нужно ввести следующую формулу:

=IF (\$G5=H\$1; H\$2; "")

Если же был выбран вариант со сдвоенными русскими буквами, то выражение в “H5” должно быть другим:

=IF (\$G5=MID (H\$1; 1; 1) ; H\$2; IF (\$G5=MID (H\$1; 2; 1) ; H\$2; ""))

Данную формулу при помощи маркера заполнения необходимо сначала размножить вправо, до ячейки “AN5”, а затем – вниз, вплоть до 504-й строки. В результате получится диапазон ячеек “H5:AN504”, в котором каждый символ обрабатываемого текста будет сравниваться с перечнем замен и при наличии соответствия в ячейках этого диапазона отобразится нужный символ-заменитель, в других случаях ячейки останутся пустыми (содержащими пустую строку “”):

F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN
	Буква:	а	б	в	г	д	е	ё	ж	з	и	й	к	л	м	н	о	п	р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я
	Замена:	а	б	в	г	д	е	yo	zh	z	i	j	k	l	m	n	o	p	r	s	t	u	f	h	ts	ch	sh	shc	"	y	'	e'	yu	ya
		1. Разбивка на символы																2. Выбор символа-заменителя																
		№ символа	Символ																															
		1	С																															
		2	ъ																															
		3	е																															

При работе со сдвоенными русскими буквами будет ещё сравниваться каждая из букв пары «заглавная-строчная» в перечне, из-за чего выражение формулы для этого случая более сложное.

Этап 3. Подтверждение замены

В ячейке “AP5” необходимо сначала указать формулу вида:

=A05&H5

Далее маркером заполнения её нужно размножить вправо, до ячейки “BV5” включительно, а после – вниз, до 504-й строки. В результате в ячейках столбца “BV” окажутся заменённые латинскими русские буквы. В случае символов, не подлежащих замене (цифры, знаки препинания и т. п.) ячейки в “BV” будут содержать пустую строку. Обратите также внимание на то, что в столбце “AO” ячейки обязательно должны быть пустыми!

В ячейку “BW5” надо ввести формулу

=ЕСЛИ (BV5=" "; G5 ; BV5) ,

=IF (BV5=" "; G5 ; BV5) ,

которую также следует размножить вниз до 504-й строки. Как нетрудно видеть, эта формула выдаёт либо заменённую на латинскую русскую букву, либо возвращает символ, который не нужно подвергать замене.

AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV	AW	AX	AY	AZ	BA	BB	BC	BD	BE	BF	BG	BH	BI	BJ	BK	BL	BMBN	BO	BP	BQ	BR	BS	BT	BU	BV	BW						
э	ю	я																																								
е'	yu	ya																																								
3. Подтверждение замены символа																																										
				а	б	в	г	д	е	ё	ж	з	и	й	к	л	м	н	о	п	р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я	итог					
									е	е	е	е	е	е	е	е	е	е	е	е	е	е	е	е	е	е	е	е	е	е	е	е	е	е	е	е	е	е	е	е	е	е

Примечание: на рисунке в диапазоне “AP4:BV4” буквы русского алфавита проставлены для удобства контроля работы программы.

Этап 4. Сборка фразы

Для соединения всех символов воедино, в ячейку “BX5” нужно ввести формулу (ячейка “BX4” при этом должна быть пуста!):

=BX4&BW5

Если её размножить вниз до 504-й строки, то в ячейке “BX504” в результате постепенного «нанизывания» символов

BW	BV	BW	BX
			4. Сборка фразы
ю	я	итог	
s	s	s	s
"	"	"	s"
e	e	e	s"e
sh	sh	sh	s"esh
'	'	'	s"esh'

соберётся транслитерованная фраза, которую останется только отобразить в “D3” при помощи формулы:

=BX504

Выглядит результат работы программы следующим образом:

	A	B	C	D	E
1					
2		Исходный текст:		Обработанный текст:	
3		Съешь ещё этих мягких		s''esh' eshchyo e'tih	
4		французских булочек, да		myagkih frantsuzskih	
5		выпей чаю!		bulochek, da vypej chayu!	
6					
7					
8					
9					
10					
11					
12					

Как видно, «транслитератор на двух функциях» получился вполне работоспособным. Несмотря на заметную ограниченность по возможностям (в сравнении с первой версией), он позволяет варьировать способы замены русских букв. Так, если в перечне замен для «ц» указать “с” вместо “ts”, то выглядеть результат обработки будет следующим образом:

	A	B	C	D	E
1					
2		Исходный текст:		Обработанный текст:	
3		Широкая		shirokaya e'lektrifikaciya	
4		электрификация южных		yuzhnyh gubernij dast	
5		губерний даст мощный		moshchnyj tolchok	
6		толчок подъёму		pod''yomu sel'skogo	
7		сельского хозяйства.		hozyajstva.	
8		Аэрофотосъёмка		ae'rofotos''yomka	
9		ландшафта уже		landshafta uzhe vyyavila	
10		выявила земли богатых		zemli bogatyh i	
11		и процветающих		procvetayushchih	
12		крестьян.		krest'yan.	
13					

© Широков Александр, 25.11.2021