

Как читать кириллицу на Интернетe и пересылать по электронной почте: некоторые советы и пояснения

Автор: [Григорий Наумовец](#)
(E-mail: g_n@online.com.ua).

Данный документ не претендует на роль всеобъемлющей инструкции и отражает личный опыт автора. В нем не рассматриваются кодировки кириллицы, специфичные для конкретных языков, за исключением русского и украинского. В разделе '[Литература](#)' можно найти ссылки на другие материалы по этому вопросу. Автор будет признателен за любые [дополнения, исправления или комментарии](#).

Дата последней редакции: 2 июля 2000 г.

ПРИМЕЧАНИЕ: В данном документе под 'кириллицей' подразумеваются главным образом тексты на русском или украинском языке. Вопросы, специфичные для болгарского, сербского или македонского языков, здесь не затрагиваются (насколько я знаю, тексты на этих языках обычно представлены на Интернетe в кодировке Windows-1251).

СОДЕРЖАНИЕ

- [Вступление](#)
- [Кодовая таблица ASCII и наиболее распространенные кодировки кириллицы](#)
 - [7-битная кодировка ASCII](#)
 - [Наиболее распространенные 8-битные кодировки кириллицы](#)
 - ['Унифицированная' кодировка UNICODE](#)
- [7-битные представления 8-битных текстов: 'Quoted-printable' и 'base64'](#)
- [Как настроить браузер для просмотра Web-страниц на русском языке](#)
 - [Netscape Navigator \(v.2,3\)](#)
 - [Microsoft Internet Explorer \(v.3, 4, 5\)](#)
 - [Netscape Communicator \(v.4.*\)](#)
 - [HTML-документы с явным указанием кодировки \(charset'a\)](#)
- [Как настроить почтовую программу для переписки на русском языке](#)
 - [В каком виде отправлять письма?](#)
 - [В какой кодировке?](#)
 - [В 8-битном или 7-битном формате?](#)
 - [Пересылка приложенных файлов](#)
 - [Как отправлять письма в кодировке КОИ-8: возможные подходы](#)
 - [Как набирать текст в кодировке КОИ-8 под Windows](#)
 - ['Родной' переключатель клавиатуры Windows95](#)
 - [CyrWin 95](#)

- [ParaWin 95/98](#)
 - [WinKey](#)
 - [Перекодировка текста с помощью почтовой программы](#)
 - [Обработка кириллицы некоторыми почтовыми программами](#)
 - [Таблица сравнения почтовых программ](#)
 - [Eudora Pro \(v2.2 & 3.*\)](#)
 - [Netscape Mail \(Mozilla\) v.2](#)
 - [Netscape Mail \(Mozilla\) v.3](#)
 - [Microsoft Internet Mail](#)
 - [Netscape Messenger \(Mozilla\) v.4](#)
 - [Outlook Express](#)
 - [Pegasus Mail for Windows](#)
 - [The Bat!](#)
 - [Vmail/UUPC for DOS](#)
 - [Перекодировка текста с помощью прокси-сервера](#)
 - [Перекодировка текстов с помощью специальных 'внешних' программ](#)
 - [некоторые перекодировщики для DOS](#)
 - [CP Tuner 97](#)
 - [Total Recode II и некоторые другие](#)
 - [Онлайновые перекодировщики](#)
 - [Проблемы, специфичные для украинского языка](#)
 - [Русификация Windows98](#)
 - [Литература](#)
 - [Благодарности](#)
-

Вступление

При обмене текстовыми сообщениями по электронной почте (в отличие от факса) пересылаются не изображения букв, а их цифровые коды. Чтобы адресат мог правильно прочесть отправленное ему письмо, нужно, чтобы на его компьютере и на компьютере отправителя одним и тем же кодам соответствовали одни и те же буквы (иначе говоря, чтобы отправитель и получатель письма пользовались одной и той же кодировкой текста). Если же они пользуются различными кодировками, нужно, чтобы на каком-то этапе прохождения от отправителя к адресату письмо было правильным образом перекодировано – так, чтобы полученные коды были другими, но соответствовали на компьютере получателя тем же буквам и символам, что и исходные коды на компьютере отправителя. С латинскими буквами таких проблем обычно не возникает, поскольку используемый для их кодирования стандарт является общепринятым. Зато для букв кириллицы приходится иметь дело с несколькими различными кодировками. Чтобы понять, какие из-за этого могут возникать проблемы и как их решать, нужно иметь представление об основных кодовых таблицах кириллицы и о том, что может происходить с кириллицей при прохождении сообщения через почтовые программы и почтовые серверы.

Кодовая таблица ASCII и наиболее распространенные кодировки кириллицы

7-битная кодировка ASCII

Для кодировки букв латинского алфавита общепринятым является стандарт ASCII (American Standard Code for Information Interchange - американский стандартный код обмена информацией). В нем буквам латинского алфавита, а также цифрам, знакам препинания и основным символам соответствуют цифровые коды, лежащие в диапазоне от 32 до 127. 'Латинские' тексты еще называют *7-битными*, поскольку для их кодирования достаточно семи двоичных разрядов. (Возведение двойки в седьмую степень дает 128, поэтому семью двоичными разрядами можно закодировать 128 различных букв и символов). Если такой текст пересылается по электронной почте, то его кодировка обычно указывается в служебных полях письма строкой: `Content-Type: text/plain;`
`charset='us-ascii'`.

Наиболее распространенные 8-битные кодировки кириллицы

Коды, лежащие в диапазоне от 128 до 255 ('8-битные коды' - они состоят уже из восьми двоичных разрядов), используются для представления дополнительных или 'национально-специфичных' символов. Существует большое количество кодовых таблиц (их еще называют 'кодowymi страницами'), различающихся в зависимости от используемой операционной системы и от ее настройки на конкретные языки. Чтобы правильно определить букву или символ, которым соответствует конкретный код, нужно знать, какая кодовая страница используется в данном случае для кодирования текста.

В среде Windows для работы с русскоязычными текстами обычно используется кодовая страница **1251** (ее иногда упрощенно называют '**кодировка Windows**'). Например, если набрать текст в MS Word for Windows, пользуясь 'обычными' кириллическими шрифтами для Windows, и сохранить его как 'Text Only', то полученный текстовый файл будет записан именно в этой кодировке.

В операционной системе DOS для работы с русскоязычными текстами наиболее распространена кодовая страница **866**, известная также под названием 'модифицированной альтернативной кодировки' (упрощенно ее называют 'Альтернативная' или 'кодировка DOS'). При 'русских' региональных настройках Windows сохранение текста из MS Word for Windows в виде 'MS-DOS Text' запишет текстовый файл именно в этой кодировке. Чтобы правильно считывать такие файлы в MS Word for Windows, нужно включить через меню Tools/ Options/ General/ установку 'Confirm conversion at open', а при открытии файла в ответ на запрос программы указывать преобразование из 'MS-DOS Text'.

В операционных системах семейства UNIX для русскоязычных текстов обычно используется кодовая страница **КОИ-8** ('Код обмена информацией 8-битный'). Точнее говоря, именем 'КОИ-8' называют целое семейство немного отличающихся друг от друга кодовых страниц - например, русская KOI8-R ([RFC 1489](http://www.ietf.org/rfc/rfc1489.txt)) и украинская KOI8-U ([RFC 2319](http://www.ietf.org/rfc/rfc2319.txt)). Именно эта кодировка (в варианте 'koi8-r') принята как стандартная в русскоязычных телеконференциях USENET (например, в группах новостей иерархии `telcom.*`) и чаще всего используется для пересылки сообщений на русском языке по электронной почте. Очень много полезной информации об этой кодировке и о том, как настроить на нее различные программы, можно найти на Web-странице Андрея Чернова (<http://www.nagual.pp.ru/~ache/koi8.html>).

Под Windows тексты в кодировке КОИ-8 можно читать и распечатывать, используя шрифты типа КОИ-8 для Windows. Однако для того, чтобы набирать тексты в этой кодировке, нужно иметь специальную раскладку клавиатуры.

На компьютерах типа **Макинтош** используется своя собственная кодировка кириллицы, известная как кодовая страница 10007 (или просто 'Макинтош') . О руссификации Макинтош-совместимых компьютеров можно прочесть, например, на www.relcom.ru/Russification/MacKoi8-r, <http://www.friends-partners.org/partners/rusmac/>, <http://www.hf.uib.no/smi/files/eudtab.html>.

Кодировка кириллицы **ISO-8859-5** (в просторечии 'ISO') известна также под названием 'основная кодировка DOS ' (хотя фактически 'основной' для DOS давно уже стала 'альтернативная' кодировка 866). Преимуществом кодировки ISO-8859-5 является строго алфавитное размещение в ней русских букв, удобное для сортировки записей в базах данных.

'**Транслит**' (транслитерация) - это не кодовая страница, это просто замена букв кириллицы 'фонетически аналогичными' латинскими буквами и их комбинациями. Транслит используют как крайнее средство в том случае, если на компьютере пользователя (или адресата письма) нет кириллических шрифтов или раскладок клавиатуры.

Ниже показаны таблицы размещения символов в различных кодировках кириллицы

Кодовая страница Windows-1251

0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	0	1	2	3	4	5	6	7	8	9	A	B	C
Ъ	Г	,	г	„	...	†	‡	‰	‰	Ль	<	Ь	К	Ъ	Ц	ђ	'	'	“	”	•	—	—	™	ль	>	ь	
	У	ў	Ј	Ѡ	Ї	Ї	§	Е	©	Є	«	¬	-	®	Ї	°	±	І	і	г	μ	¶	·	ё	№	е	»	ј
А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь
а	б	в	г	д	е	ж	з	и	й	к	л	м	н	о	п	р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь

Кодовая страница 866 (DOS)

0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	0	1	2	3	4	5	6	7	8	9	A	B	C	
А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	
а	б	в	г	д	е	ж	з	и	й	к	л	м	н	о	п	•	•	•	•	•	•	•	•	•	•	•	•	•	•
Л	⊥	т	†	-	†	†	†	℄	℄	℄	℄	=	℄	±	⊥	т	т	℄	℄	℄	℄	℄	℄	℄	℄	℄	℄		
р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я	Ё	ё	Є	е	Ї	ї	У	ў	°	•	•	√	№	

Кодовая страница KOI8-R

0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	0	1	2	3	4	5	6	7	8	9	A	B	C
—		Г	Г	Л	Ј	†	†	†	⊥	†	■	■	■	■	■	•	•	•	•	•	•	•	•	•	•	•	•	•
=		Г	ё	Г	Г	Г	Г	Г	Г	Г	Г	Г	Г	Г	Г	†	†	†	Ё	†	†	†	†	†	†	†	†	
ю	а	б	ц	д	е	ф	г	х	и	й	к	л	м	н	о	п	я	р	с	т	у	ж	в	ь	ы	з	ш	э
Ю	А	Б	Ц	Д	Е	Ф	Г	Х	И	Й	К	Л	М	Н	О	П	Я	Р	С	Т	У	Ж	В	Ь	Ы	З	Ш	Э

Макинтош-кодировка кириллицы

0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	0	1	2	3	4	5	6	7	8	9	A	B	C
А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь
†	°	Ї	£	§	•	¶	І	®	©	™	Ъ	ђ	≠	Г	г	∞	±	≤	≥	і	μ	г	Ј	Є	е	Ї	ї	Ль
ј	S	¬	√	f	≈	Δ	«	»	...	ђ	ђ	К	ќ	s	—	—	“	”	'	'	÷	„	У	ў	Ц	ц	№	
а	б	в	г	д	е	ж	з	и	й	к	л	м	н	о	п	р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь

Кодировка ISO-8859-5

0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	0	1	2	3	4	5	6	7	8	9	A	B	C	
	Ё	Ъ	Г	Є	Ѕ	І	Ї	Ј	Љ	Њ	Ћ	-	У	Ц		А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	
	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	а	б	в	г	д	е	ж	з	и	й	к	л	м
	р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я	№	ё	ђ	ѓ	ѓ	ѓ	ѓ	ѓ	ѓ	ѓ	ѓ	ѓ	ќ

'Унифицированная' кодировка UNICODE

Кодировка [Unicode](http://www.unicode.org), как можно догадаться по самому ее названию, представляет собой попытку создать ЕДИНУЮ кодовую таблицу, в которой нашлось бы место для букв и символов ВСЕХ человеческих языков (см. <http://www.unicode.org>). Естественно, такая таблица должна быть достаточно большой, поэтому для кодировки каждого ее элемента отводится уже не 8, а 16 двоичных разрядов, что позволяет кодировать максимум 65536 элементов. Таким образом, 'унификация' кодов букв достигается за счет двукратного увеличения размера текстовых файлов. Стандарт Unicode согласуется со стандартом ISO/IEC 10646-1, определяющим так называемый 'Универсальный набор символов' - Universal Character Set (UCS) и его двухбайтное (два байта на каждый символ) представление UCS-2; поэтому кодировку Unicode иногда называют также UCS-2. Однако протоколы передачи данных по сети Интернет и электронной почте не позволяют передавать 16-битные символы; для пересылки текстов, записанных в Unicode, используются специальные 7-битные или 8-битные представления Unicode **UTF-7** ([RFC2152](http://www.ietf.org/rfc/rfc2152.txt)) или **UTF-8** ([RFC2279](http://www.ietf.org/rfc/rfc2279.txt)) (UTF означает 'UCS Transformation Format'). Новые версии Веб-браузеров (и связанных с ними почтовых программ) - например, Netscape Communicator v.4.* или Microsoft Internet Explorer v.4.* или 5.*, 'понимают' кодировки UTF-7 и UTF-8. Использование таких представлений Unicode может стать радикальным решением проблемы с 'вавилонской башней' различных кодировок, но пока что лишь немногие почтовые программы (например, Outlook Express, Netscape Messenger или Forte Agent) могут автоматически декодировать UTF-7 и UTF-8.

Unicode в Windows 95/98/2000 и MS Office 97/2000

Широкого распространения на Интернетe кодировка Unicode пока не получила, однако, например, в MS Office 97/2000 она используется в качестве внутренней кодировки текста (кроме того, в Microsoft Word 97/2000 включена возможность сохранения текстовых файлов в кодировке Unicode (*Save As / Unicode text* или *Save As / Encoded text*). В MS Office 95 и более старых версиях MS Office внутренняя кодировка кириллицы -- Windows-1251. С этим различием кодировок связаны проблемы совместимости шрифтов в документах, созданных с помощью разных версий MS Office. Microsoft Word 97/2000 корректно работает только со шрифтами для Windows типа Unicode - т.е. такими, в которые могут быть включены наборы символов одновременно для нескольких кодовых страниц Windows (Western, Cyrillic, Greek, Turkish и т.д.). 'Старые' (не-юникодовские) 'англо-русские' шрифты можно преобразовать в формат Unicode программой TTF Converter ([ftp://kgb.quarta.com/WinUtil/ttfconv.zip](http://kgb.quarta.com/WinUtil/ttfconv.zip)) -- после этого их можно использовать в Word 97/2000 (хотя, разумеется, в них есть лишь два script'a - Western и Cyrillic). Однако даже при использовании только юникодовских шрифтов возможны проблемы с распечаткой кириллицы из приложений MS Office 97 на некоторых принтерах. Об этих проблемах и их решении можно прочесть на Веб-странице Павла Городянского (<http://ourworld.compuserve.com/homepages/PaulGor/>). Там же есть информация и о том, как решать проблемы с копированием и вставкой текста между прикладными программами, использующими и не использующими Unicode. Для решения проблем с документами, созданными в более ранних версиях Microsoft Word с использованием шрифтов, не входящих в стандартный комплект Microsoft Windows 3.x (например, Baltica, Pragmatica и т. д.) можно использовать специальный конвертор, который есть на

<http://www.referats.corbina.ru/referats/00/CYRFONTS.ZIP>.

Что касается обратной совместимости, то, например, специальную программу-конвертор, позволяющую открывать в Word 6.0 и 7.0 файлы, записанные в формате Word 97, можно найти на <ftp://ftp.microsoft.com/Softlib/MSLFILES/wrd97cnv.exe>; однако некоторые документы Word97 со сложным форматированием не удастся открыть в Word 6.0 и 7.0 даже с помощью этого конвертора.

Подробнее о шрифтах, кодовых страницах кириллицы и поддержке различных кодировок в Windows можно прочесть на Веб-странице Константина Казарновского (<http://www.glasnet.ru/~kazarn/fontsmain.htm>).

7-битные представления 8-битных текстов: 'Quoted-printable' и 'base64'

8-битные тексты можно пересылать по электронной почте в своем исходном 8-битном виде; при этом в

служебных полях письма обычно указывается: `Content-Transfer-Encoding: 8bit`. Однако надо иметь в виду, что не все почтовые серверы (особенно в '7-битном' англоязычном мире) правильно обрабатывают 8-битные тексты. Некоторые из них вообще просто 'обрезают' восьмой бит, уменьшая все 8-битные коды на 128. Поэтому в современных почтовых программах предусмотрен режим преобразования пересылаемых 8-битных текстов в 7-битную форму - в расчете на то, что почтовая программа получателя автоматически выполнит обратное преобразование и восстановит исходный текст. (Например, в программе Eudora Pro v.2.2 и 3.* такой режим включается через меню Tools/ Options/ Sending Mail/ May Use Quoted-Printable). В представлении '*quoted-printable*' стандартные 7-битные символы оставляются без изменений, а 8-битные символы заменяются на последовательность из символа '=' и пары латинских букв и/или цифр, изображающих их шестнадцатичный код. В результате отсылаемый текст может выглядеть примерно так: `=DD=EB=E5=EA=F2=F0=EE=ED=ED=FB=E5=F2=E0=E1=EB=E8=F6=FB`, а в служебных полях письма почтовая программа укажет: `Content-Transfer-Encoding: quoted-printable`.

Другим распространенным методом 7-битного представления 8-битных текстов является '*base64*'. При его использовании почтовая программа указывает в служебных полях письма: `Content-transfer-encoding: base64`.

Современные почтовые программы при получении письма, где текст представлен в виде *base64* или *quoted-printable*, автоматически восстановят его исходную 8-битную форму; тем же получателям, чьи почтовые программы не поддерживают такое преобразование, придется расшифровывать текст с помощью специальных декодирующих программ.

Как настроить браузер для просмотра Web-страниц на русском языке

Просматривая русскоязычные Веб-сайты, можно заметить несколько способов решения проблемы различных кодировок кириллицы, используемых в различных операционных системах.

При одном подходе, документ доступен на Веб-сервере в нескольких различных кодировках, так чтобы посетители Веб-сайта могли выбрать наиболее удобный для себя вариант. (При этом не обязательно держать на сервере отдельную копию документа для каждой кодировки; на самом деле, сервер часто хранит лишь один экземпляр документа в какой-нибудь одной кодировке, зато доступ к

6

этому документу возможен через несколько 'портов', которые выполняют различные перекодировки). Так или иначе, на Веб-странице есть меню выбора кодировки; в это меню наверняка будут включены кодировки WIN-1251 и KOI8-R, а иногда и другие -- DOS (866), ISO-8859-5, Macintosh и 'транслит'.

Другой подход используется, например, Веб-сервером '[Russian Apache](#)', который пытается определить 'родную' кодировку обращающегося к серверу браузера и затем 'отдает' пользователю запрашиваемый документ именно в этой кодировке. При этом на Веб-странице может и не быть меню выбора кодовой страницы. Если пользователь работает в операционной системе Windows, то он, скорее всего, увидит документ в кодировке Win-1251.

Третий подход -- сделать документ доступным в какой-нибудь одной кодировке, в расчете на то, что пользователь должен уметь правильно настроить свой браузер для чтения текстов в этой кодировке. Многие русскоязычные Веб-страницы доступны либо только в KOI8-R, либо только в Win-1251.

огон лишь
KOI-8



www.israel.ru/sgt/koi8.html www.homepage.techno.ru/barsic/antikoi.htm

Если Вы хотите создать свою собственную Веб-страницу на русском языке, Вы должны знать, как работает тот сервер, на котором Вы ее хотите разместить. Если этот сервер делает перекодировку, то предполагается, что загружать туда документы нужно в определенной кодировке - иначе они станут нечитабельными. Если же сервер не делает никакой перекодировки, то Вы сами можете выбирать, в какой кодировке (или кодировках) будет доступна Ваша Веб-страница.

Имейте также в виду, что обычно легче настроить на чтение KOI-8 браузеры, работающие под Windows, чем настроить на Win-1251 браузеры, работающие под Unix. Некоторые пользователи Unix вообще с предубеждением относятся к Веб-страницам, доступным **только** в кодировке Win-1251 и считают, что вряд ли там может быть что-нибудь заслуживающее внимания.

Для чтения русскоязычных Веб-страниц достаточно 'научить' Ваш браузер понимать кодировки KOI-8 и Windows-1251 - другие кодировки кириллицы вряд ли можно найти на Интернетe в качестве **единственного** варианта кодировки Веб-страницы. (Правда, на Интернетe можно найти множество русскоязычных текстовых файлов в кодировке DOS - но они обычно предназначены для загрузки, а не для просмотра в режиме online).

Netscape Navigator (v.2,3)

Настройка браузера Netscape Navigator (версия 3) на кириллицу фактически сводится к установке шрифтов нужного типа, которыми будет изображаться исходный текст. Процедура настройки подробно описана Павлом Городянским на <http://www.relcom.ru/Russification/WinNetscape>, и здесь мы остановимся на ней лишь вкратце.

Для настройки Netscape Navigator (версия 3) нужно войти в меню Options/ General Preferences/ Fonts и установить для кодировки 'Cyrillic' кириллические шрифты (пропорциональный и постоянной ширины) типа Windows-1251, а для кодировки 'Cyrillic (KOI8-R)' шрифты (пропорциональный и постоянной ширины) типа KOI-8. При этом чаще всего проблемы возникают с установкой шрифта постоянной ширины типа KOI-8, поскольку бывает, что ни один из имеющихся шрифтов типа KOI-8 не распознается системой как шрифт постоянной ширины и не виден в окошке выбора шрифтов. В случае таких проблем, комплект 'подходящих' шрифтов можно загрузить с той же Web-страницы Павла Городянского: <http://www.relcom.ru/Russification/WinNetscape/ForWWW.zip>. В дальнейшем нужно через меню Options/ Document Encoding выбирать кодировку Cyrillic для просмотра Web-страниц в кодировке Windows-1251 и кодировку Cyrillic (KOI-8) для просмотра Web-страниц в кодировке KOI-8.

Если у Вас все еще установлен Netscape Navigator (версия 2), давно пора заменить его на более 'свежую' версию. Если же у Вас нет такой возможности, можно аналогичным образом установить шрифты типа KOI8-8 для кодировки 'Latin 2' и шрифты типа Windows-1251 для кодировки 'Korean' (в меню этой версии Netscape Navigator'a нет кириллических кодировок) . После этого выбирайте из меню Options/ Document Encoding 'Latin 2' для просмотра текстов в кодировке KOI8-8 и 'Korean' для просмотра текстов в кодировке Windows-1251.

Microsoft Internet Explorer (v.3,4,5)

В отличие от Netscape Navigator'a, Microsoft Internet Explorer (версия 3) для просмотра Web-страниц в кодировках KOI8-8 и Windows-1251 не нуждается в установке каких-либо специальных шрифтов. Достаточно через меню View/ Options/ General/ Font Settings выбрать кириллический набор символов, а MIME-Encoding установить на Windows-1251 или KOI8-R. При просмотре Web-страниц можно щелкнуть левой кнопкой мыши по иконке, находящейся в правом нижнем углу окошка Internet Explorer'a v.3, и выбрать нужную кодовую страницу из всплывающего меню. (Если этой иконки не видно, скорее всего, в системе не установлен [Multilanguage Support](#) - нужные файлы можно загрузить с <ftp://ftp.microsoft.com/Softlib/MSLFILES/lang.exe>). При этом Internet Explorer выполняет соответствующую перекодировку исходного текста.

Тот же принцип сохранен и в следующих версиях - Internet Explorer v.4 и 5, которые, кроме Windows-1251 или KOI8-R, поддерживают еще несколько кодовых страниц кириллицы, а также 7-битные и 8-битные представления Unicode UTF-7 и UTF-8. В Internet Explorer v.5 нужная кодировка выбирается через меню View/ Encoding (если нужной кодировки нет в списке, нужно выбрать пункт 'More'). Шрифт, используемый для показа кириллицы, выбирается через меню Tools/ Internet Options/ General/ Fonts.

Netscape Communicator (v.4.*)

Netscape Navigator v.4 (входящий в комплект Netscape Communicator) также не требует наличия шрифтов типа KOI8-8, и может сам 'на лету' перекодировать документы из KOI8-8. Нужная кодировка выбирается через меню View/ Character Set, а шрифт, используемый для показа кириллицы, выбирается через меню Edit/ Preferences/ Appearance/ Fonts. Настройка этого браузера для чтения кириллицы также подробно описана Павлом Городянским на <http://www.relcom.ru/Russification/WinNetscape>.

Пожалуй, обработка кириллицы в Netscape Communicator сопровождается большим количеством ошибок по сравнению с MS Internet Explorer. Иногда бывает, что кириллица на экране нечитабельна даже несмотря на то, что все настройки заданы совершенно правильно. Если такое случается, попробуйте следующие варианты:

- переключите кодировку на какую-нибудь другую, а потом назад
- перезагрузите документ (reload)
- посмотрите HTML-код документа (View/ Page Source) и попробуйте понять причину проблемы (возможно, в заголовке документа неправильно [указана](#) его кодировка (charset), или сам текст неправильно закодирован)
- если проблема все равно не решается, попробуйте прочитать тот же документ в Internet Explorer вместо Netscape.

HTML-документы с явным указанием кодировки (charset=...)

Если в заголовке HTML-документа явно указывается кодировка при помощи строки типа `<META HTTP-EQUIV='Content-Type' CONTENT='text/html; charset=windows-1251'>`

, браузер должен автоматически включить режим показа документа в данной кодировке. В этом случае попытки задать кодировку 'вручную' (через меню браузера) могут не срабатывать. Если при

этом указанная в заголовке кодировка не соответствует той, которая фактически используется в тексте, документ может стать нечитабельным. Вообще, если браузер не может правильно прочесть текст, попробуйте просмотреть исходный HTML-код (через меню View/ Page source и т.п., в зависимости от конкретного браузера). Часто это помогает понять суть проблемы. В случае неправильно указанной кодировки можно сохранить документ у себя на диске, а потом отредактировать его, исправив charset либо перекодировав текст.

Вообще говоря, явное указание кодировки в заголовке HTML-документа нужно использовать очень внимательно. Помните:

1. Лучше совсем не указывать charset, чем указать его неправильно.

2. При явно указанной кодировке документ может стать нечитабельным в некоторых старых версиях браузеров, которые не распознают название данной кодировки.

3. Если Вы конвертируете HTML-документ с явным указанием кодовой страницы в другую кодировку, не забудьте исправить в заголовке и название кодировки.

4. Не указывайте кодировку явно, если документ будет находиться на сервере, который может его перекодировать.

5. Если Ваша страница сделана в кодировке Win-1251, даже с явным указанием charset=windows-1251, не думайте, что теперь она будет читабельна абсолютно для всех. Многие версии браузеров, работающих под операционной системой UNIX, **не** поддерживают эту кодировку.

Кроме того, если Вы хотите, чтобы русский текст на Вашей Веб-странице был читабелен для всех, никогда не используйте явное задание шрифта через tag FONT FACE= . . . Это считается дурным тоном, поскольку может создать проблемы для тех, у кого на компьютере шрифт с таким названием не русифицирован (либо его вообще нет).

Как настроить почтовую программу для переписки на русском языке

В каком виде отправлять письма?

Выбор кодировки

Если два человека переписываются между собой по электронной почте, разумеется, они вправе использовать любую кодировку, приемлемую для них обоих (и для почтовой системы, через которую их письма проходят). Однако если Вы хотите отправить письмо незнакомому человеку или сразу многим адресатам через список рассылки, следует учитывать сложившиеся на Интернете стандарты и традиции. Для пересылки писем на русском языке стандартом де-факто стала кодовая страница KOI8-R, закрепленная также и в уставах русскоязычных телеконференций USENET. Использование другой кодировки может быть удобно для Вас или для Вашего конкретного адресата, но в общем случае рекомендуется, чтобы:

1. письма уходили в кодировке KOI8-R
2. в служебных полях письма была правильно указана используемая в нем кодировка ('charset=koi8-r')

В 8-битном или 7-битном виде?

Отдельно рассмотрим вопрос о том, в какой форме лучше отправлять такие письма – 8-битной или 7-битной (см. параграф ['7-БИТНЫЕ ПРЕДСТАВЛЕНИЯ 8-БИТНЫХ ТЕКСТОВ'](#)).

С одной стороны, как уже говорилось, не все почтовые серверы пропускают 8-битные тексты. Некоторые из них автоматически преобразуют 8-битный текст в 7-битную форму (например, в base64), а некоторые просто 'обрезают' восьмой бит. (Между прочим, интересная особенность кодировки КОИ-8 заключается в том, что текст в этой кодировке удается прочесть и после 'обрезания' восьмого бита, поскольку буквы кириллицы в этой кодовой странице расположены со сдвигом на 128 от 'фонетически сходных' латинских букв). Кроме того, не совсем 'легальным' является использование 8-битного текста в поле Subject: и в других служебных полях письма, поскольку действующие на Интернет стандарты предписывают представление текста в служебных полях письма в 7-битной форме.

С другой стороны, в странах бывшего СССР сложилась практика отсылки писем в 8-битной форме (включая и текст в поле Subject), и подавляющее большинство серверов без проблем пропускает такие письма. При этом многие люди пользуются почтовыми программами, которые не могут автоматически преобразовать в исходную 8-битную форму текст, полученный в виде quoted-printable или base64. Кроме того, в 8-битной форме объем пересылаемого текста меньше, чем в виде base64 и особенно quoted-printable. Наконец, тексты, закодированные в quoted-printable или base64, могут не перекодироваться при прохождении через перекодировщики из КОИ8 в Windows-1251 или DOS, которыми пользуется адресат (такие перекодировщики могут стоять, например, на его POP3-сервере) и в результате придти к нему в неправильной кодировке. Поэтому третья рекомендация сформулируем так:

3. отправляйте письма в 8-битном виде (в том числе в поле Subject). Если выяснится, что Ваш почтовый сервер или почтовый сервер получателя 'портит' 8-битные тексты, попробуйте включить режим преобразования текста в quoted-printable или base64. Если Ваш адресат не может расшифровать такие письма, попробуйте послать текст не в теле письма, а приложить текстовый файл к письму в формате UUENCODE. Использование русского текста в поле 'From' настоятельно не рекомендуется - в этом случае многие адресаты просто не смогут прочесть имя отправителя.

Пересылка текстовых файлов в виде приложений

Рекомендация пересылать текст в кодировке КОИ-8 **не** относится к случаю, когда текст отправляется не в 'теле' письма, а прилагается к письму в виде отдельного файла (закодированного в формате UUencode или base64). 'Тело' письма может перекодироваться при прохождении через разные почтовые серверы и почтовые программы, тогда как с прилагаемыми к письму файлами этого не происходит. Кодировку прилагаемых к письму текстовых файлов обычно имеет смысл выбирать в зависимости от операционной системы и программного обеспечения, которыми пользуется Ваш адресат. Учтите, что письма, текст которых прилагается в виде отдельных файлов, менее удобно читать и отвечать на них, чем на 'обычные' письма. Кроме того, если Вы хотите написать письмо не в виде обычного текстового файла, а, допустим, как документ MS Word, учтите, что такие документы гораздо больше по объему и могут переносить компьютерные вирусы (а сам формат MS Word является приемлемым далеко не для всех адресатов).

Как отправлять письма в кодировке КОИ-8: возможные подходы

Итак, письма на русском языке, как правило, рекомендуется пересылать в кодировке КОИ-8. Однако как это сделать, если Вы работаете под Windows?

Фактически, можно либо сразу набирать письма в кодировке КОИ-8, либо сначала набирать их в той кодировке, в какой это удобно делать (например, Windows-1251 или DOS), а затем преобразовывать их в КОИ-8. Тут возможны несколько подходов:

- 'Коизация' Windows (т.е. установка шрифтов и раскладок клавиатуры типа КОИ-8 для Windows)
- Использование почтовых программ, которые могут перекодировать отправляемые и принимаемые письма
- Работа через специальные почтовые серверы, которые могут перекодировать отправляемые и принимаемые письма
- Использование 'внешних' перекодирующих программ.

Как набирать текст в кодировке КОИ-8 под Windows

Чтобы набирать текст в кодировке КОИ-8 в почтовой программе, работающей под Windows, нужно иметь шрифты типа КОИ-8 для Windows и переключатель клавиатуры с соответствующими раскладками. Можно использовать либо 'родной' переключатель клавиатуры Windows95, либо специальные программы.

'Родной' переключатель клавиатуры Windows95

Процедура установки раскладки клавиатуры КОИ-8 под Windows95 описана на Web-странице Андрея Чернова <<http://www.nagual.pp.ru/~ache/koi8.html>> и (гораздо более подробно) на <<http://www.netsight.net/~ryba/cyrillic/index.html>>. После того, как такая раскладка установлена, можно

добавить к переключателю клавиатуры еще один язык (например, польский или африкаанс) и настроить его на использование раскладки клавиатуры КОИ-8. (У Андрея Чернова описана и процедура более радикальной 'коизации' Windows95, основанную на подмене системного файла gdi.exe. В результате одну из национальных кодовых страниц, поддерживаемых в шрифтах типа Unicode (например, центральноевропейскую), можно заменить на кодовую страницу КОИ8-R. После этого в системе наряду со шрифтами, например, Arial Cyrillic и Times New Roman Cyrillic появляются шрифты Arial КОИ8-R и Times New Roman КОИ8-R).

Надо сказать, что при работе с 'родным' переключателем клавиатуры Windows95 иногда возникают проблемы. Например, он может не срабатывать в окнах некоторых программ; кроме того, он не позволяет редактировать имеющиеся раскладки клавиатуры. (Редактор раскладок клавиатуры Windows95 можно найти, например, на <<http://www.kiarchive.ru/pub/cyrillic/windows/jkbd9542.zip>>, но он позволяет редактировать не всякие файлы раскладок). Наконец, если приходится пользоваться более чем тремя раскладками клавиатуры, работать с 'родным' переключателем клавиатуры Windows95 становится просто неудобно, поскольку он не позволяет использовать разные клавиши для доступа к разным раскладкам.

Если Вы испытываете неудобства при работе с 'родным' переключателем клавиатуры Windows95, имеет смысл установить специальные программы для переключения клавиатуры.

CyrWin 95

Эта программа платная. Она позволяет включать разные раскладки клавиатуры через разные клавиши быстрого доступа, а также настраивать раскладки по своему вкусу с помощью редактора клавиатуры.

Но, кроме этого, она обладает еще одним важным достоинством. Если Вы пользуетесь американской версией Windows95, у Вас могут возникать проблемы при установке русскоязычных программ (например, русские буквы в меню могут отображаться в виде подчеркиваний или других нечитабельных символов). Для решения этих проблем не всегда достаточно установить Multilanguage Support для Windows95 и кириллические системные шрифты. Может потребоваться еще и модификация некоторых системных файлов Windows95. При установке CyrWin95 эта работа выполняется автоматически, и проблем с русскоязычными программами больше не возникает.

ParaWin 95/98

Эта программа также платная (пробную версию можно найти на <http://www.smartlinkcorp.com>). В отличие от первой версии ParaWin 95 (v.1), она позволяет переключать и редактировать различные раскладки клавиатуры. Индикатор ParaWin 95 выводится на место, обычно занимаемое на панели задач 'родным' переключателем клавиатуры Windows95, и выглядит очень похоже, но работать с ним значительно удобнее. ParaWin 95 также устраняет большинство проблем, имеющихся у американской версии Windows95 при работе с русскоязычными программами. ParaWin 95 можно использовать только под Windows95, а ParaWin 98 -- и под Windows95, и под Windows98. Для Windows 2000 или NT нужен ParaWin 2000.

Иногда при использовании ParaWin для переключения клавиатуры в окне составления нового сообщения Outlook Express задача 'подвисает', и приходится ее снимать через Ctrl-Alt-Del. Возможно, в таких случаях стоит использовать ParaWin только для русификации системы (если это необходимо), а для переключения раскладок клавиатуры установить другую программу -- например, RusLat (www.ruslat.wm.ru).

WinKey

Эта программа бесплатная и довольно удобная, и может работать как под Windows95, так и под Windows 3.*. Ее можно найти, например, на <http://www.kiarchive.ru/pub/cyrillic/windows/>. В отличие от ParaWin 95 и CyrWin 95, она не позволяет 'кириллизировать' системные файлы Windows95, хотя и включает ряд полезных системных шрифтов (в том числе шрифты типа DOS-866 для Windows). Однако она не работает с теми программами, которые используют шрифты типа Unicode, и поэтому ее можно считать устаревшей.

Из других бесплатных переключателей клавиатуры можно назвать RusLat (www.ruslat.wm.ru).

Перекодировка текста с помощью почтовой программы

Вместо того, чтобы 'коизировать' Windows, можно использовать почтовые программы, которые сами могут перекодировать отправляемый и получаемый текст. Такие программы могут иметь 'внутреннюю' и 'внешнюю' кодировку текста, отличающиеся друг от друга. В качестве 'внутренней' кодировки имеет смысл выбрать ту, в которой удобно набирать и редактировать текст под используемой операционной системой (например, 1251 для Windows или 866 для DOS), а в качестве 'внешней' -- ту, которая принята для пересылки писем (для русского языка -- KOI8-R). При отправке писем почтовая программа должна преобразовывать кодировку текста из 'внутренней' во 'внешнюю' (например, из 1251 в KOI8-R), а при их получении производить обратное преобразование (соответственно, из KOI8-R в 1251).

Есть почтовые программы, которые выполняют 'фиксированную' перекодировку текста (например, из все входящие сообщения из KOI8-R в 1251, а все исходящие - из 1251 в KOI8-R), а есть и такие, у которых есть меню перекодировки с поддержкой нескольких кодовых страниц. При 'фиксированной' перекодировке для того, чтобы правильно перекодировать получаемые письма, нужно знать, в какой кодировке они приходят. В принципе, здесь возможны два основных подхода:

- предполагать, что все письма приходят в одной 'внешней' кодировке (скорее всего, в KOI8-R). В этом случае будут неправильно перекодированы те письма, которые приходят в других кодировках.
- пытаться определить кодировку входящего письма по его заголовку (служебное поле charset=...). В этом случае будут неправильно перекодированы те письма, в заголовках которых неправильно указана кодовая страница.

На Web-странице Тимура Кадышева <<http://blue.iris.mipt.ru/timur/trueway.htm>> сделана попытка сформулировать те требования, которым должны соответствовать почтовые программы и браузеры, чтобы они могли полностью взять на себя задачу перекодировки и отпала необходимость в 'коизации' Windows.

Обработка кириллицы некоторыми почтовыми программами

Таблица сравнения почтовых программ

Почтовая программа	Перекодировка текста сообщения	Указание кириллической кодировки (charset'a) в заголовке	Проблемы и комментарии
Eudora Pro v.2.2	Нет Необходима внешняя перекодировка или шрифты koi8	неправильно: iso-8859-1 (can be hacked to koi8-r)	Очень удобный интерфейс; Проблемы с: шрифтами типа Unicode, переключателем клавиатуры Win95, сообщениями, переправленными из Netscape, разбитыми на куски приложенными файлами; не очень удобна для работы с кириллицей
Eudora Pro v.3.* , 4.*	Нет Можно использовать с koi8-plugin'ом для перекодировки win<->koi	неправильно: iso-8859-1 (с plugin'ом: koi8-r или windows-1251)	То же, что и для Eudora v.2.*; добавлена поддержка нескольких account'ов, а также plugin'ов, в том числе koi8-plugin; у v.4.* есть ошибки в обработке кириллицы, которые нужно исправлять
Netscape Mail (Mozilla) v.3.*	Входящие: нет Исходящие: win to koi, koi to koi Ответы: koi to koi нужны шрифты koi8; не использовать внешнюю перекодировку	koi8-r (может отличаться от фактической кодировки)	Одноуровневые папки для хранения почты; переправляемые сообщения могут создавать проблемы для адресатов, использующих Eudora 2.2; не очень удобен для работы с кириллицей

Netscape Messenger (Mozilla) v.4.*	<p>Входящие: из любой в win Исходящие: win to koi (в зависимости от конкретной версии)</p> <hr/> <p>шрифты koi8 или внешняя перекодировка не нужны</p>	koi8-г (может отличаться от фактической кодировки)	<p>Многоуровневые папки для хранения почты; Возможность создания нескольких профилей для разных пользователей; контраинтуитивный выбор кодировки для исходящих сообщений в v.4.0*; 8-битный Subject включается через правку prefs.js</p>
MS Internet Mail	<p>Входящие: koi или win to win Outgoing: win to koi или win</p> <hr/> <p>шрифты koi8 или внешняя перекодировка не нужны</p>	<p>правильно: koi8-r или Windows-1251</p>	<p>Одноуровневые папки для хранения почты; нет выбора кодировки в режиме ответа на письмо; письма с неправильно указанной кодировкой могут быть нечитабельны</p>
Outlook Express v.4.*, 5.*	<p>Входящие: из любой кодировки в Unicode Исходящие: из Unicode в любую кодировку</p> <hr/> <p>шрифты koi8 или внешняя перекодировка не нужны</p>	<p>правильно: koi8-r, Windows-1251 и др.</p>	<p>Многоуровневые папки для хранения почты; возможность конфигурации нескольких account'ов; возможна поддержка koi8-u; одна из наиболее удобных для работы с кириллицей почтовых программ</p>
The Bat!	<p>Входящие: из любой кодировки в win Исходящие: из win в любую кодировку</p> <hr/> <p>шрифты koi8 или внешняя перекодировка не нужны</p>	<p>правильно: koi8-r, koi8-u, Windows-1251 и др.</p>	<p>Маленькая но весьма мощная почтовая программа; многоуровневые папки для хранения почты; возможность конфигурации нескольких account'ов; поддержка koi8-u; одна из наиболее удобных для работы с кириллицей почтовых программ</p>
Pegasus Mail	<p>Нет</p> <hr/> <p>после настройки - фиксированная: Входящие: koi to win; Исходящие: win to koi</p> <hr/> <p>после настройки шрифты koi8 или внешняя перекодировка не нужны</p>	<p>После настройки: koi8-г</p>	<p>Довольно мощный, но интерфейс не очень удобен; требует настройки; многоуровневые папки для хранения почты; проблемы с 8-битными заголовками; Не очень удобен для работы с кириллицей</p>
Forte Agent	После настройки:	После настройки:	Один из лучших newsreader'ов, который

(not Free!)	меню перекодировок	koi8-г, koi8-и и др.	может использоваться и как почтовая программа; одноуровневые папки для хранения почты; о настройке см. www.glasnet.ru/~kazarn/soft.htm
	В зависимости от настройки, может использоваться внутренняя либо внешняя перекодировка		

Eudora Pro (v.2.2, 3.* ,4)

Eudora 'ничего не знает' о кириллице и сама по себе не производит никакой перекодировки кириллического текста. Поэтому для чтения и отправки писем в кодировке КОИ-8 приходится либо устанавливать шрифты и раскладки клавиатуры типа КОИ-8, либо пользоваться 'внешней' перекодировкой. Настройка этой почтовой программы для переписки по-русски заключается в задании для отображения текста кириллических шрифтов (типа koi-8 или Windows-1251) через меню Tools/ Options/ Fonts & Display. Режим преобразования исходящих 8-битных сообщений в 7-битную форму (quoted-printable) контролируется через меню Options/ Sending Mail/ May Use Quoted-Printable. Если 8-битный текстовый файл посылается как приложение, надо обратить внимание на опцию 'Put text attachments in body of message' в меню Tools/ Options/ Attachments : если она включена, текст будет вставлен в тело сообщения, а если выключена, то приложен в виде отдельного файла.

Проблемы.

Пожалуй, самое привлекательное в Eudora - ее удобный пользовательский интерфейс; но эта почтовая программа имеет и ряд недостатков:

- В заголовке исходящих сообщений на кириллице Eudora указывает неверную кодировку ('charset=iso-8859-1' вместо 'koi8-r' or 'Windows-1251'). В некоторых случаях, такие сообщения неправильно обрабатываются почтовой программой адресата. Для версии 2.2 это можно исправить редактированием файла Eudora.exe (см., например, [здесь](http://www.nagual.pp.ru/~ache/koi8.html) или Веб-страницу Андрея Чернова на www.nagual.pp.ru/~ache/koi8.html). Однако не нужно изменять объявленную кодировку сообщений на koi8-г, если на самом деле Вы собираетесь отсылать их в кодировке Windows-1251!! (Для версии 3.* или 4.* лучше не редактировать файл Eudora.exe, а использовать koi8-plugin, речь о котором будет идти ниже).
- Eudora плохо работает со шрифтами типа Unicode: обычно для того, чтобы нормально читать кириллицу, нужно использовать 'старые' неюникодовские либо экранные системные шрифты.
- Иногда также случается, что в окне составления письма Eudora блокирует работу 'родного' переключателя клавиатуры Win95 (для того, чтобы переключиться на русский язык, приходится переходить в другое окно Eudora, переключаться там и потом возвращаться в окно составления письма). Эту проблему обычно удается решить путем замены шрифта на неюникодовый либо системный шрифт, но радикальное решение - использовать вместо 'родного' переключателя клавиатуры Win95 другую программу (например, ParaWin 95/98, CyrWin 95 и т.п.).

Стоит отметить еще один недостаток Eudora 2.2, не связанный с обработкой кириллицы. Eudora автоматически декодирует файлы, приложенные к входящим письмам, и сохраняет их в заданном каталоге (обычно в EudoraAttach). При этом Eudora 2.2 часто не может автоматически (по двойному щелчку мыши) открывать приложенный файл с длинным именем, содержащим пробелы (в таких файлах, например, приходят сообщения, переправленные из Netscape Mail), и его приходится открывать 'вручную', что довольно неудобно. В Eudora v.3.* этот недостаток исправлен. Автоматическое декодирование приложенных двоичных файлов также неудобно, когда приходит файл в UUencode, 'нарезанный' на несколько кусочков. При этом первый кусочек автоматически декодируется и сохраняется под именем исходного двоичного файла (фактически будучи лишь его 'огрызком'), а остальные куски остаются неперекодированными. Для того, чтобы восстановить

оригинальный файл, нужно закодировать этот 'огрызок' обратно в UUencode, 'приклеить' к нему по порядку все последующие куски и потом уже декодировать 'склеенный' файл.

Еще один недостаток Eudora - это невозможность читать или набирать письма во время отсыдки или приема почты.

Важное преимущество **Eudora Pro v.3.*** над предыдущими версиями Eudora - это поддержка нескольких 'personalities' с различными установками (имя пользователя, pop3 и smtp сервер и другие установки), а также поддержка дополнительных модулей (plugin'ов). Обработка кириллицы в ней идет практически так же, как и в v.2.2, включая указание в заголовке сообщений неверной кодировки iso-8859-1. Однако в этом случае лучше не редактировать файл Eudora.exe, а использовать koi8-plugin, разработанный Евгением Суroveгиным (см. www.glasnet.ru/~ebs/koi8plugin.html). Этот plugin добавляет в меню 'Edit/ Message plug-ins' команды для перекодировки текста из koi8-r в windows-1251 и наоборот; кроме того, он указывает кодировку исходящих сообщений как koi8-r или windows-1251 вместо iso-8859-1.

koi8-plugin может использоваться и с Eudora Pro v.4.*; об ошибках в обработке кириллицы, которые есть в v.4.* и их исправлении можно прочесть на www.glasnet.ru/~ebs/koi8plugin.html.

Из-за своей неспособности к 'внутренней' перекодировке Eudora часто используется с pop3/smtp прокси-серверами, которые перекодируют входящие и исходящие сообщения на кириллице из koi8-r в Windows-1251 и наоборот. Если использовать Eudora v.3.* или 4.* с koi8-plugin'ом, может быть удобно получать почту неперекодированной, а отсылать сообщения через перекодирующий smtp прокси-сервер. При этом входящие письма можно в случае необходимости перекодировать koi8-plugin'ом, а исходящие будут автоматически перекодироваться прокси-сервером в koi8-r. (Однако в этом случае нужно внимательно посмотреть настройки koi8-plugin'a - если не выключить опцию 'Analyze charset when sending email', то письма могут уходить с указанием в заголовке кодировки Windows-1251, но их текст после перекодировки на smtp-сервере будет в koi8-r).

Netscape Mail (Mozilla) - v.2

Эта версия Netscape Mail уже устарела, и давно пора заменить ее на более новую. Она не производит никакой перекодировки кириллицы, и для чтения и отправки писем в КОИ-8 нужны шрифты и раскладки клавиатуры типа КОИ-8. Настройка - та же, что и для браузера Netscape Navigator (v.2). Кодировка кириллицы в заголовке исходящих писем указывается неверно.

Netscape Mail (Mozilla) - v.3

Настройка для переписки на кириллице аналогична настройке браузера Netscape Navigator (v.3). Если Document Encoding в меню Options установлено на 'Cyrillic (Win1251)', исходящие сообщения можно набирать в кодировке Windows-1251, при отправке они будут перекодированы в koi8-r, а в их заголовках будет указано 'charset=koi8-r'. Если настройка Document Encoding установлена на Cyrillic (KOI8-R), то исходящие письма нужно набирать в кодировке КОИ-8. Входящие письма НЕ перекодируются.

Режим отправки писем в 8-битной форме включается в Netscape Mail через меню *Options/ Mail and News Preferences/ Composition/ Allow 8-bit*.

Netscape Mail v.3 не следует использовать с pop3/smtp прокси-серверами, перекодирующими входящие и исходящие письма из koi8-r в Windows-1251 и обратно.

Когда Netscape Mail перенаправляет сообщение другому адресату (Forward), по умолчанию исходный текст пересылается в виде приложенного файла с длинным именем, определяемым заголовком

исходного сообщения. Для некоторых получателей такие письма создают проблемы; они будут Вам весьма признательны, если Вы не просто нажмете на кнопку 'Forward', а перенаправите письмо в процитированном виде, используя меню Message/ Forward Quoted.

Microsoft Internet Mail

Данная почтовая программа вместе с браузером Microsoft Internet Explorer (v.3) включена в состав операционной системы Windows95 (версия OSR2). Эта программа позволяет использовать Windows-1251 в качестве 'внутренней' кодировки, а отправлять письма в кодировке Windows-1251 или KOI8-R по выбору, причем с указанием правильной кодовой страницы. Кодировка, в которой Вы читаете входящие письма и отправляете исходящие, выбирается через меню *View/ Language*. (Если Вы отвечаете на полученное письмо, то эта настройка недоступна – письмо будет отправлено в той же кодировке, что и исходное). Чтобы пересылать кириллицу (в том числе в поле '*Subject*') в 8-битном виде, в меню *Mail/ Options/ Send* настройку '*Mail Sending Format*' нужно установить на '*Plain text*', при этом в '*Settings*' задать '*Message Format*'=*MIME* и '*Encode text using*'=*None*, а также включить опцию '*Allow 8-bit characters in headers*'.

Если Вы пользуетесь почтовой программой Microsoft Internet Mail, то Вам не нужно ни 'коизировать' Windows, ни работать через специальные перекодирующие прокси-серверы.

Netscape Messenger (Mozilla) v.4

В отличие от Netscape Mail v.3, для Netscape Messenger v.4.* не нужны шрифты и раскладки клавиатуры типа koi8. Он может сам перекодировать сообщения из различных кодировок через меню *View/ Encoding* или *View/ Character Set*. В вопросе отсылки сообщений есть одна тонкость: чтобы отослать письмо в кодировке koi8-r, для версий 4.0* нужно выбрать в окне составления сообщения кодировку Windows-1251 и набрать текст, используя обычные шрифты и раскладки клавиатуры; тогда при отправке письмо будет автоматически перекодировано в koi8-r. То, что произойдет при выборе в окне составления письма другой кириллической кодировки, зависит от версии Netscape. Для версий 4.0* при выборе кодировки koi8-r (или Western) почтовая програма 'думает', что текст уже набран в нужной перекодировке, и не преобразует его из Win-1251 в koi8-r. (Таким образом, получается, что для отправки письма в кодировке koi8-r нужно выбрать кодировку Win-1251, зато при выборе koi8-r письмо будет отправлено в кодировке Win-1251, но с предоставленным в заголовке сообщения `charset=koi8-r`, что может вызывать проблемы при чтении такого письма). Для версий 4.5*- 4.7* кириллическая кодировка, выбираемая при составлении письма, не имеет значения -- в любом случае письмо преобразуется в koi8-r. Таким образом, управление кодировкой исходящих сообщений в Netscape Messenger реализовано 'криво' и неполно.

Netscape Messenger v.4 не нужно использовать с перекодирующими pop3 и smtp серверами.

Чтобы заставить Netscape Messenger v.4.0* отсылать сообщения в 8-битном формате, без преобразования в quoted-printable, нужно в меню *Edit/ Preferences/ Mail & Newsgroups/ Messages/* установить переключатель '*Send messages that use 8-bit characters*' в положение '*As is*'. Кроме того, в меню *Edit/ Preferences/ Mail & Newsgroups/ Formatting/* рекомендуется отключить HTML-форматирование исходящих писем ('*Use the plain text editor to compose messages*').

Некоторые проблемы и их решение.

Проблема: Netscape Messenger v.4.5*- 4.7* отправляет Subject закодированным в 7-битную форму. Как заставить его отправлять Subject в обычном 8-битном формате?

Решение: Эту установку нельзя исправить через конфигурационные меню. Придется открыть для

редактирования конфигурационный файл `prefs.js` file (он должен быть в каталоге `...NetscapeUsers<пользователь>`) и дописать туда следующую строку:
`user_pref('mail.strictly_mime_headers', false);`
Имеет смысл также дописать строку
`user_pref('mailnews.start_page.enabled', false);`
которая отменяет автоматическое обращение Netscape Messenger'a к Netscape NetCenter для показа приветственного сообщения.

Более подробно настройка описана Павлом Городянским на <http://www.relcom.ru/Russification/WinNetscape/>.

Outlook Express

Outlook Express включен в состав Microsoft Internet Explorer v.4.* и 5.*, и его возможности по обработке кириллицы гораздо мощнее, чем у MS Internet Mail & News, включенного в IE v.3.*. Настраивается Outlook Express для работы с кириллицей через меню *Tools/ Options/ Read/ Fonts*, *Tools/ Options/ Send/ International Settings*, *Tools/ Options/ Send/ Mail (News) Sending Format* (здесь и далее названия меню указаны для версии 5; в версии 4 они могут несколько отличаться). Рекомендуется в меню *Tools/ Options/ Send/ International Settings* установить KOI8-R как кодировку по умолчанию, а в меню *Tools/ Options/ Send/ Mail (News) Sending Format* указать *'Plain text'*; в меню *Tools/ Options/ Send/ Mail (News) Sending Format/ Plain text settings* задать *'Message format' = 'MIME'*, *'Encode text using' = 'None'*, а также *'Allow 8-bit characters in headers'*. Если нужно отправлять приложенные двоичные файлы в формате UUencode вместо MIME/base64, выберите *'Message format' = 'Uuencode'* вместо *MIME*. Однако при этом в заголовке письма может не проставляться кодировка. При чтении или составлении писем (в том числе ответов), кодировку текста легко выбрать через меню *View/ Encoding* или *Format/Encoding*.

При отправке сообщений из Outlook Express нужно помнить, что внутренней кодировкой этой программы является Unicode, и текст сообщения в принципе может содержать любые символы, которые есть в Unicode. Кодировка, в которой сообщение будет отправлено, выбирается через меню *Format/Encoding* в окне составления сообщения. При преобразовании отправляемого текста из Unicode во 'внешнюю' кодировку возможна ситуация, когда некоторых символов, использованных в тексте письма, во 'внешней' кодировке просто нет. Такие символы будут заменены вопросительными знаками, и передаваемая ими информация будет безвозвратно утеряна. Часто такая ситуация возникает, когда получатель письма, написанного латинскими буквами, отвечает на него (либо переправляет кому-то со своими комментариями) на кириллице. По умолчанию кодировка ответа будет той же, что и кодировка исходного письма; если исходное письмо пришло, например, в кодировке iso-8859-1 или us-ascii, то ответ будет отправлен в той же кодировке (если ее не изменить 'вручную'), и все буквы кириллицы будут заменены вопросительными знаками. Чтобы этого не произошло, при отправке сообщений на кириллице (особенно в ответ на чужие сообщения) нужно проверять кодировку через меню *Format/Encoding*. При составлении нового сообщения по умолчанию обычно выбирается кодировка KOI8-R; если отправляется текст на украинском языке, то нужно не забыть изменить эту кодировку на другую, в которой есть украинские буквы (KOI8-U или Windows-1251), иначе они будут заменены вопросительными знаками.

Радикальным решением проблемы потери части символов при переходе от Unicode к внешней кодировке было бы использование 7- или 8-битных представлений Unicode UTF-7 или UTF-8 (при этом можно пересылать многоязычные тексты); однако пока что эти кодировки поддерживаются лишь немногими почтовыми программами (Outlook Express, Netscape Messenger, Forte Agent), да и большинство перекодировщиков их 'не понимают'. Поэтому использовать UTF-7 или UTF-8 стоит только в том случае, когда вы уверены, что ваш адресат сможет их прочесть.

Разумеется, при использовании Outlook Express не нужно подключаться к провайдеру через перекодирующие pop3 и smtp сервера - иначе письмо может быть перекодировано лишний раз.

18

Некоторые проблемы и решения

Проблема: Иногда текст нечитабелен даже если в меню *View/ Language* установлена правильная кодировка.

Решение: Попробуйте переключиться на другую кодировку (например, Win-1251 вместо koi8-r), а потом обратно на правильную.

Проблема: Некоторые (или все) символы кириллицы (особенно в поле Subject) при печати заменяются знаками вопроса, хотя на экране видны нормально.

Решение: Эта проблема (и некоторые другие) с распечаткой кириллицы может наблюдаться для некоторых бета-версий OE4 и OE5. Установите более новую версию программы. Можно попробовать также различные шрифты, хотя это и не всегда помогает. Кроме того, попробуйте в Internet Explorer (меню *Tools/Internet Options/General/Accessibility*) включить *'Ignore font styles specified on Web Pages'*.

Проблема: Иногда OE v.5.0 показывает русскоязычные заголовки в списке сообщений телеконференции в неправильной кодировке (не перекодированными из koi8 в win-1251). Переключение кодировки не помогает.

Решение: Это известная ошибка в OE v.5.0 (исправлена в OE v.5.01). Методы борьбы с ней в OE v.5.0 не всегда эффективны и несколько отдают 'шаманизмом' (см. следующий список предлагаемых действий):

- прежде чем загружать какие-либо новые сообщения, зайдите в Inbox или Outbox и просмотрите какое-нибудь сообщение, которое там уже есть (можно включить опцию *'When starting, go directly to my Inbox folder'* в меню *Tools/ Options/ General*);

- в меню *Tools/ Options/ Read/ Fonts* задать *Cyrillic (Windows)* как кодировку по умолчанию;

- в меню *Tools/ Options/ Read/ International Settings* задать *Cyrillic (KOI8-R)* как кодировку по умолчанию и включить опцию *'Use default encoding for all incoming messages'*.

- в меню *Tools/ Options/ Send/ International Settings* задать *Cyrillic (KOI8-R)* как кодировку по умолчанию.

- все это должно помочь (хотя иногда может и не помочь) только после перезагрузки списка сообщений в телеконференции (для этого можно, например, щелкнуть правой кнопкой мыши на название телеконференции в окошке список папок и телеконференций, а затем выбрать из контекстного меню *Properties/ Local File/ Reset*).

Радикальное решение проблемы -- установить Outlook Express v.5.01.

Проблема: При цитировании сообщения, на которое пишется ответ, OE иногда помечает цитируемый текст символом цитирования ('>'), а иногда нет.

Решение: Это обычно случается при ответе на сообщения, которые написаны в формате Rich-Text или HTML. Попросите автора отсылать сообщения в обычном текстовом формате (в Outlook Express 5, для этого нужно в меню *Tools/ Options/ Send/ Mail sending format* и *News sending format* задать *'Plain text'*).

Pegasus Mail for Windows

Почтовая программа Pegasus Mail (www.pegasus.usa.com/) невелика по размеру и бесплатна, но обладает довольно богатыми возможностями. В ней есть много различных настроек, которые весьма привлекательны для опытных пользователей, хотя и могут несколько запутать новичков. Pegasus Mail может использовать Windows-1251 как 'внутреннюю' кодировку и KOI8-R как 'внешнюю'. Однако для его настройки нужно отредактировать конфигурационные файлы, где заданы таблицы перекодировки. См. Веб-страницу Андрея Чернова (www.nagual.pp.ru/~ache/koi8.html). О русификации Пегасуса можно прочесть на <http://severov.atom.ru/guests/pegasus/>.

Pegasus Mail не позволяет отправлять кириллицу в поле Subject в 8-битном формате. Чтобы отправить тело письма в 8-битной форме, нужно включить опцию *'Allow 8-bit MIME message encoding'*

в меню *Tools/ Options/ Advanced Settings/* (при этом Pegasus Mail может предупредить Вас, что отправка текстов в 8-битном формате 'не вполне законна'). В окошке *'MIME character set'* нужно задать *'koi8-r'*.

The Bat!

[The Bat!](http://www.ritlabs.com/the_bat/) (www.ritlabs.com/the_bat/)- маленькая (инсталляционный файл занимает всего около полутора мегабайт) почтовая программа с очень богатыми возможностями (в том числе и по обработке кириллицы). В программе можно включить меню на нескольких языках (в том числе на русском), есть поддержка нескольких account'ов (доступ к каждому из которых можно защитить отдельным паролем), многоуровневая система папок, настраиваемые шаблоны сообщений, а также поддержка всех основных кодировок кириллицы (koi8-r, Win-1251, DOS-866, ISO-8859-5 и Mac) и даже нестандартных кодировок типа '1251 autoconverted from KOI8' и 'KOI8 autoconverted from 1251'. Кроме того, The Bat! поддерживает и украинскую кодировку koi8-u (с версии 1.18). К более ранним версиям программы поддержку koi8-u можно добавить с помощью дополнительной перекодировочной таблицы - см. [здесь](#)).

Пожалуй, The Bat! - одна из лучших почтовых программ (если не лучшая) для работы с кириллицей (даже несмотря на то, что поддержка юникодовских кодировок UTF-8 и UTF-7, введенная начиная с версии 1.42, односторонняя -- т.е. сообщения, приходящие в таком формате, автоматически декодируются, но сам The Bat! не может отправлять сообщения в UTF-8 и UTF-7 и, соответственно, не позволяет составлять 'многоязычные' письма, как Outlook Express).

The Bat! настраивается для '8-битной' переписки на кириллице через меню:

Account/ Properties:

Transport -- 8-bit characters are treated: Without changes.

Options -- Allow 8-bit characters in message headers.

Templates/ New message -- Use character set: Cyrillic (KOI-8).

Options/ Editor preferences/ Display/ Font -- задать кириллический шрифт.

Bmail/UUPC for DOS

В странах бывшего СССР многие все еще работают с электронной почтой при помощи программ типа Bmail/UUPC (или Demos Mail/UUPC) для DOS. Эти программы автоматически дозваниваются до почтового сервера, отправляют исходящие письма, принимают входящие и сразу же автоматически отключаются. В случае обрыва связи они могут при следующем соединении продолжить пересылку письма, начиная с того места, на котором связь прервалась. Благодаря этому, 'UUPC'-режим отправки и получения электронной почты может удовлетворительно работать и на плохих телефонных линиях (и, кроме того, обычно он значительно дешевле, чем on-line доступ к Интернету).

Если Вы хотите отправить письмо адресату, который пользуется программой типа Bmail/UUPC, надо иметь в виду, что такие программы НЕ понимают 7-битных представлений 8-битного текста в виде quoted-printable или base64, поэтому письма им нужно отправлять в 8-битной форме. Кроме того, такие программы автоматически декодируют лишь файлы-приложения, приложенные по методу UUENCODE.

Вообще, для минимизации проблем с Bmail, имеет смысл перейти на наиболее свежую версию **Demos Mail for DOS** -- там есть многие возможности, которых нет в Bmail.

В ранних версиях программы Bmail/UUPC (до 5.09g включительно) была жестко заложена перекодировка отправляемых писем из DOS-866 в KOI8-R и принимаемых -- из KOI8-R в DOS-866. В более поздних версиях заложена возможность выбора таблицы перекодировки. При этом программа

перекодирует по заданной ей таблице не все принимаемые письма, а только письма, в служебных полях которых указаны определенные кодовые страницы (charset). В результате письма с неправильно указанной кодовой страницей (либо письма типа 'multipart-mixed') часто попадают в почтовый ящик неперекодированными. На практике часто бывает удобнее пользоваться такой версией, которая перекодировала бы все письма без учета указанного в них charset'a. Из новых версий Vmail/UUCP, с этой точки зрения удобнее всего версия 6.18m (она не перекодирует только текст в письмах типа 'multipart-mixed').

Современные почтовые программы типа Outlook Express, The Bat!, Netscape Messenger, Eudora и т.п. сами по себе не поддерживают передачу электронной почты по протоколу UUCP (UUCP означает 'Unix-to-Unix Copy Protocol', а UUPC -- реализация UUCP для PC), однако существуют специальные программы, играющие роль 'шлюзов' между протоколами передачи почты POP3/SMTP и UUCP. При этом можно быть подключенным к провайдеру по протоколу UUCP, но для чтения и составления писем использовать 'нормальную' почтовую программу. Одна из таких программ-'шлюзов' -- **UUCP Connect** (www.novgorod.net/~kvantsoft/), а другая -- **Mailserver** (насколько я знаю, автор Mailserver'a прекратил работу над этой программой в 1998 г.; ее последнюю бесплатную версию можно найти здесь: www.geocities.com/g_naumovets/mlsrv212.zip).

Перекодировка текста с помощью прокси-сервера

Во времена, когда не было почтовых программ, способных самостоятельно перекодировать текст, например, из koï8 в windows-1251 и обратно, некоторые провайдеры установили у себя специальные перекодирующие прокси-серверы, чтобы помочь своим клиентам решать проблемы с перекодировкой кириллицы и сделать ненужной 'коизацию' Windows. Если указать в настройках почтовой программы такие серверы в качестве POP3- и SMTP-серверов, то проходящая через них почта будет автоматически перекодироваться по заданной провайдером таблице. О функции таких серверов часто можно догадаться по их названию (например, сервер pop-1251.gu.net перекодирует входящие письма из koï8-r в Windows-1251, сервер smtp-1251.gu.net перекодирует исходящие письма из Windows-1251 в koï8-r, а сервер news-1251.gu.net выполняет аналогичную перекодировку при работе с телеконференциями).

Однако если письмо приходит не в той кодировке, на которую рассчитан прокси-сервер (например, Windows-1251 вместо koï8-r), либо с неправильно указанной кодовой страницей (если данный прокси-сервер обращает внимание на поле charset), то оно будет перекодировано неправильно. Поэтому тем, кто работает через перекодирующий прокси-сервер, наверняка придется держать под рукой и специальные перекодирующие программы, которые смогут исправить последствия неправильной перекодировки.

При работе с 'умной' почтовой программой типа Outlook Express или The Bat! использовать перекодирующие прокси-серверы нет никакого смысла -- наоборот, это может только породить письма в 'извращенных' кодировках.

Перекодировка текстов с помощью специальных программ

Некоторые перекодирующие программы для DOS

Для перекодировки кириллицы разработано много различных программ. Если Вы умеете работать с программами для DOS, то можно порекомендовать маленькую программку под названием Recode,

которую можно найти, например, на <http://www.kiarchive.ru/pub/archive-tools/msdos/convert/char/>>. Кроме обычной для программ такого рода функции преобразования текста из одной кодовой страницы кириллицы в другую, эта программка позволяет еще и восстанавливать исходную 8-битную форму текста, закодированного по методу Quoted-printable или с помощью специальных команд языка HTML (в последнем случае текст может выглядеть примерно так: 'Âèñîêîîáð' и т.п.). Эта программа особенно полезна для тех, кто пользуется почтовыми системами типа Vmail/UUPC.

CP Tuner 97, CP Tuner 2000

CP Tuner 97/2000 (www.smartlinkcorp.com) может работать как plugin для Word 7.0 и Word 97, а также как самостоятельная программа. В нее заложены различные схемы перекодировки текста (в том числе многократной), которые могут встретиться на практике. Если в меню CP Tuner 97 *Tools/Options/* выбрать режим *'On Decode command: Autodetect the best named scheme and use it'*, то по команде *Tools/Decode* программа попытается автоматически преобразовать текст в кодировку Windows 1251. В большинстве случаев (включая и описанный в предыдущем параграфе), автоматический режим перекодировки срабатывает успешно. Если нет, можно добавить к имеющимся схемам перекодировки свои собственные. CP Tuner -- это платная программа (с www.smartlinkcorp.com можно загрузить пробную версию с ограниченным сроком действия). Однако в Интернете можно найти и множество бесплатных перекодировщиков (см., например, www.download.ru), многие из которых по своим возможностям не уступают CP Tuner'у.

Tot-Recode II и некоторые другие перекодировщики

По принципу и возможностям [Tot-Recode II](http://smartline.ru/recode/) (<http://smartline.ru/recode/>) довольно сходен с CP Tuner 97 (хотя пользовательский интерфейс выглядит немного иначе). Важное преимущество Tot-Recode II - то, что он бесплатен.

Он может также декодировать quoted-printable в 8-битный формат (в ручном режиме; удобно создать новую схему, включающую преобразование из QP в 8-bit и из KOI в Win).

Как и CP Tuner 97, он может автоматически 'транслитерировать' кириллицу (т.е. заменять буквы кириллицы 'фонетически сходными' латинскими буквами или их комбинациями).

На www.download.ru можно найти и множество других перекодировщиков. Из них имеет смысл выбрать такой, который поддерживает кодировки UTF-7 и UTF-8 (например, Russian Anywhere - см. <http://private.peterlink.ru/dll/rus/re.html>)

Онлайновые перекодировщики

Некоторые перекодировщики кириллицы доступны через Интернет -- см., например, Automatic Online Decoder (www.design.ru/free/decoder) или Ilya Sandler's Universal Cyrillic Converter (www.friends-partners.org/~isandler/cyrconv/cyrconv.html).

Литература

Веб-страница Андрея Чернова: 'Библия' по 'коизации' различных операционных систем и программного обеспечения <<http://www.nagual.pp.ru/~ache/koi8.html>>

How to Russify Netscape Navigator for Windows

<<http://www.relcom.ru/Russification/WinNetscape>> -- WinNetscape Russification Bible by Paul Gorodyanski

Paul Gorodyansky's Web page <http://ourworld.compuserve.com/homepages/PaulGor/> - Russification of Netscape, Office 97 and other helpful stuff

Konstantin Kazarnovskii about fonts and languages <<http://shlimazl.newmail.ru/fontsmain.htm>> -- very educational and well written, definitely worth reading

Lots of helpful information on Russification of various programs and operation systems <<http://www.siber.com/sib/russify/>>

The Cyrillic Charset Soup - о различных кодировках кириллицы <<http://czyborra.com/charsets/cyrillic.html>>

Cyrillic, Russian, Slavic and East European fonts and tools for Windows
Страница Кристофа Зингера на немецком, английском и русском языках по вопросам кириллизации и украинизации в Windows.
<<http://homepages.uni-tuebingen.de/student/christoph.singer/slovo/>>

Win95 Russification FAQ - <<http://www.hackzone.ru/rtw95>>

WinNT Russification - <<http://rwntug.quarta.msk.ru/questions.htm>>

Russification of Macintosh computers <<http://www.relcom.ru/Russification/MacKoi8-r>>

Настройка Forte Agent для работы с кириллицей
<<http://www.glasnet.ru/~kazarn/soft.htm>>
<<http://blue.iris.mipt.ru/timur/agent.htm>>

How mailers should handle Cyrillic messages (and how they actually do it)
<<http://blue.iris.mipt.ru/timur/trueway.htm>>

A lot of useful programs for working with Cyrillic texts <<http://www.kiarchive.ru/pub/cyrillic/>>

Some frequently asked questions about Russification of DOS and Windows
<<http://www.maths.monash.edu.au/~kig/ruscom/general/rusfaq.html>>

Installation of various keyboard layouts for native Windows95 keyboard switcher
<<http://www.netsight.net/~ryba/cyrillic/index.html>>

Internet standards (RFC -- requests for comments) <<http://www.rfc-editor.org>>
[RFC 1489](#): description of KOI8-R code page
[RFC 1521](#): MIME format for messages sent over the Internet
[RFC 1522](#): MIME format for non-ASCII characters in the message headers
[RFC 2319](#): Ukrainian Character Set KOI8-U

[RFC 2152](#): UTF-7 -- A Mail-Safe Transformation Format of Unicode
[RFC 2279](#): UTF-8, a transformation format of ISO 10646