

Национальный стандарт Российской Федерации

ГОСТ Р 7.0.29-2010

Система стандартов по информации, библиотечному и издательскому делу

**Электронные издания. Представление расширенного
кириллического алфавита для обмена информацией**

Дата введения – 01.07.2011



Федеральное агентство по техническому регулированию и метрологии

Предисловие

Цели и принципы стандартизации в Российской Федерации установлены Федеральным законом от 27 декабря 2002 г. № 184-ФЗ «О техническом регулировании», а правила применения национальных стандартов Российской Федерации – ГОСТ Р 1.0-2004 «Стандартизация в Российской Федерации. Основные положения».

Сведения о стандарте

1. Подготовлен Федеральным государственным бюджетным учреждением науки Всероссийским институтом научной и технической информации Российской академии наук (ВИНИТИ РАН).

2. Внесен Техническим комитетом по стандартизации ТК 191 «Научно-техническая информация, библиотечное и издательское дело».

3. Утвержден и введен в действие Приказом Федерального агентства по техническому регулированию и метрологии от 9 ноября 2010 г. № 346-ст.

4. Настоящий стандарт включает в себя идентичные основные нормативные положения следующих международных стандартов:

- раздел 5 – ИСО 5427:1984 «Расширение набора кодированных знаков кириллицы для обмена библиографической информацией» (ISO 5427:1984 «Extension of the Cyrillic alphabet coded character set for bibliographic information interchange»);
- раздел 6 – ИСО 10754:1996 «Расширение набора кодированных знаков кириллицы неславянских языков для обмена библиографической информацией» (ISO 10754:1996 «Information and documentation – Extension of the Cyrillic alphabet coded character set for non-Slavic languages for bibliographic information interchange»).

Сопоставление структуры и нумерации структурных элементов настоящего стандарта и основных нормативных положений указанных международных стандартов приведено во введении.

Применение указанных стандартов в настоящем стандарте обусловлено совпадением объекта стандартизации и принципов установления требований к нему.

5. Введен впервые.

Информация об изменениях к настоящему стандарту публикуется в ежегодно издаваемом информационном указателе «Национальные стандарты», а текст изменений и поправок – в ежемесячно издаваемых информационных указателях «Национальные стандарты». В случае пересмотра (замены) или отмены настоящего стандарта соответствующее уведомление будет опубликовано в ежемесячно издаваемом информационном указателе «Национальные стандарты». Соответствующая информация, уведомление и тексты размещаются также в информационной системе общего пользования – на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет.

Содержание

1. Область применения
 2. Нормативные ссылки
 3. Термины и определения
 4. Общие положения
 5. Символы алфавитов славянских языков
 6. Символы алфавитов неславянских языков
- Приложение А (справочное). Использование символов расширенного набора в кириллических алфавитах различных языков

Введение

Настоящий стандарт устанавливает расширение набора кодов для представления букв и других символов, используемых в алфавитах языков, продолжающих традицию славянской письменности, идущую от свв. Кирилла и Мефодия в IX в. н.э. В число этих языков входит русский язык, другие восточнославянские и южнославянские языки, монгольский язык и множество языков народов бывшего Советского Союза.

Эти алфавиты включают в себя основные буквы современного русского языка, коды которых установлены ГОСТ 27463-87, а также дополнительные символы, коды которых устанавливает настоящий стандарт в соответствии с требованиями двух международных стандартов.

В разделе 5 настоящего стандарта установлена кодовая таблица для символов славянских языков (включая их старославянские алфавиты), идентичная кодовой таблице ИСО 5427:1984. В разделе 6 настоящего стандарта установлена кодовая таблица для алфавитов неславянских языков, идентичная таблице 1 ИСО 10754:1996.

В совокупности со стандартом ГОСТ 27463-87 коды, установленные настоящим стандартом, достаточны для передачи текстовых данных на языках, указанных в приложении А, а также в других случаях, когда набор необходимых символов не выходит за пределы кодовых таблиц настоящего стандарта.

ГОСТ Р 7.0.29-2010

Система стандартов по информации, библиотечному и издательскому делу Электронные издания. Представление расширенного кириллического алфавита для обмена информацией

1. Область применения

1.1. Настоящий стандарт устанавливает набор из 135 графических символов, включая их наименования, обозначения и 7-битные коды, а также указания об их использовании – для обмена информацией на машиночитаемых носителях и в информационных сетях. Данный набор символов используется совместно с набором КОИ-7 НИ по ГОСТ 27463. При использовании 8-битного кода настоящий стандарт применяется совместно с ГОСТ 19768.

1.2. Эти символы вместе с основным набором кириллических символов, зарегистрированным под номером 37 в международном регистре ИСО, составляют набор символов для международного обмена информацией, включая библиографические данные, аннотации и полные тексты документов, электронные издания на языках, указанных в 1.3.

1.3. Набор символов предназначен для представления текстовых данных и обмена ими на следующих языках славянской группы: русском, белорусском, украинском, болгарском, македонском, сербском, а также на других языках, которые используют алфавиты на русской графической основе (кириллические алфавиты) в пределах установленного здесь набора. Перечень соответствующих языков см. в приложении А.

1.4. Настоящий стандарт не включает в себя символы, использовавшиеся в алфавитах некоторых языков в XIX веке и ранее и к XX веку вышедшие из употребления.

1.5. Настоящий стандарт не устанавливает типы шрифтов графических символов и не регламентирует методы обнаружения и обработки ошибок кодового представления символов.

1.6. Настоящий стандарт предназначен для органов научно-технической информации, а также для других организаций и предприятий, осуществляющих научно-информационную деятельность.

2. Нормативные ссылки

В настоящем стандарте использованы нормативные ссылки на следующие стандарты:

- ГОСТ 19768-93 Информационная технология. Наборы 8-битных кодированных символов. Двоичный код обработки информации;
- ГОСТ 27463-87 Системы обработки информации. 7-битные кодированные наборы символов;
- ГОСТ 27466-87 Системы обработки информации. Наборы символов в 7-и 8-битных кодах. Методы расширения кодов.

Примечание – При пользовании настоящим стандартом целесообразно проверить действие ссылочных стандартов в информационной системе общего пользования – на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет или по ежегодно издаваемому информационному указателю «Национальные стандарты», который опубликован по состоянию на 1 января текущего года, и по соответствующим ежемесячно издаваемым информационным указателям, опубликованным в текущем году.

Если ссылочный стандарт заменен (изменен), то при пользовании настоящим стандартом следует руководствоваться заменяющим (измененным) стандартом. Если ссылочный стандарт отменен без замены, то положение, в котором дана ссылка на него, применяется в части, не затрагивающей эту ссылку.

3. Термины и определения

В настоящем стандарте применены следующие термины с соответствующими определениями.

3.1. **кириллический алфавит**: Алфавит, основанный на использовании русских букв и некоторых дополнительных символов и восходящий к традиции письменности, установленной просветителями славянских народов свв. Кириллом и Мефодием в IX в.

3.2. **символ алфавита**: Элемент письменности, представленный на машиночитаемом носителе 7-битным кодом и визуализируемый как изображение определенной буквы или другого орфографического знака.

4. Общие положения

4.1. Настоящий набор кодов применяется для записи текстовых данных на носителях информации и для передачи по линиям связи в соответствии с национальными и международными стандартами на обработку данных, которые, в частности, определяют меры обнаружения и исправления ошибок.

4.2. Применение настоящего стандарта осуществляется совместно с ГОСТ 27466. Кодирование символов алфавитов славянских языков определено в разделе 5. Кодирование символов алфавитов неславянских языков определено в разделе 6.

4.3. В кодовых таблицах 1 и 2 внутри жирной рамки указаны изображения символов алфавитов, а в боковике и головке таблиц – значения битов соответствующего кода $b_7b_6b_5b_4b_3b_2b_1$, где $[b_n$ – значения n-го бита (ноль или единица) при счете, начиная от младшего разряда]. Коды, соответствующие незаполненным клеткам таблицы, для представления алфавитных символов не используют. В пятой колонке боковика указаны 16-ричные значения первых четырех битов; в четвертой строчке – 8-ричные значения последних трех битов кода.

4.4. Идентификация настоящего набора графических символов в потоке данных осуществляется переключающими последовательностями согласно ГОСТ 27466.

5. Символы алфавитов славянских языков

5.1. Коды символов для расширенных алфавитов славянских языков (русского, белорусского, болгарского, македонского, сербского и украинского) устанавливает кодовая таблица 1.

5.2. Для идентификации настоящего набора символов в потоке данных установлены следующие переключающие последовательности:

- AP2 2/8 3/2 – в качестве набора ГО;
- AP2 2/9 3/2 – в качестве набора Г1.

5.3. В таблице 3 указаны наименования символов и языки, в алфавите которых они употребляются.

6. Символы алфавитов неславянских языков

6.1. Коды символов для расширенных алфавитов неславянских языков (указанных в приложении А) устанавливает кодовая таблица 2.

6.2. В колонках 2 и 3 таблицы 2 (позиции 21-27 и 31-37) указаны диакритические знаки, которые используются в комбинации с одним из остальных символов в качестве надстрочного или подстрочного дополнения.

В алфавитах неславянских языков иногда не соблюдаются точное написание и различение подстрочных знаков «хвостик направо», «хвостик налево», «огонек», «левый огонек», а также надстрочных знаков «верхняя запятая», «апостроф», «слабое ударение», но

они все приведены в таблицах 2 и 4, с тем чтобы обеспечить передачу орфографически правильного текста.

Таблица 1. Коды символов расширенного набора для алфавитов славянских языков

				b ₇	0	0	0	0	1	1	1	1
				b ₆	0	0	1	1	0	0	1	1
				b ₅	0	1	0	1	0	1	0	1
b ₄	b ₃	b ₂	b ₁		0	1	2	3	4	5	6	7
0	0	0	0	0					г	ѣ	Ґ	Ђ
0	0	0	1	1					ђ	ѳ	Ђ	Ѳ
0	0	1	0	2					ѓ	ѵ	Ѓ	Ѵ
0	0	1	1	3					є	ж	Є	Ж
0	1	0	0	4					ë		Ё	
0	1	0	1	5					s		S	
0	1	1	0	6					i		I	
0	1	1	1	7					ï		Ї	
1	0	0	0	8					j		J	
1	0	0	1	9					љ		Љ	
1	0	1	0	10					њ		Њ	
1	0	1	1	11					ћ	[Ћ	
1	1	0	0	12					ќ		Ќ	
1	1	0	1	13					ђ]	Ђ	
1	1	1	0	14					џ		Џ	
1	1	1	1	15							Ђ	

6.3. Переключающие последовательности для идентификации в потоке данных настоящего набора кодов установлены Международным регистром наборов символов¹⁾.

Таблица 2. Коды символов расширенного набора для алфавитов неславянских языков

				b ₇	0	0	0	0	1	1	1	1
				b ₆	0	0	1	1	0	0	1	1
				b ₅	0	1	0	1	0	1	0	1
b ₄	b ₃	b ₂	b ₁		0	1	2	3	4	5	6	7
0	0	0	0	0					ı	Ɔ	ø	⊘
0	0	0	1	1		ó	ö		ĵ	Ɔ	œ	⊘
0	0	1	0	2		ö	ō		к	Ɔ	п̣	Ɔ
0	0	1	1	3		ȳ	ȳ		к	Ɔ	р̣	Ɔ
0	1	0	0	4		ȳ	ȳ		k	Ɔ	ç	Ɔ
0	1	0	1	5		ȳ	ȳ		Ɔ	Ɔ	т̣	Ɔ
0	1	1	0	6		ȳ	ȳ		q	Q	γ	Υ
0	1	1	1	7		ȳ	ȳ		l	L	Ƴ	Υ
1	0	0	0	8		æ	Æ		љ	Љ	w	W
1	0	0	1	9		Ɔ	Ɔ		љ̣	Љ̣	ц̣	Ɔ
1	0	1	0	10		ђ̣	Ђ̣		л̣	Љ̣	ч̣	Ɔ
1	0	1	1	11		d	đ		н	Н	h	h
1	1	0	0	12		đ	đ		ң	Ɔ	е̣	Ɔ
1	1	0	1	13		з̣	З̣		н̣	Н̣	ə	Ə
1	1	1	0	14		з̣	З̣		н̣̣	Н̣̣	я̣	Ɔ
1	1	1	1	15		з̣̣	З̣̣		ө	Ө	ı	

¹⁾ International Register of Characters Sets. Обратиться к нему можно через секретариат регистрационного агентства по адресу: ECMA, Rue du Rhône 114, CH-1204, Geneva, Switzerland.

6.4. В таблице 4 приведены наименования или описания символов для расширенных алфавитов неславянских языков.

Таблица 3. Наименования символов расширенного набора для алфавитов славянских языков

	Код в 16-значном представлении		Наименование	Язык	
	Строчная буква	Заглавная буква			
40	г	60	Г	взрывное Г	украинский, белорусский
41	ђ	61	Ђ	мягкое Д	сербский
42	ѓ	62	Ѓ	мягкое Г	македонский
43	є	63	Є	мягкое Е	украинский
44	ё	64	Ё	Ё (йо)	русский, белорусский
45	s	65	S	зепо	македонский
46	i	66	I	И с точкой	украинский, белорусский, русский
47	ï	67	Ï	И с двумя точками	украинский
48	j	68	J	йот	сербский, македонский
49	љ	69	Љ	мягкое Л	сербский, македонский
4A	њ	6A	Њ	мягкое Н	сербский, македонский
4B	ћ	6B	Ћ	мягкое Т	сербский
4C	ќ	6C	Ќ	мягкое К	македонский
4D	ў	6D	Ў	краткое У	белорусский
4E	џ	6E	Џ	звонкое Ч	сербский
		6F	Ъ	твердый знак	русский, болгарский
50	ѣ	70	Ь	ять	русский, болгарский
51	ѓ	71	Ѧ	фита	русский, болгарский
52	v	72	V	ижица	русский, болгарский
53	ж	73	Ж	юс	болгарский
5B	[скобка открывающая	все
5D]			скобка закрывающая	все
5F	—			нижняя черта	все

Таблица 4. Символы расширенного набора для алфавитов неславянских языков

	Код в 16-значном представлении			Описание, уточняющее характер изображения
	Строчная буква		Заглавная буква	
21	◌̇			ударение (акут)
22	◌̈			диерезис (трема, умляут); две точки над буквой
23	◌̋			огонэк; хвостик снизу буквы
24	◌̌			хвостик справа от буквы
25	◌̍			кратка (бревис); плавная дуга над буквой
26	◌̎			седиль
27	◌̏			слабое ударение (гравис)
28	æ	38	Æ	лигатура А и Е
29	ƒ	39	F	перечеркнутая Г
2A	ɸ	3A	Ɔ	Г с крюком
2B	ɔ	3B	ɔ	совпадает с латинской строчной d
2C	ɔ̣	3C	ɔ̣	латинская строчная d с хвостиком вверх
2D	z̄	3D	Z̄	З с плоской верхушкой
2E	z̋	3E	Z̋	З с хвостиком вниз
2F	ž	3F	Ž	З с хвостиком вверх
31	◌̈́			двойное ударение
32	◌̈̄			долгота (макрон); прямая черта сверху буквы
33	◌̣̋			левый огонэк; хвостик налево под буквой
34	◌̣̌			левый хвостик
35	◌̣̍			корона; ломаная черта над буквой
36	◌̣̎			кружок над буквой
37	◌̣̏			запятая над буквой
40	ı̇	50	İ	латинская I с чертой и точкой сверху
41	ȝ	41	J	перечеркнутый йот
42	ƙ	52	K	перечеркнутая К
43	ƙ̣	53	Ḳ	К с флажком слева
44	ƙ̣̣	54	Ḳ̣	К с крестом
45	ƙ̣̣̣	55	Ḳ̣̣	К с кружком сверху
46	q	56	Q	совпадает с латинской Q
47	l	57	L	совпадает с латинской L
48	ł	58	Ł	Л с хвостиком вверх
49	ł̣	59	Ł̣	Л с крюком
4A	ł̣̣	5A	Ł̣̣	лигатура Л и Х
4B	h	5B	H	лигатура Н и Г
4C	ḥ	5C	Ḥ	Н с флажком справа
4D	ḥ̣	5D	Ḥ̣	Н с хвостиком вверх
4E	ḥ̣̣	5E	Ḥ̣̣	Н с крюком
4F	ȯ	5F	Ȯ	перечеркнутая О
60	œ	70	Œ	напоминает рукописную заглавную Q
61	œ̣	71	Œ̣	лигатура О и Е
62	ɸ̣	72	Ɔ̣	П с крюком
63	ɸ̣̣	73	Ɔ̣̣	лигатура Р и К
64	ɸ̣̣̣	74	Ɔ̣̣̣	С с чертой; напоминает латинскую заглавную G
65	ɸ̣̣̣̣	75	Ɔ̣̣̣̣	Т с хвостиком вверх
66	у	76	Y	прямая У, совпадает с латинской Y
67	у̇	77	Ẏ	прямая У перечеркнутая
68	w	78	W	совпадает с латинской W
69	ц̣	79	Ц̣	Ц с крышкой
6A	ч̣	7A	Ч̣	перечеркнутая Ч
6B	ḥ	7B	Ḥ	совпадает с латинской строчной Н
6C	ѣ̣	7C	Ѧ̣	напоминает рукописную X
6D	э̣	7D	Ѧ̣̣	шва
6E	ѣ̣̣	7E	Ѧ̣̣	лигатура Я и Е
6F	ı̇			гуттуральный знак

Язык	Простые буквы и комбинации с диакритическими знаками, используемые в данном языке
Табасаранский	І
Таджикский	ғ қ х ч й ў
Татарский (казанский)	жң ө ү һ ә
Татарский (кряшенский)	ä ö ŷ н
Татский	І
Тувинский	ң ө ү
Туркменский	жң ө ү ә
Удмуртский	ж з й ө ч
Узбекский	ғ қ х ў
Уйгурский	ғ ж қң ө ү ә һ
Украинский	г і ї є
Хакасский	г і ј н, ö ŷ ч һ (ц)
Хантыйский (вахский)	ä ö ө ә э у қ һ
Хантыйский (казымский)	ä ө ә э пң
Хантыйский (сургутский)	ä ө ә э ў қң
Хантыйский (шурушкарский)	ң
Цыганский	рғ
Чеченский	Б һ h j к k н ц q һ х ц ч І
Чувашский	ă ě ё ъ ѣ њ њ с т г њ ъ ѱ
Чукотский	қң
Шорский	ј н ö ŷ
Эвенкийский	ң
Эвенский	ң ө ө
Эскимосский	г'к'л'н'х'ў ъ
Якутский	а о у ү ө ҕ н һ

Ключевые слова: набор кодированных символов, кириллический алфавит, расширенный набор символов, символы алфавитов, славянские языки, неславянские языки, обмен информацией, текстовые данные, представление данных

информация для всех



Электронная версия создана
МОО «Информация для всех»
www.ifap.ru