ФЕДЕРАЛЬНОЕ АГЕНТСТВО

ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ

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

(ISO 5427:1984, MOD)
(ISO 10754:1996, MOD)

Издание официальное

Москва
Стандартинформ
2019

gost.e-ecolog.ru

Предисловие

1 ПОДГОТОВЛЕН Учреждением Российской академии наук Всероссийским институтом научной и технической информации РАН (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», MOD);

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», MOD).

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

6 ПЕРЕИЗДАНИЕ. Декабрь 2018 г.

*Правила применения настоящего стандарта установлены в статье 26 Федерального закона от 29 июня 2015 г. № 162-ФЗ «О стандартизации в Российской Федерации». Информация об изменениях к настоящему стандарту публикуется в ежегодном (по состоянию на 1 января текущего года) информационном указателе «Национальные стандарты», а официальный текст изменений и поправок — в ежемесячном информационном указателе «Национальные стандарты». В случае пересмотра (замены) или отмены настоящего стандарта соответствующее уведомление будет опубликовано в ближайшем выпуске ежемесячного информационного указателя «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 (*gost.e-ecolog.ru

© Стандартинформ, оформление, 2011, 2019

Настоящий стандарт не может быть полностью или частично воспроизведен, тиражирован и распространен в качестве официального издания без разрешения Федерального агентства по техническому регулированию и метрологии

Содержание

1 Область применения 1
2 Нормативные ссылки 1
3 Термины и определения 2
4 Общие положения 2
5 Символы алфавитов славянских языков 2
6 Символы алфавитов неславянских языков 2
Приложение А (справочное) Использование символов расширенного набора в кириллических алфавитах различных языков 8

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 входят 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 старописьменные 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.

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

Система стандартов по информации, библиотечному и издательскому делу

ЭЛЕКТРОННЫЕ ИЗДАНИЯ

Представление расширенного кириллического алфавита для обмена информацией

System of standards on information, librarianship and publishing. Electronic publishing.
Representation of extended Cyrillic alphabet for information interchange

Дата введения — 2011—07—01

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 NI 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 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 ссылки на следующие стандарты:

ГОСТ 19768 Информационная технология. Наборы 8-битных кодированных символов. Двоичный код обработки информации

ГОСТ 27463 Системы обработки информации. 7-битные кодированные наборы символов

ГОСТ 27466 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 ежегодному информационному указателю «Национальные стандарты», который опубликован по состоянию на 1 января текущего года, и по выпускам

Издание официальное

ежемесячного информационного указателя «Национальные стандарты» за текущий год. Если заменен ссылочный стандарт, на который дана недатированная ссылка, то рекомендуется использовать действующую версию этого стандарта с учетом всех внесенных в данную версию изменений. Если заменен ссылочный стандарт, на который дана датированная ссылка, то рекомендуется использовать версию этого стандарта с указанным выше годом утверждения (принятия). Если после утверждения настоящего стандарта в ссылочный стандарт, на который дана датированная ссылка, внесено изменение, затрагивающее положение, на которое дана ссылка, то это положение рекомендуется применять без учета данного изменения. Если ссылочный стандарт отменен без замены, то положение, в котором дана ссылка на него, рекомендуется применять 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 в соответствии 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 3 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:

АР2 2/8 3/2 — в качестве набора Г0;

АР2 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 4 (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					ґ	ь	Ґ	Ь
0	0	0	1	1					ђ	ѳ	Ђ	Ѳ
0	0	1	0	2					ѓ	v	Ѓ	V
0	0	1	1	3					є	ж	Є	Ж
0	1	0	0	4					ё		Ё	
0	1	0	1	5					s		S	
0	1	1	0	6					і		I	
0	1	1	1	7					ї		Ї	
1	0	0	0	8					j		J	
1	0	0	1	9					љ		Љ	
1	0	1	0	A					њ		Њ	
1	0	1	1	B					ћ	[	Ћ	
1	1	0	0	C					ќ		Ќ	
1	1	0	1	D					ў	]	Ў	
1	1	1	0	E					џ		Џ	
1	1	1	1	F						_	Ъ	

6.3 Переключающие последовательности для идентификации в потоке данных настоящего набора кодов установлены Международным регистром набором символов[1)].

[1)] International Register of Charactres Sest. Обратиться к нему можно через секретариат регистрационного агентства по адресу: ECMA, 114 Rue du Rone, CH01294, Geneve, Suisse (Switzerland).

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					ī	Ī	ɷ	ɷ
0	0	0	1	1			ǘ	ǘ	ɟ	ɟ	œ	Œ
0	0	1	0	2			ǖ	ǖ	к	К	ҧ	Ҧ
0	0	1	1	3			Ǘ	Ǘ	ҡ	Ҡ	ҩ	Ҩ
0	1	0	0	4			ǘ	ǘ	ҟ	Ҟ	ҫ	Ҫ
0	1	0	1	5			ǚ	ǚ	ҝ	Ҝ	ҭ	Ҭ
0	1	1	0	6			Ǘ	Ǘ	ҷ	Ҷ	ҳ	Ҳ
0	1	1	1	7			ǜ	ǜ	ӏ	Ӏ	ҵ	Ҵ
1	0	0	0	8			æ	Æ	ҧ	Ҧ	ҩ	Ҩ
1	0	0	1	9			ҥ	Ҥ	ҧ	Ҧ	ҭ	Ҭ
1	0	1	0	A			ҕ	Ҕ	ҧ	Ҧ	ҷ	Ҷ
1	0	1	1	B			d	d	ҥ	Ҥ	h	h
1	1	0	0	C			ɗ	ɗ	ҥ	Ҥ	ҽ	Ҽ
1	1	0	1	D			ʒ	Ʒ	ҥ	Ҥ	ə	Ə
1	1	1	0	E			ʒ	Ʒ	ҥ	Ҥ	ҙ	Ҙ
1	1	1	1	F			ʒ	Ʒ	ҩ	Ҩ	Ӏ	

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	Строчная буква	Код в 16-значном представлении	Заглавная буква	Наименование	Язык
40	ґ	60	Ґ	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	ѕ	65	Ѕ	Зело	Македонский
46	і	66	І	И с точкой	Украинский, белорусский, русский
47	ї	67	Ї	И с двумя точками	Украинский
48	ј	68	Ј	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	Ѣ	Ять	Русский, болгарский
51	ѳ	71	Ѳ	Фита	Русский, болгарский
52	ѵ	72	Ѵ	Ижица	Русский, болгарский
53	ѫ	73	Ѫ	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	Строчная буква	Код в 16-значном представлении	Заглавная буква	Описание, уточняющее характер изображения
21	◌́			Ударение (акут)
22	◌̈			Диерез (трема, умляут). Две точки над буквой
23	◌̨			Огонек. Хвостик снизу буквы
24	◌̩			Хвостик справа от буквы
25	◌̆			Кратка (бревис). Плавная дуга над буквой
26	◌̧			Седиль
27	◌̀			Слабое ударение (гравие)
28	æ	38	Æ	Лигатура А и Е
29	ғ	39	Ғ	Перечеркнутая Г
2A	ҕ	3A	Ҕ	Г с крюком
2B	d	3B	d	Совпадает с латинской строчной d
2C	d	3C	d	Латинская строчная d с хвостиком вверх
2D	ҙ	3D	Ҙ	З с плоской верхушкой
2E	ʒ	3E	Ʒ	З с хвостиком вниз
2F	ȝ	3F	Ȝ	З с хвостиком вверх
31	◌̋			Двойное ударение
32	◌̄			Долгота (макрон). Прямая черта сверху буквы
33	◌̨			Левый огонек. Хвостик налево под буквой
34	◌̦			Левый хвостик
35	◌̌			Корона. Ломаная черта над буквой
36	◌̊			Кружок над буквой
37	◌̓			Запятая над буквой
40	i̇̄	50	İ̄	Латинская I с чертой и точкой сверху
41	j̄	41	J̄	Перечеркнутый йот
42	к	52	К	Перечеркнутая К

Окончание таблицы 4

Код в 16-значном представлении	Строчная буква	Код в 16-значном представлении	Заглавная буква	Описание, уточняющее характер изображения
43	ҡ	53	Ҡ	К с флажком слева
44	ҟ	54	Ҟ	К с крестом
45	ҝ	55	Ҝ	К с кружком сверху
46	ԛ	56	ԚQ	Совпадает с латинской Q
47	l	57	L	Совпадает с латинской L
48	лҳ	58	Лҳ	Л с хвостиком вверх
49	лһ	59	Лһ	Л с крюком
4A	лк	5A	Лк	Лигатура Л и К
4B	нг	5B	Нг	Лигатура Н и Г
4C	нһ	5C	Нһ	Н с флажком справа
4D	нҳ	5D	Нҳ	Н с хвостиком вверх
4E	нһ	5E	Нһ	Н с крюком
4F	⊖	5F	⊖	Перечеркнутая О
60	ԛ	70	Ԛ	Напоминает рукописную заглавную Q
61	œ	71	Œ	Лигатура О и Е
62	пһ	72	Пһ	П с крюком
63	рк	73	Рк	Лигатура Р и К
64	ҫ	74	Ҫ	С с чертой, напоминает латинскую заглавную G
65	тҳ	75	Тҳ	Т с хвостиком вверх
66	ү	76	Ү	Прямая У, совпадает с латинской Y
67	ұ	77	Ұ	Прямая У перечеркнутая
68	w	78	W	Совпадает с латинской w
69	тц	79	Тц	Ц с крышкой
6A	ч	7A	Ч	Перечеркнутая Ч
6B	h	7B	h	Совпадает с латинской строчной h
6C	ҽ	7C	Ҽ	Напоминает рукописную Ч
6D	ә	7D	Ә	Шва
6E	яе	7E	Яе	Лигатура Я и Е
6F	I			Гуттуральный знак

Приложение А
(справочное)

Использование символов расширенного набора в кириллических алфавитах различных языков

Таблица А.1

Язык	Простые буквы и комбинации с диакритическими знаками, используемые в данном языке
Абазинский	I
Абхазский	ҕ, ҷ, ҽ, ҿ, ҙ, ҟ, ҡ, ҩ, ҧ, ҥ, ҭ, ҳ, ҵ, ҷ, ә
Аварский	I
Адыгейский	I
Азербайджанский	ғ, ә, j, қ, ө, ү, h, ч
Айсорский	d, ə, j, l, q, s, h
Алтайский	j, ң, ӧ, ӱ
Башкирский	ғ, ҙ, ҡ, ң, ө, ҫ, ү, h, ә
Белорусский	і, ў, ё, ґ
Болгарский (до 1945 г.)	і, ѣ, v, ѳ, ж
Бурятский	й, ө, ӱ, ү, h
Даргинский	I
Дунганский	ә, җ, ң, ў, ү
Ингушский	I
Кабардинский	I
Казахский	ә, ғ, і, ҕ, ң, ө, ү, ұ, h
Калмыцкий	ä, ӧ, у, җ, ң, ө, ү, h, ə
Каракалпакский	ә, ғ, қ, ң, ө, ү, ў, х
Карачайский	ӳ
Карельский	ä, ӧ, ӱ
Киргизский	ң, ө, ү
Коми	і, ӧ
Корякский	ӷ, ӈ
Курдский	г', ә, ə', к', ӧ, п', р', т', h, h', ч', q, w
Лакский	I
Лезгинский	I
Македонский	ѓ, s, j, ќ
Мансийский	ң, (н')
Марийский (горный)	ӓ, ӧ, ӱ, ӈ, ӹ

Окончание таблицы А.1

Язык	Простые буквы и комбинации с диакритическими знаками, используемые в данном языке
Марийский (луговой)	ӓ, ӧ, ӱ, ӈ
Молдавский	ӝ
Монгольский	ө, ү
Мордовский (мокша)	ӕ, ӗ, лк, ӈ, ӧ, ҏ, ӹ, ӭ
Мордовский (эрзя)	ӈ
Ненецкий	ӊ
Осетинский	æ
Русский	ё
Русский (до 1918 г.)	ё, і, ѣ, v, ѳ
Саамский	ӊ
Селькупский	å, œ
Сербский	ђ, ћ, љ, њ, џ, ј
Табасаранский	І
Таджикский	ғ, қ, ҳ, ҷ, ӣ, ӯ
Татарский (казанский)	җ, ң, ө, ү, h, ә
Татарский (кряшенский)	ӓ, ӧ, ӱ, ӈ
Татский	І
Тувинский	ң, ө, ү
Туркменский	җ, ӈ, ө, ү, ә
Удмуртский	ӝ, ӟ, й, ӧ, ӵ
Узбекский	ғ, қ, ҳ, ӯ
Уйгурский	ғ, җ, қ, ң, ө, ү, ә, h
Украинский	ґ, і, ї, є
Хакасский	ғ, і, ј, ӈ, ӧ, ӱ, ҷ, h, (ц)
Хантыйский (вахский)	ӓ, ӧ, ӫ, ө, ә, ә, у, қ, ӊ
Хантыйский (казымский)	ӓ, ӫ, ө, ӭ, ә, п’, ң
Хантыйский (сургутский)	ӓ, ӫ, ө, ә, ӱ, қ, ң
Хантыйский (шурушкарский)	ӊ
Цыганский	р̆, ғ
Чеченский	ҕ, Һ, h, ј, қ, ḱ, ң, ц, q, ҭ, х̆, ц̆, ч̆, І
Чувашский	ӑ, ӗ, ӗ, җ, ӈ, ӊ, ӊ, ҫ, ҭ, т̆, ҭ, ӳ, ӱ
Чукотский	қ, ӊ
Шорский	ј, ӈ, ӧ, ӱ
Эвенкийский	ӊ
Эвенский	ӊ, ө, ӫ
Эскимосский	г', к', л', н', х', ӯ, (җ)
Якутский	ӓ̄, ӧ̄, ӱ̄, ү, ө, ҕ, ӈ, h

УДК 025.43.026.036:006.354 ОКС 01.140.40 Т62

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

Редактор *Н.В. Таланова*
Технический редактор *В.Н. Прусакова*
Корректор *Е.Д. Дульнева*
Компьютерная верстка *И.А. Налейкиной*

Сдано в набор 19.12.2018. Подписано в печать 23.01.2019. Формат 60×84 1/8. Гарнитура Ариал.
Усл. печ. л. 1,86. Уч.-изд. л. 1,49.
Подготовлено на основе электронной версии, предоставленной разработчиком стандарта

ИД «Юриспруденция», 115419, Москва, ул. Орджоникидзе, 11.
gost.e-ecolog.ru gost.e-ecolog.ru

Создано в единичном исполнении ФГУП «СТАНДАРТИНФОРМ» для комплектования Федерального информационного фонда стандартов, 117418 Москва, Нахимовский пр-т, д. 31, к. 2.
gost.e-ecolog.ru gost.e-ecolog.ru

gost.e-ecolog.ru