

Екатерина Пиджакова, Надежда Казанцева¹

Ekaterina Pidzhakova, Nadezhda Kazantseva

**ОЦЕНКА РАЗМЕРА ИНФОРМАЦИИ В ЭЛЕКТРОННОЙ БИБЛИОТЕКЕ
СТАНДАРТОВ**

**ESTIMATION OF THE SIZE OF INFORMATION IN THE ELECTRONIC
LIBRARY OF STANDARDS**

Уральский федеральный университет имени первого Президента России
Б. Н. Ельцина, г. Екатеринбург

Ural Federal University named after the first President of Russia B. N. Yeltsin, Ekaterinburg

В статье рассматривается алфавитный подход измерения размера информации в стандартах. Для оценки размера информации в действующих стандартах были использованы стандарты групп 17.040 «Линейные и угловые измерения» и 17.140 «Акустика и акустические измерения» в соответствии с общероссийским классификатором стандартов ОК 001-2021 (ИСО МКС).

The article discusses the alphabetical approach to measuring the size of information in standards. To estimate the amount of information in the current standards, the standards of groups 17.040 «Linear and angular measurements» and 17.140 «Acoustics and acoustic measurements» were used in accordance with the all-Russian classifier of standards ОК 001-2021 (ISO MKS).

Ключевые слова: стандарт, количество информации, алфавитный подход

Keywords: standard, amount of information, alphabetic approach

В современных условиях цифровой трансформации стандарты как источник достоверной, объективной информации приобретают особое значение и для успешного присоединения этих сведений к различным информационным системам необходимо получить представление о размере информации, содержащейся в стандартах [1, 2].

Целью исследования является оценка размера информации для стандартов из числа подгрупп 17 раздела ОК 001-2021 (ИСО МКС) – «Метрология и измерения. Физические явления» на основе использования расчетов алфавитного подхода. Данный подход позволяет определить количество знаков, содержащихся в тексте, и поэтому является объективным, так как не зависит от субъекта, воспринимающего текст. Алфавитный подход не связывает

¹ Пиджакова Е. – студент магистратуры
Казанцева Н. – канд. техн. наук, доцент

количество информации с содержанием сообщения, он позволит рассчитать размер информации, содержащейся в стандартах, и с помощью полученных данных станет возможным оценить полученный размер данных в стандартах групп 17.040 «Линейные и угловые измерения» и 17.140 «Акустика и акустические измерения».

В соответствии с общероссийским классификатором стандартов ОК 001-2021 (ИСО МКС) группа стандартов 17.040 «Линейные и угловые измерения» состоит из следующих подгрупп:

- 17.040.01 «Линейные и угловые измерения в целом»;
- 17.040.10 «Допуски и посадки»;
- 17.040.20 «Свойства поверхностей»;
- 17.040.30 «Измерительные приборы»;
- 17.040.40 «Геометрические характеристики продукции»;
- 17.040.99 «Другие стандарты, связанные с линейными и угловыми измерениями».

Общее количество действующих стандартов на текущий момент составляет 294 стандарта. Состав базы стандартов группы 17.040 по подгруппам и временным интервалам принятия по текущему состоянию на 2022 год представлен в таблице 1.

Таблица 1

Состав стандартов группы ОКС 17.040 Линейные и угловые измерения

Временной интервал принятия стандартов, гг.	Количество стандартов по подгруппам, шт.						Всего
	17.040.01	17.040.10	17.040.20	17.040.30	17.040.40	17.040.99	
1950-1959	-	1	-	3	-	-	4
1960-1969	-	4	-	36	-	-	40
1970-1979	-	1	-	75	-	-	76
1980-1989	-	6	1	94	-	-	101
1990-1999	-	1	1	23	-	-	25
2000-2009	13	2	-	4	-	1	20
2010-2019	4	4	5	14	-	-	27
2020-2022	-	-	-	1	-	-	1
Общее количество действующих стандартов, шт.	17	19	7	250	0	1	294

Для определения объема информации, содержащейся в стандартах группы 17.040 необходимо оценить ее объем в репрезентативной выборке стандартов.

В расчетах использована стратифицированная выборка – выборка, при которой, перед тем как начать отбор, генеральная совокупность разбивается на однородные части (страты) по определенному признаку, а затем из них производится систематический или простой случайный отбор. В качестве признака расслоения приняты признак деления по шести подгруппам в соответствии с ОКС и временным интервалам принятия действующих стандартов.

Генеральной совокупностью в исследовании является совокупность стандартов группы ОКС 17.040 «Линейные и угловые измерения». Их число составляет 294 стандарта (действующих). Объем выборки составил 60 стандартов или 20 % от всей базы данной группы стандартов.

Разделив объем выборки на объем генеральной совокупности, получим значение 0,2. Теперь можно число 0,2 умножать на каждую из внутриклеточных частот в таблице, представляющей генеральную совокупность. Так, 13 мы умножим на 0,2 и получим 2,6. Значит, в подгруппе 17.040.01 2000-2009 годов мы должны отобрать 3 стандарта. Таким же образом необходимо произвести вычисления для каждой клетки таблицы.

Следовательно, следует при помощи таблицы случайных чисел отобрать следующее число стандартов из подгрупп:

- из подгруппы 17.040.01 – 3 стандарта;
- из подгруппы 17.040.10 – 4 стандарта;
- из подгруппы 17.040.20 – 1 стандарт;
- из подгруппы 17.040.30 – 51 стандарт;
- из подгруппы 17.040.40 – 0 стандартов;
- из подгруппы 17.040.99 – 1 стандарт.

Процент, который составляют в объеме выборочной совокупности стандарты каждой страты, полностью совпадает с процентом данной страты в количестве всех стандартов, то есть в генеральной совокупности. Так число стандартов в подгруппе 17.040.01 2010-2019 годов в выборке – 1, общий объем выборки – 60 стандартов, следовательно, процент стандартов данной подгруппы в выборочной совокупности составляет 2 %. Этот процент идентичен соответствующему показателю по генеральной совокупности. То же самое можно сказать о каждой из оставшихся страт. По таблице случайных чисел выбраны стандарты и составлена структура выборочной совокупности, представленная в таблице 2.

В соответствии с общероссийским классификатором стандартов ОК 001-2021 (ИСО МКС) группа стандартов 17.140 «Акустика и акустические измерения» состоит из следующих подгрупп:

- 17.140.01 «Акустические измерения и борьба с шумами в целом»;
- 17.140.20 «Шум от машин и оборудования»;
- 17.140.30 «Шум от транспорта»;
- 17.140.50 «Электроакустика»;
- 17.140.99 «Другие стандарты, связанные с акустикой».

В результате указанного отбора каждый стандарт, входящий в конкретную страту в генеральной совокупности, имеет равные шансы с другими стандартами, входящими в данную страту, попасть в выборочную совокупность [3, 4].

Таблица 2

Состав выборочной совокупности стандартов

Временной интервал принятия стандартов, гг.	Количество стандартов по подгруппам, шт.						Всего
	17.04 0.01	17.040.10	17.040.20	17.040.30	17.040.40	17.040.99	
1950-1959	-	0	-	1	-	-	1
1960-1969	-	1	-	7	-	-	8
1970-1979	-	0	-	15	-	-	15
1980-1989	-	2	0	19	-	-	21
1990-1999	-	0	0	5	-	-	5
2000-2009	2	0	-	1	-	1	4
2010-2019	1	1	1	3	-	-	6
2020-2022	-	-	-	0	-	-	0
Общее количество стандартов, шт.	3	4	1	51	0	1	60

В сформированной выборке, по проведенным расчетам согласно методике, представленной в статье [5], средний размер одной страницы стандарта составляет примерно 3 Кбайт. Общее количество действующих стандартов по текущему состоянию 2022 года составляет 183 стандарта. Состав базы стандартов группы 17.140 по подгруппам и временным интервалам принятия представлен в таблице 3.

Объем выборки составил 55 стандартов от всей базы данной группы стандартов.

При помощи таблицы случайных чисел отобрать следующее число стандартов из подгрупп:

- из подгруппы 17.140.01 – 8 стандартов;
- из подгруппы 17.140.20 – 27 стандартов;
- из подгруппы 17.140.30 – 10 стандартов;
- из подгруппы 17.140.50 – 10 стандартов;
- из подгруппы 17.140.99 – 0 стандартов.

Таблица 3

Состав стандартов группы 17.140 Акустика и акустические измерения в соответствии с ОКС (ОК001-2021 (ИСО МКС))

Временной интервал принятия стандартов, гг.	Количество стандартов по подгруппам, шт.					Всего
	17.140.01	17.140.20	17.140.30	17.140.50	17.140.99	
1970-1979	-	-	1	2	-	3
1980-1989	-	4	7	2	-	13
1990-1999	-	9	1	1	-	11
2000-2009	10	36	6	3	-	55
2010-2019	17	36	18	23	1	95
2020-2022	1	4	-	1	-	6
Общее количество действующих стандартов, шт.	28	89	33	32	1	183

Структура выборочной совокупности представлена в таблице 4.

Таблица 4

Состав выборочной совокупности стандартов

Временной интервал принятия стандартов, гг.	Количество стандартов по подгруппам, шт.					Всего
	17.140.01	17.140.20	17.140.30	17.140.50	17.140.99	
1970-1979	-	-	0	1	-	1
1980-1989	-	1	3	1	-	5
1990-1999	-	3	0	0	-	3
2000-2009	3	11	2	1	-	17
2010-2019	5	11	5	7	0	28
2020-2022	0	1	-	0	-	1
Общее количество стандартов, шт.	8	27	10	10	0	55

По аналогичным расчетам выборки средний размер одной страницы стандарта составляет примерно 4 Кбайт.

По результатам исследования определено, что размер информации, содержащейся в стандарте, приходящийся на одну страницу текста документа в выбранных подгруппах стандартов составил от 3 до 4 Кбайт. Можно оценить, что размер информации в стандартах подгруппы 17.040 составляет 18464 Кбайт (всего в 294 стандартах 4616 страниц), а в подгруппе 17.140 – 18904 Кбайт (всего в 183 стандартах 4726 страниц). Полученные данные необходимы при создании и проектировании электронных библиотек стандартов на предприятиях. Такие библиотеки создают возможность получения актуальной информации, содержащаяся в стандартах, на любых рабочих местах по ходу технологических процессов, осуществляемых на предприятиях.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Количественная оценка информации как основа построения базы данных (на примере стандартов на металлопродукцию) / М. А. Полякова, Т. В. Казанцева, Н. К. Казанцева, Е. Н. Пиджакова // Черная металлургия. Бюллетень научно-технической и экономической информации. – 2022. – Т. 78. – № 4. – С. 338-344. – DOI 10.32339/0135-5910-2022-4-338-344.
2. Пиджакова, Е.Н. Анализ нормативной базы стандартов РФ на цветные металлы / Е.Н. Пиджакова, Т.В. Казанцева, Н.К. Казанцева : сборник статей IX Всероссийской научно-практической конференции с международным участием, Екатеринбург, 19 мая 2022 г. – Екатеринбург : Издательский дом «Ажур», 2022. – С. 61-67.
3. Казанцева, Т. В. Разработка подхода для цифровизации стандартов на метизную продукцию / Т. В. Казанцева, Н. К. Казанцева, М. А. Полякова // Международная научно-практическая конференции им. Д.И. Менделеева, посвященная 90-летию профессора Р.З. Магарила : Материалы конференции, Тюмень, 25–27 ноября 2021 года / Отв. редактор А.Н. Халин. – Тюмень: Тюменский индустриальный университет, 2022. – С. 219-220.
4. Казанцева, Т. В. Предпосылки создания электронной библиотеки стандартов / Т. В. Казанцева, Н. К. Казанцева, М. А. Полякова // Актуальные проблемы современной науки, техники и образования : Тезисы 80-й международной научно-технической конференции, Магнитогорск, 18–22 апреля 2022 года. – Магнитогорск: Магнитогорский государственный технический университет им. Г.И. Носова, 2022. – С. 199.
5. Пиджакова, Е. Оценка размера информации в стандартах / Е. Пиджакова, Т. В. Казанцева // Роль технического регулирования и стандартизации в эпоху цифровой экономики : сборник статей участников III Международной научно-практической конференции молодых ученых, Екатеринбург, 25 ноября 2021 года. – Екатеринбург: Издательский дом «Ажур», 2021. – С. 83-88.