МЕЖГОСУДАРСТВЕННЫЙ СТАНДАРТ

СИСТЕМА СТАНДАРТОВ ПО ИНФОРМАЦИИ, БИБЛИОТЕЧНОМУ И ИЗДАТЕЛЬСКОМУ ДЕЛУ

ТЕЗАУРУС ИНФОРМАЦИОННО-ПОИСКОВЫЙ МНОГОЯЗЫЧНЫЙ

СОСТАВ, СТРУКТУРА И ОСНОВНЫЕ ТРЕБОВАНИЯ К ПОСТРОЕНИЮ

Издание официальное

БЗ 9—2000

ИПК ИЗДАТЕЛЬСТВО СТАНДАРТОВ
Москва

gost.e-ecolog.ru

УДК 025.48:006.354 Группа Т62

М Е Ж Г О С У Д А Р С Т В Е Н Н Ы Й С Т А Н Д А Р Т

Система стандартов по информации, библиотечному и издательскому делу

ТЕЗАУРУС ИНФОРМАЦИОННО-ПОИСКОВЫЙ МНОГОЯЗЫЧНЫЙ

Состав, структура и основные требования к построению

ГОСТ 7.24—90

System of standards on information, librarianship and publishing. Multilingual thesaurus for information retrieval. Composition, structure and basic requirements for development

ОКСТУ 0007

Дата введения 01.01.91

Настоящий стандарт распространяется на многоязычные информационно-поисковые тезаурусы (МИПТ) и устанавливает состав, структуру и основные требования к построению МИПТ, применяемых в информационно-поисковых системах.

1. ОБЩИЕ ПОЛОЖЕНИЯ

1.1. Многоязычный информационно-поисковый тезаурус (МИПТ) — согласованная совокупность одноязычных информационно-поисковых тезаурусов, содержащая эквивалентные дескрипторы на языках — компонентах МИПТ, необходимые и достаточные для межъязыкового обмена, и включающая средства для указания их эквивалентности.

МИПТ предназначен для обработки и поиска информации, а также обмена данными между информационными системами, пользующимися различными языками.

1.2. Одноязычная версия МИПТ — каждый из одноязычных информационно-поисковых тезаурусов (ИПТ), входящих в состав МИПТ.

1.3. Язык-компонент МИПТ — язык, на основе которого разработана какая-либо из одноязычных версий.

1.4. Дескриптор МИПТ — основная лексическая единица МИПТ, представляющая собой совокупность эквивалентных дескрипторов одноязычных версий, связанных средствами указания эквивалентности.

2. СОСТАВ И СТРУКТУРА ТЕЗАУРУСА

2.1. Обязательными составными частями МИПТ являются:

1) вводная часть;

2) алфавитные лексико-семантические указатели, включающие средства указания эквивалентности дескрипторов одноязычных версий (п. 2.4). Допускается в состав тезауруса вводить систематические, пермутационные, иерархические и другие указатели и списки специальных категорий лексических единиц, а также приложения, содержащие дополнительные сведения о разработке и использовании МИПТ.

2.2. Вводная часть включает титульный лист и введение, составленные согласно ГОСТ 7.25 и изложенные на одном или нескольких языках-компонентах.

На титульном листе должны быть указаны все языки-компоненты. Рекомендуется наименование тезауруса на титульном листе формулировать на одном языке, а дополнительно к основному титульному листу включать во вводную часть титульные листы на всех языках-компонентах. Введение должно содержать краткое описание методических принципов установления эквивалент-

Издание официальное ★ **Перепечатка воспрещена**

© Издательство стандартов, 1990
© ИПК Издательство стандартов, 2002

ности дескрипторов одноязычных версий, а также указание на то, что тезаурус разработан в соответствии с настоящим стандартом. Рекомендуется в состав вводной части включать полный или сокращенный перевод введения на всех языках-компонентах.

2.3. Алфавитный лексико-семантический указатель в каждой из одноязычных версий содержит перечень тезаурусных статей, расположенных в алфавитном порядке заглавных лексических единиц на соответствующем языке-компоненте. Состав и структура тезаурусных статей в лексико-семантических указателях одноязычных версий — по ГОСТ 7.25.

2.4. Средства указания эквивалентности дескрипторов одноязычных версий реализуются введением в тезаурусную статью кода эквивалентности дескрипторов (идентификационного кода) и (или) эквивалентных дескрипторов на языках-компонентах.

При использовании идентификационных кодов в дополнение к алфавитному лексико-семантическому указателю МИПТ составляют указатель идентификационных кодов, в котором каждому коду приписывают соответствующие дескрипторы одноязычных версий.

2.5. Систематические, иерархические и пермутационные указатели строятся по ГОСТ 7.25 на языках-компонентах и содержат средства указания эквивалентности дескрипторов одноязычных версий. Примеры использования средств указания эквивалентности дескрипторов приведены в приложении.

2.6. Состав и структура одноязычных версий должны удовлетворять следующим требованиям:

1) в каждой одноязычной версии должны содержаться все дескрипторы, необходимые с точки зрения разработчика для обработки и поиска, а также обмена информацией между системами, использующими различные языки;

2) классы условной эквивалентности лексических единиц одноязычных версий должны соответствовать друг другу по объему выражаемых понятий;

3) соответствующие друг другу дескрипторы отдельных одноязычных версий должны иметь по возможности максимально совпадающие парадигматические отношения.

Одноязычные версии могут быть дополнены дескрипторами, не являющимися необходимыми для межъязыкового обмена и отсутствующими в других одноязычных версиях; при этом указывают иерархическую подчиненность их какому-либо дескриптору МИПТ.

Одноязычные версии могут включать парадигматические отношения, отсутствующие в других одноязычных версиях, если они на данном языке-компоненте уточняют объем понятия, выраженного дескриптором МИПТ, и не противоречат его пониманию в других языках-компонентах.

3. ОСНОВНЫЕ ТРЕБОВАНИЯ К ПОСТРОЕНИЮ МИПТ

3.1. Содержание работ по формированию одноязычных версий определяется наличием или отсутствием одноязычных тезаурусов по данной тематике на языках-компонентах.

3.1.1. При отсутствии одноязычных тезаурусов по данной тематике на языках-компонентах, как правило, сначала создают одну одноязычную версию и на ее основе строят остальные.

Возможна также совместная разработка одновременно всех или нескольких одноязычных версий при взаимном согласовании состава дескрипторов и их парадигматических связей.

3.1.2. При наличии тезауруса по данной тематике на одном из языков-компонентов, как правило, он берется за основу разрабатываемого МИПТ. Другие одноязычные версии формируются подбором эквивалентов дескрипторов этого тезауруса на языках-компонентах. При этом возможны различные степени семантического соответствия эквивалентов дескрипторов на различных языках.

3.1.3. При наличии нескольких исходных тезаурусов по данной тематике на языках-компонентах разработка МИПТ ведется подбором эквивалентных дескрипторов в данных тезаурусах (пп. 3.2.1 и 3.2.2). Дескрипторы, не имеющие эквивалентов в отдельных одноязычных версиях, переводят на соответствующие языки и включают в одноязычные версии, в которых этот дескриптор первоначально отсутствовал (п. 3.2.3). При этом следует вносить в исходные тезаурусы изменения, необходимые для согласования одноязычных версий.

3.1.4. Рекомендуется одну из одноязычных версий выбрать в качестве базовой и устанавливать эквивалентность дескрипторов других одноязычных версий по отношению к дескрипторам базовой версии.

3.2. При установлении эквивалентности дескрипторов различных одноязычных версий необходимо различать на разных языках-компонентах следующие степени эквивалентности терминов:

1) полная;
2) неполная;
3) частичная;
4) отсутствие эквивалентного термина.

3.2.1. При наличии в языках-компонентах полностью эквивалентных терминов они считаются представителями одного дескриптора МИПТ.

3.2.2. При отсутствии в языках-компонентах полных эквивалентов для выражения одного и того же понятия в качестве дескриптора МИПТ в одноязычных версиях используют неполные и частичные эквиваленты.

Неполными эквивалентами являются термины, для которых объемы выражаемых ими понятий пересекаются.

Частичными эквивалентами являются термины, для которых объем понятия, выражаемого одним эквивалентом, входит в объем понятия, выражаемого другим эквивалентом.

3.2.2.1. Если различие в объеме понятий между неполными или частичными эквивалентами не существенно для данного МИПТ, они могут быть использованы для представления дескриптора МИПТ в разных одноязычных версиях, например:

русск.	франц.
АННОТАЦИЯ	RESUME

3.2.2.2. Если различие в объеме понятий между неполными или частичными эквивалентами существенно, при представлении дескриптора МИПТ в одноязычных версиях значение используемого термина уточняется одним из следующих способов:

1) добавлением релятора (п. 3.2.2.3);

2) добавлением примечания (п. 3.2.2.4);

3) включением термина в уточняющее словосочетание с определяющими словами (п. 3.2.2.5);

4) введением условно эквивалентных аскрипторов в тезаурусную статью дескриптора одноязычной версии (п. 3.2.6);

5) использованием комбинации двух или более терминов (п. 3.2.2.7).

3.2.2.3. Добавлением релятора, который является частью дескриптора, применяется в тех случаях, когда для более полного соответствия содержанию дескриптора необходимо ограничить объем понятия, выражаемого термином, например:

англ.	русск.
TREE	ДЕРЕВО (РАСТЕНИЕ)
WOOD (MATERIAL)	ДЕРЕВО (МАТЕРИАЛ)
WOOD (LANDSCAPE)	ЛЕС (ЛАНДШАФТ)

3.2.2.4. Добавление примечания, которое не является частью дескриптора, используют в случаях уточнения объема понятия, выраженного соответствующими дескрипторами на разных языках-компонентах, например:

франц.	русск.
AFFILIATION	МЕСТО РАБОТЫ (как элемент библиографического описания, обозначающий место работы автора публикации)

3.2.2.5. Включение термина в уточняющее словосочетание применяется в тех случаях, когда требуется ограничить объем понятия и данное словосочетание устойчиво употребляется в документах, например:

англ.	русск.
CRANE (MACHINE)	ПОДЪЕМНЫЙ КРАН

3.2.2.6. Введение условно эквивалентных аскрипторов в тезаурусную статью дескриптора одноязычной версии применяется в тех случаях, когда они необходимы для однозначного представления понятия в данной версии, например:

русск.	англ.
ГРАФИК	GRAPH UF diagram chart

3.2.2.7. Если исходный дескриптор обозначает понятие, не выраженное в виде единого термина на другом языке-компоненте, допускается использовать в одноязычной версии в качестве эквивалента комбинации двух или более дескрипторов одноязычной версии, например:

русск.	франц.
ПЕРФОКАРТА РУЧНОГО ОБРАЩЕНИЯ РЕФЕРИРОВАНИЕ	CARTE PERFOREE+ SELECTION MANUELLE RESUME ANALYTIQUE + ELABORATION

3.2.3. При отсутствии лексического эквивалента в одном из языков-компонентов допускаются решения, приведенные в пп. 3.2.3.1—3.2.3.5.

3.2.3.1. Создание нового термина переводом или калькированием иноязычного дескриптора, например:

русск.	англ.	франц.
ПЯТИЛЕТНИЙ ПЛАН	FIVE-YEAR PLAN	PLAN QUINQUENNAL

3.2.3.2. Использование в качестве эквивалента словосочетания, например:

англ.	русск.
CORER	МАШИНА ДЛЯ УДАЛЕНИЯ СЕРДЦЕВИНЫ ПЛОДА

3.2.3.3. Заимствование иноязычного термина, например:

англ.	русск.
FILE	ФАЙЛ

3.2.3.4. Использование в качестве эквивалента термина с близким значением, например:

русск.	франц.
ЛИНГВИСТИЧЕСКОЕ ОБЕСПЕЧЕНИЕ	OUTILS LINGUISTIQUES
русск.	англ.
ПОДРОСТОК	TEENAGER

3.2.3.5. Использование в качестве эквивалента термина более широкого по значению и не входящего в лексический состав данной тематической области, например, в МИПТ по библиотечному делу:

франц.	русск.
ANIMATEUR	КУЛЬТМАССОВЫЙ РАБОТНИК

3.3. Форма представления одноязычных версий — по ГОСТ 7.25 и п. 2.4.

3.4. Процесс ведения МИПТ сводится к ведению его одноязычных версий и к обеспечению их согласованности. Изменения вносят одновременно во все одноязычные версии после их согласования на уровне языков-компонентов.

ПРИЛОЖЕНИЕ
Справочное

ПРИМЕРЫ ИСПОЛЬЗОВАНИЯ СРЕДСТВ УКАЗАНИЯ ЭКВИВАЛЕНТНОСТИ ДЕСКРИПТОРОВ В ОДНОЯЗЫЧНЫХ ВЕРСИЯХ

1. Использование цифровых идентификационных кодов, не зависящих от языка

STALA
011420
a CHARAKTERYSTYKA
INDEKS
MODUL
STALY
WSKAZNIK (LICZBA)

2. Использование эквивалентов дескрипторов на другом языке

КОНСТАНТА
STALA
c Постоянная
a КОЭФФИЦИЕНТ
МОДУЛЬ
ПОКАЗАТЕЛЬ
ПОСТОЯННЫЙ
ХАРАКТЕРИСТИКА

3. Использование эквивалентов дескрипторов на других языках и идентификационных кодов

БЕНЗИНЫ
fre ESSENCE MOTEUR
eng GASOLINE
pol BENZYNA
benzynawysokooktanowa
011420
в НЕФТЕПРОДУКТЫ СВЕТЛЫЕ
ТОПЛИВО ЖИДКОЕ
ТОПЛИВО КАРБЮРАТОРНОЕ
н БЕНЗИНЫ АВИАЦИОННЫЕ
БЕНЗИНЫ АВТОМОБИЛЬНЫЕ
БЕНЗИНЫ АЛКИЛИРОВАННЫЕ

4. Использование совместного расположения статей на двух языках

АМАЛЬГАМА СЕРЕБРА	SILBERAMALGAM
в МЕТАЛЛЫ	OB METALLE
н КУПРОАРКВЕРИТ	UB CUPROARQUERIT

ИНФОРМАЦИОННЫЕ ДАННЫЕ

1. **РАЗРАБОТАН И ВНЕСЕН Государственным комитетом СССР по науке и технике Академией наук СССР Государственным комитетом СССР по управлению качеством продукции и стандартам**

2. **УТВЕРЖДЕН И ВВЕДЕН В ДЕЙСТВИЕ Постановлением Государственного комитета СССР по управлению качеством продукции и стандартам от 26.01.90 № 96**

3. **Стандарт полностью соответствует СТ СЭВ 6574—89**

4. **ВЗАМЕН ГОСТ 7.24—80**

5. **ССЫЛОЧНЫЕ НОРМАТИВНО-ТЕХНИЧЕСКИЕ ДОКУМЕНТЫ**

Обозначение НТД, на который дана ссылка	Номер пункта
ГОСТ 7.25—2001	2.2, 2.3, 2.5, 3.3

6. **ПЕРЕИЗДАНИЕ. Октябрь 2001 г.**

Редактор *Р.Г. Говердовская*
Технический редактор *Н.С. Гришанова*
Корректор *В.И. Варенцова*
Компьютерная верстка *Л.А. Круговой*

Изд. лиц. № 02354 от 14.07.2000. Сдано в набор 21.11.2001. Подписано в печать 18.12.2001. Усл. печ. л. 0,93. Уч.-изд. л. 0,60. Тираж 140 экз. С 3203. Зак. 1144.

ИПК Издательство стандартов, 107076, Москва, Колодезный пер., 14.
http://gost.e-ecolog.ru mail: info@standards.ru
Набрано в Издательстве на ПЭВМ
Филиал ИПК Издательство стандартов — тип. "Московский печатник", 103062, Москва, Лялин пер., 6.
Плр № 080102

gost.e-ecolog.ru