



**ГОСУДАРСТВЕННЫЙ СТАНДАРТ
СОЮЗА ССР**

**СИСТЕМА СТАНДАРТОВ ПО ИНФОРМАЦИИ,
БИБЛИОТЕЧНОМУ И ИЗДАТЕЛЬСКОМУ ДЕЛУ**

**ТЕЗАУРУС
ИНФОРМАЦИОННО-ПОИСКОВЫЙ
ОДНОЯЗЫЧНЫЙ**

**ПРАВИЛА РАЗРАБОТКИ, СТРУКТУРА, СОСТАВ
И ФОРМА ПРЕДСТАВЛЕНИЯ**

**ГОСТ 7.25—80
{СТ СЭВ 174—85}**

Издание официальное

**ГОСУДАРСТВЕННЫЙ КОМИТЕТ СССР ПО СТАНДАРТАМ
Москва**

Система стандартов по информации,
библиотечному и издательскому делу

**ТЕЗАУРУС ИНФОРМАЦИОННО-ПОИСКОВЫЙ
ОДНОЯЗЫЧНЫЙ**

Правила разработки, структура, состав и форма
представления

System of standards on information, librarianship
and publishing. Monolingual thesaurus for information
retrieval. Rules for development, structure, composition
and form of presentation

ГОСТ

7.25—80

[СТ СЭВ 174—85]

Постановлением Государственного комитета СССР по стандартам от 31 декабря
1980 г. № 6340 срок введения установлен

с 01.01.82

Настоящий стандарт устанавливает правила разработки, структуру, состав и форму представления информационно-поисковых тезаурусов (ИПТ), ориентированных на использование лексики русского языка и разрабатываемых в рамках автоматизированных систем научно-технической информации.

Стандарт содержит все требования СТ СЭВ 174—85.

В стандарт дополнительно включены требования к форме лексических единиц на русском языке и к взаимодействию с ГАСНТИ (см. приложение 2).

1. ОБЩИЕ ПОЛОЖЕНИЯ

1.1. **Одноязычный информационно-поисковый тезаурус** — контролируемый словарь лексических единиц дескрипторного языка, основанный на лексике одного естественного языка, отображающий семантические отношения между лексическими единицами и предназначенный для организации поиска информации путем индексирования документов и/или запросов.

Целью создания ИПТ является повышение показателей качества поиска информации в автоматизированных информационно-поисковых системах.

1.2. **Лексическая единица (ЛЕ) ИПТ** — выбранное для использования в ИПТ слово, словосочетание или лексически значимая компонента сложного слова естественного языка.

(Измененная редакция, Изм. № 1).

Издание официальное

Печелатка воспрещена

★

* Переиздание (февраль 1988 г.) с Изменением № 1,
утвержденным в декабре 1986 г. (ИУС 3—87).

© Издательство стандартов, 1988

1.3. **Дескриптор** — лексическая единица ИПТ, предназначенная для использования в поисковых образах документов и/или запросов.

1.4. **Аскриптор (недескриптор)** — лексическая единица ИПТ, которая в поисковых образах документов (запросов) подлежит замене на дескриптор при поиске или обработке информации.

1.5. ИПТ подразделяют на два типа:

1 — тезаурусы, выделяющие среди своих лексических единиц дескрипторы и аскрипторы;

2 — тезаурусы, все лексические единицы которых являются дескрипторами.

2. ПРАВИЛА ПОСТРОЕНИЯ ИПТ

2.1. Построение ИПТ состоит из следующих этапов:

определение тематического охвата ИПТ;

сбор массива лексических единиц (ЛЕ);

формирование словника ИПТ;

построение словарных статей и указателей;

оформление ИПТ;

экспертиза и регистрация ИПТ.

2.2. **Определение тематического охвата ИПТ**

Определение тематического охвата ИПТ осуществляется на основе Рубрикатора ГАСНТИ путем анализа информационной потребности абонентов (потребителей). При анализе состава имеющегося или проектируемого информационного фонда и информационных потребностей абонентов отмечаются все рубрики нижнего уровня Рубрикатора, по которым производится комплектование фондов или соответствующие запросам абонентов.

2.3. **Сбор массива лексических единиц**

До начала работ по сбору массива лексических единиц разработчик должен обратиться в центральную службу Автоматизированной системы ведения информационных языков (АСВИЯ) с целью определения наличия зарегистрированных тезаурусов по заданной тематике.

При наличии таких тезаурусов производится оценка возможности внедрения их в данной системе.

Если в АСВИЯ отсутствуют тезаурусы по данной тематике или экспертом АСВИЯ будет признана невозможность внедрения имеющихся тезаурусов в данной системе, разработчик производит сбор лексики и дальнейшие процедуры по созданию тезауруса.

Первоначальный сбор лексики осуществляется выделением лексических единиц (ЛЕ) из представительной коллекции документов и/или запросов.

В полученный массив ЛЕ дополнительно должны быть включены соответствующие тематике лексические единицы, выделенные

в соответствии с методикой построения данного тезауруса из следующих источников:

Рубрикатора ГАСНТИ;

рубрикатора соответствующей АСНТИ;

энциклопедических, толковых и терминологических словарей и справочников;

терминологических стандартов;

классификаторов технико-экономической информации;

таблиц УДК, МКИ и других систем классификации.

Наименования стран и других географических единиц должны соответствовать требованиям СТ СЭВ 251—76, наименования языков — требованиям СТ СЭВ 252—76.

2.4. Формирование словника ИПТ

2.4.1. В словник ИПТ могут быть включены следующие типы лексических единиц:

одиночные слова (существительные, прилагательные, глаголы, наречия);

именные словосочетания;

лексически значимые компоненты сложных слов;

аббревиатуры;

сокращения слов и словосочетаний.

2.4.2. Словосочетания могут включаться в словник, если в качестве опорного слова содержат существительное, и выполнено одно из следующих условий:

значение словосочетания не выводится из значений его компонентов, например:

ЧЕРНЫЙ ЯЩИК

АБСОЛЮТНО ЧЕРНОЕ ТЕЛО

ЦАРСКАЯ ВОДКА;

хотя бы один из компонентов словосочетания не употребляется в составе других сочетаний или употребляется всегда в другом смысле, например:

ТОРГОВЛЯ НА ВЫНОС

ЛЕГКАЯ ПРОМЫШЛЕННОСТЬ;

для данного словосочетания в словнике ИПТ существуют полные синонимы, например:

НАТРИЯ ХЛОРИД = ПОВАРЕННАЯ СОЛЬ;

данное словосочетание является устойчивым словосочетанием с именем собственным, например:

ТАБЛИЦА МЕНДЕЛЕЕВА

ЗАКОН БОЙЛЯ-МАРИОТТА;

отдельные слова словосочетания имеют слишком широкое значение, например, слово **МАШИНЫ** в словосочетаниях:

СТРОИТЕЛЬНЫЕ МАШИНЫ

ЭЛЕКТРИЧЕСКИЕ МАШИНЫ;

для данного словосочетания в словнике ИПТ существует общепринятая аббревиатура, как например:

ПОВЕРХНОСТНО АКТИВНЫЕ ВЕЩЕСТВА = ПАВ
УНИВЕРСАЛЬНАЯ ДЕСЯТИЧНАЯ КЛАССИФИКАЦИЯ =
= УДК

ИНФОРМАЦИОННО-ПОИСКОВЫЙ ТЕЗАУРУС = ИПТ
ЭЛЕКТРОННО-ВЫЧИСЛИТЕЛЬНАЯ МАШИНА = ЭВМ;

разбиение словосочетаний на отдельные компоненты приводит к потере важных для поиска семантических связей, например, разбиение ЛЕ «ЯЗЫК ПРОГРАММИРОВАНИЯ» не позволило бы установить связи с такими ЛЕ как «АЛГОЛ», «КОБОЛ», «ФОРТРАН» и др.

Словосочетания, не удовлетворяющие указанным требованиям, следует разбивать на компоненты.

(Измененная редакция, Изм. № 1).

2.4.3. Падежные формы одиночных существительных следует заменять формой именительного падежа. Падежные формы опорных существительных в словосочетаниях также заменяются формой именительного падежа с одновременным изменением надежной формы согласованных определений.

2.4.4. Формы прилагательных и причастий следует приводить к именительному падежу. Прилагательные и причастия в единственном числе приводятся к форме мужского рода.

2.4.5. Глаголы рекомендуется заменять отглагольными существительными. Глаголы, включенные в словник, приводят к форме инфинитива.

2.4.6. Компоненты сложных слов, включаемые в словник ИПТ, должны быть широко употребляющимися частями сложных слов, общепринятых в языке данной специальности. Отсутствующая компонента сложного слова заменяется дефисом, например:

ЭЛЕКТРО-
ГИДРО-
ВОДО-
ТЕРМО-

(Измененная редакция, Изм. № 1).

2.4.7. При наличии аббревиатур в словник должны включаться также соответствующие им полные формы.

Пример: ИПТ — ИНФОРМАЦИОННО-ПОИСКОВЫЙ ТЕЗАУРУС,
ОКГ — ОПТИЧЕСКИЙ КВАНТОВЫЙ ГЕНЕРАТОР,
УКВ — УЛЬТРАКОРОТКОВОЛНОВЫЙ.

2.4.8. Сокращения должны соответствовать требованиям ГОСТ 7.12—77.

2.4.9. Исключения из правил п. 2.4 возможны только в тех случаях, когда требуемая каноническая форма отсутствует.

Тип ссылки	Обозначение на русском языке	Допустимое символическое обозначение	Русские ссылки	Аналог на английском языке
1. Ссылка от аскриптора к эквивалентному дескриптору	см	=	смотри	USE
2. Ссылка от дескриптора к эквивалентному аскриптору	с	≠	синоним	UF (used for)
3. Ссылка от аскриптора к нескольким альтернативно заменяющим его дескрипторам	иа	= :	используй альтернативно	—
4. Ссылка от аскриптора к заменяющей его комбинации дескрипторов	ик	= +	используй комбинацию	—
5. Ссылка от дескриптора к вышестоящему дескриптору	в	<	выше	BT (broader term)
6. Ссылка от дескриптора к вышестоящему родовому дескриптору	вр	<G	выше-род	BTG (broader term generic)
7. Ссылка от дескриптора к вышестоящему дескриптору, обозначающему целое	вц	<P	выше-целое	BTP (broader term partitive)
8. Ссылка от дескриптора к нижестоящему дескриптору	н	>	ниже	NT (narrower term)
9. Ссылка от дескриптора к нижестоящему видовому дескриптору	нв	>G	ниже-вид	NTG (narrower term generic)
10. Ссылка от дескриптора к нижестоящему дескриптору, обозначающему часть	нч	>P	ниже-часть	NTP (narrower term partitive)
11. Ссылка от дескриптора к ассоциативно связанному дескриптору	а	—	ассоциация	RT (related term)
12. Ссылка от дескриптора к аскриптору, который заменяется данным дескриптором при альтернативном выборе (обратная ссылка к «иа»)	са	: =	сравни альтернативный выбор	—
13. Ссылка от дескриптора к аскриптору, который заменяется комбинацией, включающей данный дескриптор (обратная ссылка к «ик»)	ск	+ =	сравни комбинацию	—
14. Техническая обратная ссылка	ср	:	сравни	—

2.5. Построение словарных статей

2.5.1. При построении словарных статей ИПТ лексическим единицам приписываются ссылки согласно таблице.

2.5.2. Ссылки отмечают связи данной ЛЕ с другими и являются результатом выполнения следующих операций:

устранения неоднозначности ЛЕ;

установления отношений эквивалентности;

выбора дескриптора, представляющего класс эквивалентности при индексировании (для ИПТ 1-го типа, различающих дескрипторы и аскрипторы);

установления иерархических и ассоциативных отношений дескрипторов.

2.5.3. Устранение неоднозначности ЛЕ

Неоднозначность ЛЕ устраняется релятором или лексическим примечанием в зависимости от характера неоднозначности и способа использования ИПТ.

2.5.2, 2.5.3. (Измененная редакция, Изм. № 1).

2.5.3.1. Релятор является частью ЛЕ и поясняет ее значение, относя ее к определенной понятийной категории или предметно-тематической области. Допускается релятор представлять сокращением, например:

СТАБИЛИЗАТОР (ЭЛЕКТРОТЕХН.)

СТАБИЛИЗАТОР (ХИМ.)

СТАБИЛИЗАТОР (АВИАЦИЯ)

СМАЗКА (ПРОЦЕСС)

СМАЗКА (ВЕЩЕСТВО).

Допускается в качестве реляторов использовать кодовые обозначения, смысл которых поясняется во введении к ИПТ, например:

СТАБИЛИЗАТОР (45)

СТАБИЛИЗАТОР (61)

СТАБИЛИЗАТОР (55.47),

где в качестве реляторов использованы коды соответствующих рубрик Рубрикатора ГАСНТИ.

2.5.3.2. Лексическое примечание не является частью ЛЕ и представляет собой поясняющий текст на естественном языке, например:

МЛЕКОПИТАЮЩИЕ (виды на суше)

АЭРОПЛАН (только для документов до 1940 г.)

СПЛАВЫ МЕДИ (сплавы, в которых медь является основным компонентом).

2.5.3.1, 2.5.3.2. (Введены дополнительно, Изм. № 1).

2.5.4. ЛЕ объявляются эквивалентными в ИПТ, если замена одной ЛЕ на другую не приводит к изменению смысла текста, существенному для поиска информации.

(Измененная редакция, Изм. № 1).

2.5.4.1. Эквивалентность устанавливается между:

1) абсолютными синонимами:

ИПТ=ИНФОРМАЦИОННО-ПОИСКОВЫЙ ТЕЗАУРУС
 ГЕОМЕТРИЯ РИМАНА=РИМАНОВА ГЕОМЕТРИЯ
 НАГРЕВ=НАГРЕВАНИЕ
 КОНУСНЫЙ=КОНИЧЕСКИЙ=В ФОРМЕ КОНУСА

2) стилистическими синонимами (если различие в сфере употребления не существенно для задач ИПТ), например:

ПНЕВМОНИЯ=ВОСПАЛЕНИЕ ЛЕГКИХ
 АДГЕЗИЯ=ПРИЛИПАНИЕ
 ЛЕНТА ЛИПКАЯ=ЛЕНТА КЛЕЙКАЯ

3) относительными синонимами (если случаи несовпадения значений несущественны для задач ИПТ), например:

СТОЛ=ДИЕТА=ПИТАНИЕ
 БЮРО=КОНТОРА=ФИРМА
 ВИНТ=БОЛТ.

2.5.4.2. Допускается установление эквивалентности также между ЛЕ, различными по значению, но семантически связанными в тех случаях, когда отождествление этих понятий полезно для задач ИПТ, например:

УСТОЙЧИВОСТЬ=НЕУСТОЙЧИВОСТЬ
 ТОРГОВЛЯ=ПРОДАЖА
 РЕКА=РУЧЕЙ
 МАСЛО=СМАЗКА.

2.5.4.1, 2.5.4.2. (Введены дополнительно, Изм. № 1).

2.5.5. *Выбор дескрипторов*

Выбор дескрипторов производится при формировании ИПТ 1-го типа.

2.5.5.1. Для выбора дескрипторов в классах эквивалентности выбирается одна ЛЕ в качестве представителя всего класса.

Если в классе эквивалентности есть существительные, то представителем класса объявляется одно из них.

Если в классе эквивалентности более одного существительного, а также в тех случаях, когда существительные отсутствуют, критериями выбора представителя являются:

полнота выражения смыслового значения данного класса ЛЕ;
 краткость и понятность.

Если в классе эквивалентности более одного существительного, а также в тех случаях, если в классе эквивалентности существительных нет, критериями выбора представителя класса являются:

полнота выражения значений данного класса ЛЕ;
 соответствие стандартам и рекомендациям по научно-технической терминологии;

краткость и понятность;

при прочих равных условиях — наибольшая частота использования (распространенность) в документах и запросах.

(Измененная редакция, Изм. № 1).

2.5.5.2. Каждая ЛЕ получает статус дескриптора или аскриптора.

Статус аскриптора присваивают в следующих случаях:

ЛЕ входит в класс эквивалентности, но не является представителем класса;

значение ЛЕ целесообразно при поиске информации представить как пересечение значений других ЛЕ, получивших статус дескриптора;

значение ЛЕ целесообразно при поиске информации представить как объединение значений других ЛЕ, несовместимых в одном поисковом образе.

В остальных случаях ЛЕ получает статус дескриптора.

2.5.5.3. Словарные статьи ЛЕ, получивших статус аскриптора, должны быть снабжены ссылками на заменяющие их дескрипторы.

Примеры:

1. Аскриптор — синоним дескриптора

Языковедение	Фарси
см ЯЗЫКОЗНАНИЕ	см ПЕРСИДСКИЙ ЯЗЫК

2. Аскриптор, заменяемый комбинацией дескрипторов

Русские тексты	Алфавитно-цифровые знаки
ик РУССКИЙ ЯЗЫК	ик БУКВЫ
ТЕКСТЫ	ЦИФРЫ

3. Аскриптор, заменяемый несовместимыми дескрипторами

Искусственные языки	
на ИНФОРМАЦИОННО-ПОИСКОВЫЕ ЯЗЫКИ	
ЯЗЫКИ МЕЖДУНАРОДНОГО ОБЩЕНИЯ	
ЯЗЫКИ ПРОГРАММИРОВАНИЯ	

(Измененная редакция, Изм. № 1).

2.5.5.4. Если статус аскриптора получает ЛЕ, представляющая класс эквивалентности, все ЛЕ, входящие в этот класс, получают статус аскриптора, и в их словарных статьях должны быть идентичные ссылки на заменяющие дескрипторы.

Пример:

Персидские тексты	Тексты на фарси
ик ПЕРСИДСКИЙ ЯЗЫК	ик ПЕРСИДСКИЙ ЯЗЫК
ТЕКСТЫ	ТЕКСТЫ

2.5.5.5. Статьи дескрипторов снабжаются ссылками на заменяемые ими аскрипторы.

Примеры:

1. ЯЗЫКОЗНАНИЕ	ПЕРСИДСКИЙ ЯЗЫК
с лингвистика	с фарси
языковедение	

2. РУССКИЙ ЯЗЫК	ТЕКСТЫ
ск русские тексты	ск персидские тексты
	русские тексты
	тексты на фарси

3. ИНФОРМАЦИОННО-ПОИСКОВЫЕ ЯЗЫКИ ЯЗЫКИ ПРОГРАММИРОВАНИЯ
 са искусственные языки са искусственные языки

(Измененная редакция, Изм. № 1).

2.5.6. Установление иерархических и ассоциативных отношений

2.5.6.1. Для дескрипторов устанавливаются парадигматические отношения, отражающие лексико-семантические связи между понятиями, выражаемыми дескрипторами. Связь указывают путем внесения в дескрипторную статью ссылки, включающей обозначение согласно таблице ссылок и связанный дескриптор.

2.5.6.2. Основными типами связей являются:

род—вид,
 часть—целое,
 причина—следствие,
 сырье—продукт,
 административная иерархия,
 процесс—объект,
 функциональное сходство,
 процесс—субъект,
 свойство—носитель свойства,
 антонимия.

2.5.6.3. Отношения в ИПТ (исключая синонимию, рассмотренную в п. 2.5.5) могут быть разделены на 2 класса: иерархические и ассоциативные.

2.5.6.4. Иерархические отношения обладают свойствами транзитивности и антисимметричности, которые могут быть использованы при избыточном индексировании в интересах повышения эффективности информационного поиска. Поэтому является предпочтительным вводить отношения как отношения иерархического вида, если они обладают этими свойствами. Применяемые в ИПТ иерархические отношения могут дифференцироваться на отдельные виды.

2.5.6.5. В ИПТ должно фиксироваться наличие иерархического отношения дескрипторов указанием связей в соответствии с таблицей ссылок.

2.5.6.6. Родовидовая связь устанавливается между двумя дескрипторами, если объем понятия нижестоящего дескриптора входит в объем понятия вышестоящего дескриптора.

Примеры:

МЕБЕЛЬ	СТУЛЬЯ	СТОЛЫ
нв СТУЛЬЯ	вр МЕБЕЛЬ	вр МЕБЕЛЬ
СТОЛЫ		

2.5.6.7. Связь часть—целое устанавливается между двумя дескрипторами в том случае, если нижестоящий дескриптор обозначает компонент объекта, обозначаемого вышестоящим дескриптором.

Примеры:

АВТОМОБИЛИ

нч КУЗОВ

КУЗОВ

ШАССИ

вц АВТОМОБИЛИ

КУЗОВ

вц АВТОМОБИЛИ

2.5.6.8. Если для одного дескриптора можно указать более одного непосредственно вышестоящего дескриптора, в иерархических отношениях должны быть установлены связи со всеми дескрипторами.

Пример:

ТЕЛЯТА

в КРУПНЫЙ РОГАТЫЙ СКОТ

МОЛОДНЯК ЖИВОТНЫХ

КОМНАТА

вр ПОМЕЩЕНИЕ

вц ДОМ.

(Измененная редакция, Изм. № 1).

2.5.6.9. При установлении иерархических отношений должны быть указаны связи со всеми нижестоящими дескрипторами независимо от аспекта деления. Аспект деления может быть указан в примечании при ссылке.

Пример:

АНТЕННЫ

н ВЫСОКОЧАСТОТНЫЕ АНТЕННЫ /диапазон/

ШИРОКОПОЛОСНЫЕ АНТЕННЫ /диапазон/

ПАРАБОЛИЧЕСКИЕ АНТЕННЫ /конструкция/

ЩЕЛЕВЫЕ АНТЕННЫ /конструкция/.

2.5.6.10. При необходимости иерархические связи в дескрипторных статьях могут быть указаны на два и более уровня. В таких случаях уровни иерархии номеруют относительно заглавной ЛЕ.

Пример:

ПОЛК

в1 ДИВИЗИЯ

в2 АРМИЯ

н1 БАТАЛЬОН

н2 РОТА

БАТАРЕЯ

2.5.6.11. Ассоциативное отношение является объединением других отношений, не входящих в иерархические отношения или в отношения синонимии. Допускается включать в ассоциативное отношение все виды отношений, кроме синонимии и отношения род—вид.

2.5.7. В целях обеспечения ведения ИПТ и индексирования документов ссылки, фиксирующие связи ЛЕ в словарных статьях,

должны быть взаимными, т. е. для каждой ссылки, указывающей связь заглавной ЛЕ с другой ЛЕ, в словарной статье другой ЛЕ должна быть обратная ссылка. Если нецелесообразно использовать обратную ссылку при поиске информации, следует применять технологическую обратную ссылку «сравни», обеспечивающую ведение ИПТ.

2.5.7.1. В дескрипторной статье указывают все синонимы заглавного дескриптора. Для каждого синонимичного дескриптора указывают один заменяющий дескриптор.

2.5.7.2. Для каждой пары дескрипторов, связанных недифференцированной иерархической связью (ссылки «выше» и «ниже») или иерархической связью по отношению род—вид (ссылки «выше—род» и «ниже—вид»), в статье вышестоящего дескриптора должна быть ссылка на нижестоящий, а в статье нижестоящего — на вышестоящий.

2.5.7.3. Для дескрипторов, связанных иерархической связью по отношению часть — целое, может быть дана ссылка от вышестоящего к нижестоящему или наоборот. В этих случаях обратная иерархическая ссылка заменяется неиспользуемой при поиске технологической ссылкой.

2.5.7.4. Для указания дескрипторов в статьях дескрипторов, заменяемым по ссылкам «иа» и «ик», используются соответственно ссылки «са» («сравни альтернативную ссылку») и «ск» («сравни комбинацию»).

(Измененная редакция, Изм. № 1).

2.6. Построение лексико-семантического указателя

2.6.1. Лексико-семантический указатель является упорядоченной последовательностью словарных статей ИПТ и формируется путем расположения их в алфавитном порядке заглавных ЛЕ.

Устанавливается следующий приоритет символов при алфавитном расположении лексических единиц:

пробел и знаки препинания, кроме дефиса (приоритеты всех знаков препинания и их сочетаний считаются одинаковыми),

русские буквы,

цифры,

латинские буквы,

буквы других алфавитов и специальные символы.

При этом лексические единицы, написанные через дефис, занимают то же алфавитное место, что и при слитном написании.

2.6.2. Идентификаторы могут быть выделены в отдельные списки, упорядоченные по алфавиту и размещенные после основного перечня словарных статей.

2.7. Построение систематического указателя

2.7.1. Систематический указатель является перечнем дескрипторов, сгруппированных согласно принятой в ИПТ рубрикации.

2.7.2. Систематические указатели ИПТ подразделяются на три типа:

тематические,
категориальные,
смешанные.

2.7.3. При построении систематического указателя тематического и смешанного типа в его тематической части следует использовать рубрики Рубрикатора ГАСНТИ или рубрикатора конкретной АСНТИ, совместимого с Рубрикатором ГАСНТИ.

2.7.4. При построении систематического указателя категориального и смешанного типа в его категориальной части следует использовать следующие общие категории:

названия дисциплин и отраслей деятельности;
предметы, материалы;
методы, процессы, операции, явления;
свойства, величины, параметры, характеристики;
отношения, структуры, модели, законы, правила, абстрактные понятия.

2.7.5. В систематическом указателе категориального типа каждый дескриптор должен быть отнесен только к одной рубрике.

2.7.6. Внутри каждой рубрики дескрипторы располагаются в алфавитном порядке.

2.8. Построение указателя иерархических отношений

2.8.1. Иерархический указатель отражает полную структуру иерархических отношений.

2.8.2. Иерархический указатель представляет собой перечень списков дескрипторов, причем каждый список начинается с дескриптора, не имеющего вышестоящих. После каждого дескриптора приведены непосредственно нижестоящие дескрипторы с указанием их уровня в иерархии путем применения нумерации либо графического обозначения уровня.

Пример:

АРМИЯ	АРМИЯ
. ДИВИЗИЯ	1 ДИВИЗИЯ
.. ПОЛК	2 ПОЛК
... БАТАЛЬОН	3 БАТАЛЬОН
.... РОТА	4 РОТА
... ДИВИЗИОН	3 ДИВИЗИОН
.... БАТАРЕЯ	4 БАТАРЕЯ

2.9. Построение пермутационного указателя

2.9.1. Пермутационный указатель предназначен для поиска ЛЕ по отдельным словам, входящим в текст ЛЕ, в том числе и по тем, которые стоят не в начале ЛЕ.

2.9.2. Пермутационный указатель представляет собой перечень списков, каждый из которых соответствует одному из знамена-

тельных слов в составе ЛЕ и включают все ЛЕ, содержащие это знаменательное слово. Перечень списков упорядочен по алфавиту знаменательных слов, а внутри списки упорядочены по тексту ЛЕ.

2.9.3. Если число словосочетаний в ИПТ невелико, допускается указывать пермутации словосочетаний непосредственно в лексико-семантическом указателе. При этом словосочетания в канонической форме приводятся в качестве дескрипторов, а другие варианты — в качестве синонимичных аскрипторов.

2.10. Автоматизация построения ИПТ

2.10.1. Построение тезаурусов с использованием ЭВМ позволяет автоматизировать следующие процессы построения ИПТ:

- частотный анализ,
- коррекция статей,
- алфавитная сортировка словника,
- проверка взаимности и непротиворечивости ссылок,
- составление указателей,
- распечатка в требуемых формах.

2.10.2. Автоматизированное построение ИПТ должно допускать его представление в коммуникативном формате на магнитной ленте и в виде комплекта указателей, отражающих текущее состояние машинного тезауруса и напечатанных с помощью ЭВМ.

2.10.3. Состав, структура и форма представления указателей, напечатанных с помощью ЭВМ, должны соответствовать п. 2.11.

2.11. Структура, состав и форма представления тезауруса

В состав ИПТ входит вводная часть, основная часть (лексико-семантический указатель) и дополнительные части (систематический, пермутационный, иерархический и др. указатели и списки специальных категорий ЛЕ).

Обязательными составными частями являются вводная часть и лексико-семантический указатель.

Допускается в состав ИПТ вводить приложения, содержащие дополнительные сведения о разработке и использовании ИПТ.

2.11.1. Вводная часть включает титульный лист и введение.

2.11.1.1. На титульном листе должны быть приведены:

- наименование ИПТ, включающее термин «информационно-поисковый тезаурус» и указывающее область его применения;
- наименование организации-разработчика и дополнительные сведения об авторстве ИПТ;
- сведения о переиздании;
- место и год создания или издания ИПТ.

Оформление титульного листа приведено в приложении 1.

2.11.1.2. Введение должно содержать следующие данные:

- цель создания и область применения тезауруса, описанную кодами и наименованиями рубрик Рубрикатора ГАСНТИ;

ссылки на источники, использованные для сбора лексики ИПТ, включая перечень тезаурусов, представленных в АСВИЯ, с их регистрационными номерами;

ссылки на нормативно-технические и методические документы, использованные при составлении ИПТ;

описание порядка составления тезауруса, включая обоснование представительности использованных источников лексики;

описание состава и структуры ИПТ;

перечень отношений ЛЕ и методические основания для их установления;

перечень всех символов и специальных сокращений, допущенных для представления ЛЕ;

порядок алфавитного расположения ЛЕ (расположение букв разных алфавитов, небуквенных символов и др.);

количественные характеристики тезауруса (общее число статей, число дескрипторов, аскрипторов и др.);

описание состава и формы представления дополнительных данных в словарных статьях;

отдельный абзац следующего содержания:

Тезаурус подготовлен в соответствии с ГОСТ 7.25—80.

2.11—2.11.1.2. **(Измененная редакция, Изм. № 1).**

2.11.1.3. Введение к последующим изданиям (версиям) ИПТ дополнительно должно содержать следующие данные:

обоснование необходимости составления новой версии тезауруса;

указание на характер внесенных изменений.

2.11.2. *Форма представления дескрипторных и аскрипторных статей в лексико-семантическом указателе*

2.11.2.1. Дескрипторная статья состоит из заглавного дескриптора, списка дескрипторов и аскрипторов, семантически связанных с ним, с обозначением видов связи.

2.11.2.2. Аскрипторная статья состоит из аскриптора и заменяющих его при обработке и поиске информации дескрипторов или комбинации дескрипторов.

2.11.2.3. Словарная статья может включать дополнительные данные, например,

частоту использования дескриптора,

кодировый номер дескриптора,

код дескриптора по систематическому указателю,

дополнительные семантические и лексикографические пометы, иноязычные эквиваленты.

Дополнительные данные располагаются непосредственно за заглавным дескриптором и должны быть отделены от него.

Допускается указание дополнительных данных также при незаглавных ЛЕ.

2.11.2.4. Для указания языка иноязычных эквивалентов используется трехбуквенный латинский код языка по СТ СЭВ 251—76.

2.11.2.5. В рамках дескрипторной статьи термины располагаются в следующем порядке:

заглавный дескриптор,
дополнительные данные,
лексическое примечание,
аскрипторы или дескрипторы-синонимы,
вышестоящие дескрипторы,
нижестоящие дескрипторы,
ассоциативные дескрипторы,
дескрипторы, связанные другими видами отношений.

2.11.2.6. Внутри каждой группы лексических единиц, связанных с заглавным дескриптором одним видом парадигматических отношений, должен быть алфавитный порядок расположения. Если незаглавные ЛЕ имеют примечания, предусмотренные п. 2.5.6.9, они могут быть сгруппированы по алфавитному порядку примечаний.

2.11.2.7. В дескрипторной статье должны быть приведены все непосредственно связанные с заглавным дескриптором вышестоящие и нижестоящие дескрипторы. Иерархические связи на втором и следующих уровнях могут быть указаны только при необходимости.

2.11.2.8. Для указания связи между лексическими единицами тезауруса должны быть использованы ссылки, обозначение и упорядочение которых установлено таблицей ссылок (п. 2.5.1).

Допускается также указывать связи других типов, смысл которых должен быть пояснен во вводной части ИПТ. Обозначение этих ссылок должно отличаться от установленных в настоящем стандарте.

2.11.2.9. Обозначение ссылки в словарной статье может быть приведено только один раз перед первой ЛЕ, связанной с заглавным дескриптором данным отношением.

2.11.2.10. Каждая ЛЕ в словарной статье занимает новую строку. Если она не помещается на строке, все последующие строки с ее продолжением должны быть сдвинуты относительно начала ЛЕ.

2.11.2.11. При печатном издании ИПТ необходимо соблюдать следующие требования:

дескрипторы должны быть выделены шрифтом или иным способом, позволяющим легко отличать их от аскрипторов и лексических примечаний;

реляторы должны быть напечатаны в круглых скобках в конце дескриптора тем же шрифтом, что и остальной текст дескриптора, через пробел (допускается другой вид скобок, отличных от косых);

лексические примечания должны быть напечатаны после текста дескриптора в косых скобках через пробел;

заглавная лексическая единица статьи должна выделяться левым сдвигом на расстояние, равное шести строчным буквам относительно начала незаглавной ЛЕ.

Примеры:

1. Дескрипторная статья
АЛГОРИТМИЧЕСКИЕ ЯЗЫКИ /формализм теории алгоритмов
 см. ТЕОРИЯ АЛГОРИТМОВ

с языки алгоритмические
 машиноориентированные языки
 проблемноориентированные языки
 в ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ
 ФОРМАЛЬНЫЕ ЯЗЫКИ
 н АВТОКОДЫ
 АЛГОЛ
 ИПЛ
 КОБОЛ
 КОМИТ
 ПЛ/1
 ФОРТРАН
 а АЛГОРИТМЫ
 ПРОГРАММИРОВАНИЕ
 ср искусственные языки

2. Аскрипторные статьи

1) Языки программирования
 см АЛГОРИТМИЧЕСКИЕ ЯЗЫКИ

2) Языки
 исп а ФОРМАЛЬНЫЕ ЯЗЫКИ
 ЕСТЕСТВЕННЫЕ ЯЗЫКИ

3) Алфавитно-цифровые знаки
 исп к БУКВЫ
 ЦИФРЫ

2.11.3. ИПТ представляют в печатном виде, на микроформах и на машиночитаемых носителях.

При представлении ИПТ на магнитной ленте формат записи должен соответствовать ГОСТ 7.47—84. Обязательным является оформление вводной части в напечатанном виде. Распечатка ИПТ и его фрагментов должна соответствовать требованиям оформления, установленным настоящим стандартом.

Представление ИПТ на видеотерминалах должно соответствовать основным требованиям оформления, установленным настоящим стандартом.

Представление ИПТ на микроформах должно соответствовать требованиям к печатной форме, установленной настоящим стандартом, и требованиям стандартов на микроформы.

(Измененная редакция, Изм. № 1).

3. ЭКСПЕРТИЗА И РЕГИСТРАЦИЯ ТЕЗАУРУСОВ

3.1. Разработанные и оформленные в соответствии с настоящим стандартом тезаурусы направляются на регистрацию.

3.2. Экспертиза тезаурусов на соответствие настоящему стандарту и регистрация тезаурусов осуществляются в соответствии с порядком, установленным Государственным комитетом СССР по науке и технике.

3.3. Зарегистрированные тезаурусы направляются в Автоматизированную систему ведения информационных языков (АСВИЯ) которая осуществляет их хранение и справочно-информационно-обслуживание по фонду тезаурусов.

**ОФОРМЛЕНИЕ ТИТУЛЬНОГО ЛИСТА
ИНФОРМАЦИОННО-ПОИСКОВОГО ТЕЗАУРУСА**

наименование министерства (ведомства)

наименование организации, разработавшей тезаурус

УДК _____ Регистрационный номер _____

Рубрикатор ГАСНТИ _____

ТЕЗАУРУС
информационно-поисковый
ПО

область применения

место и год издания

(Измененная редакция, Изм. № 1).

Соответствие требований ГОСТ 7.25—80 СТ СЭВ 174—85

ГОСТ 7.25—80		СТ СЭВ 174—85	
Пункт	Содержание требований	Пункт	Содержание требований
2.3	Регламентируется обращение в АСВИЯ и использование Рубрикатора ГАСНТИ и др. классификаций для сбора массива лексических единиц	—	—
2.4.3	Регламентируется приведение существительных к именительному падежу	2.2.2	Регламентируется приведение лексических единиц к словарной форме в соответствии с национальной практикой
2.4.4	Регламентируется приведение прилагательных и причастий к именительному падежу и мужскому роду		То же
2.4.5	Регламентируется приведение глаголов к форме инфинитива		»
2.4.8, 2.5.3	Регламентируется форма сокращений слов	—	—
2.7.3	Регламентируется применение Рубрикатора ГАСНТИ при построении систематического указателя	5.5	Рекомендуется применение рубрикатора соответствующей системы НТИ
2.11.1.2	Регламентируется ссылка на источники, полученные в АСВИЯ	—	—
3	Регламентируется проведение экспертизы и регистрации тезаурусов	—	—

(Введено дополнительно, Изм. № 1).

Редактор Р. С. Федорова
Технический редактор Э. В. Митяй
Корректор С. И. Ковалева

Сдано в наб. 08.06.88 Подп. в печ. 26.07.88 1,5 усл. п. л. 1,5 усл. кр.-отт. 1,18 уч.-изд. л.
Тираж 20 000 Цена 5 коп.

**Ордена «Знак Почета» Издательство стандартов, 123840, Москва, ГСП,
Новопроспектский пер., д. 3.**

Вильнюсская типография Издательства стандартов, ул. Даряус и Гирено, 39. Зак. 1862.