

# **СИСТЕМАТИЗАЦИЯ И ПРЕДМЕТИЗАЦИЯ ДОКУМЕНТОВ**

**ГОСТ 7.59-90 ИНДЕКСИРОВАНИЕ ДОКУМЕНТОВ. ОБЩИЕ  
ТРЕБОВАНИЯ К СИСТЕМАТИЗАЦИИ И ПРЕДМЕТИЗАЦИИ**



**СИСТЕМАТИЗАЦИЯ ДОКУМЕНТА – ЭТО**  
**ОПРЕДЕЛЕНИЕ КЛАССИФИКАЦИОННОГО**  
**ИНДЕКСА ПРОИЗВЕДЕНИЯ ПЕЧАТИ ИЛИ**  
**ДРУГОГО ДОКУМЕНТА (СЛАЙДЫ,**  
**ДИАПОЗИТИВЫ, ВИДЕО, ДИСКИ, CD-ROM И**  
**Т.Д.)**

# **Основные принципы систематизации:**

**научность,**

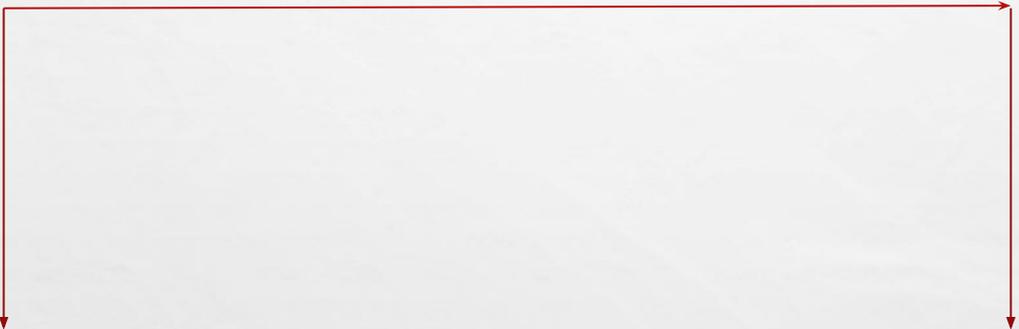
**единообразие,**

**информативность классификационных решений.**

**Объектом систематизации может быть отдельное произведение печати (издание) или другой документ, его составная часть (статья, глава и т.д.). Неопубликованные материалы (рефераты, доклады и т.д.), аудиовизуальные документы, электронные издания**

**Задача систематизации- с необходимой полнотой представить содержание документа в виде классификационных индексов, предметных рубрик, набора ключевых слов в целях создания и использования электронного и систематических каталогов, организации ресурсных фондов.**

***Из задач, которые стоят перед систематизацией,  
вытекают основные функции систематизации.***



**Организация поиска в систематических и электронных каталогах.**

**организация систематической расстановки фонда.**

**Выполняется систематизация по рабочим таблицам классификации, которые отражают особенности структуры и степень детализации конкретного систематического каталога и систематической расстановки определенных фондов.**

**Полный индекс показывает, в каком отделе (или отделах) систематического (электронного) каталога эта книга отражена. Он включает индексы всех подразделений каталога, в которых должна быть отражена книга. Все эти индексы соединяются знаком + (плюс), причем на первом месте указывается индекс отдела, в котором находится документ на полке, а потом остальные.**

# **КЛАССИФИКАЦИОННЫЙ ИПЯ**

**ИНФОРМАЦИОННО-ПОИСКОВЫЙ ЯЗЫК (ИПЯ) - ФОРМАЛИЗОВАННЫЙ ИСКУССТВЕННЫЙ ЯЗЫК, ПРЕНАЗНАЧЕННЫЙ ДЛЯ ИНДЕКСИРОВАНИЯ ДОКУМЕНТОВ, ИНФОРМАЦИОННЫХ ЗАПРОСОВ И ОПИСАНИЯ ФАКТОВ С ЦЕЛЬЮ ПОСЛЕДУЮЩЕГО ХРАНЕНИЯ И ПОИСКА.**

**Основными элементами ИПЯ являются алфавит, лексика и грамматика.**

**Алфавит** ИПЯ — система знаков, используемых для записи слов и выражений ИПЯ.

**Лексика**, или словарный состав, ИПЯ — совокупность слов, словосочетаний и выражений, используемых для построения текстов ИПЯ.

**Грамматика** ИПЯ — совокупность средств и способов построения, изменения и сочетания лексических единиц. Грамматика включает морфологию и синтаксис.

## Требования, которым должен удовлетворять ИПЯ:

- располагать лексико-грамматическими средствами для точного отображения темы документа и запроса;
- не содержать полисемии, синонимии и омонимии;
- отображать только объективные характеристики предметов и отношений между ними;
- быть удобным для алгоритмического сопоставления ПОД и ПОЗ.

Построение выражений ИПЯ требует решения по крайней мере двух проблем. Первая из них связана с **выбором лексических единиц ИПЯ**, необходимых для построения выражений.

**Выбор слов** определяется их смыслом, обусловленным парадигматическими отношениями между предметами и явлениями, которые они определяют.

**Парадигматические отношения**— это отношения, обусловленные наличием логических связей между предметами и явлениями, обозначенными данными словами.

## Парадигматические отношения:

- «вид—род», например, «шкаф—мебель»;
- «часть—целое», например, «лезвие—нож»;
- «причина—следствие», например, «лампа—свет»;
- «функциональное сходство», например, «лопата—экскаватор».

Естественный язык обладает высокой многозначностью. В ИПЯ недопустима многозначность. Поэтому здесь необходимо учитывать отношения синонимии и омонимии слов естественного языка, используемых в ИПЯ.

**Омонимия** — это совпадение слов по написанию или звучанию и несовпадение по смыслу.

**Полисемия** слова состоит в том, что одно и то же слово выражает пучок родственных понятий.

Например, слово «соль» обозначает вещество, а также понятие смысла. Оба значения близки по сути.

**Синонимия** — это совпадение слов по значению и несовпадение по написанию.

**Вторая проблема построения фраз ИПЯ связана с определением последовательности выбранных слов.**

**Синтагматические отношения** — отношения слов при соединении их в словосочетания и фразы. Для уточнения смысла документа или запроса, помимо ключевых слов, часто необходимо указывать, в каких синтагматических отношениях эти слова находятся. Так, фраза «защита окружающей среды от человека» и фраза «защита человека от окружающей среды» имеют совершенно разный смысл, хотя и состоят из одних и тех же ключевых слов.

Многообразие используемых в ИПЯ парадигматических и синтагматических отношений определяет семантическую силу ИПЯ.

# **СТРУКТУРА КЛАССИФИКАЦИОННЫХ СИСТЕМ.**

**Классификационная система** -любая упорядоченная система в соответствии с признаком совокупность классов.

**Класс** - совокупность предметов, процессов, явлений, объединенных по одному или нескольким основным классификационным признакам.

**Классификационная структура** - взаиморасположение классов, обусловленное их отношениями подчинения и соподчинения. В основе построения классификационной структуры лежат основной и вспомогательный классификационные признаки. Элементами классификационной структуры являются классификационные деления, образующие вертикальные ряды, горизонтальные ряды.

**Предкоординатные ипя, ипя классификационного типа – ИПЯ**, построенные на принципах предкоординации. Эти языки представлены известными иерархическими классификационными системами и их словарными средствами: рубрикаторами и таблицами классификации (в том числе УДК, МКИ, ГРНТИ, ББК, ТБК и др).

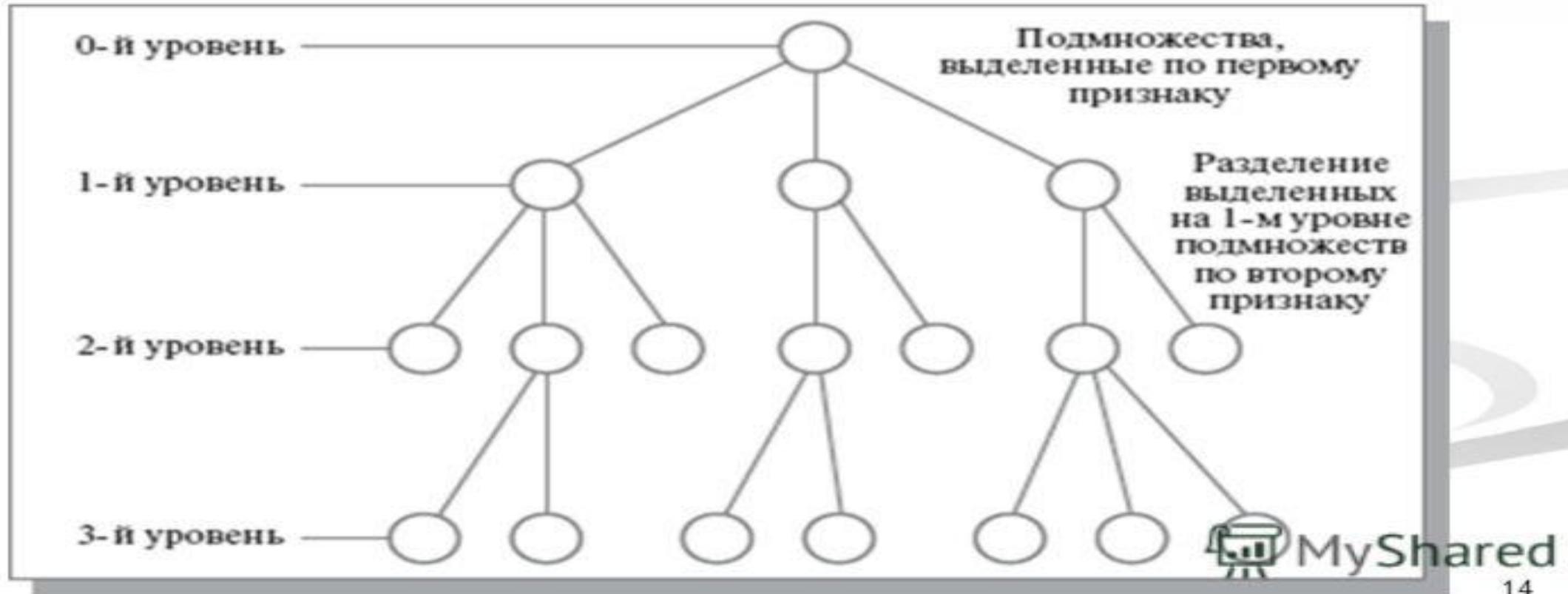
Предкоординация – построение словарного состава ИПЯ до его использования при индексировании, которое характеризуется применением словосочетаний и фраз, выражающих сложные понятия.

***Иерархическая классификация*** – вид классификации предметной области, при которой каждый подкласс имеет один и только один непосредственно предшествующий ему и включающий его класс (отношение сильной иерархии).

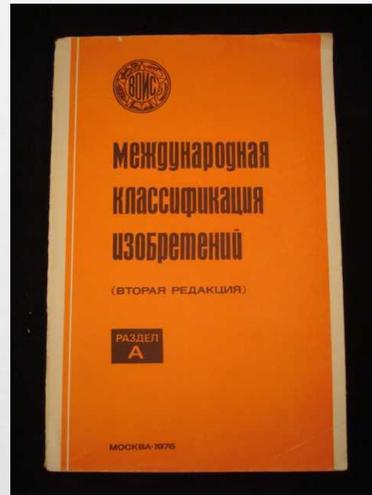
Иерархическая классификация характеризуется строгими правилами формально-логического построения для однозначного определения места документа в предметной области, а также соответствия его другим особенностям классифицируемых документов: виду издания, назначению, языку и т.п.

# Иерархический метод классификации

*Иерархический метод классификации* заключается в последовательном делении заданного множества на подчиненные множества, каждое из которых в свою очередь делится на подчиненные ему подмножества.



**Таблицы классификации** – пособия, предназначенные для определения классификационных индексов в целях перевода содержания документов и запросов с естественного языка на язык информационно-поисковой системы для последующего поиска информации.



**УДК**

**ББК**

**ГРНТИ**

**Иерархическая универсальная комбинационная классификация.**

**Иерархическая универсальная комбинационная система классификации.**

**Иерархическая универсальная перечислительная система классификации.**

**Цель создания**

**Необходимость разработать ИПЯ, чтобы упорядочить весь существующий поток информации.**

**Острая необходимость советских библиотек в классификации, которая позволила бы вести систематизацию колоссальных потоков общественно-политической литературы исходя из идеологии советского гос-ва.**

**Массовый выпуск машиночитаемых Б. Д., их обмен, развитие АИС и АСУ**

- 1. Организация документальных фондов**
- 2. Индексирование документов и запросов**
- 3. Организация и ведение каталогов и картотек**
- 4. Упорядочение записей в б.д.**
- 5. Обеспечение тематического поиска в АИС**

- 1. Организация документальных фондов**
- 2. Индексирование документов и запросов**
- 3. Организация и ведение каталогов и картотек**
- 4. Упорядочение записей в б.д.**
- 5. Обеспечение тематического поиска в АИС**

- 1. Определение тематического охвата органов НТИ в соответствии с их функциями в ГСНТИ.**
- 2. Формирование информационных массивов в органах НТИ с целью обмена.**
- 3. Систематизация материала в информационных изданиях.**
- 4. Выполнение нормативных функций при разработке политематических, отраслевых и проблемных рубрикаторов.**
- 5. Поиск по рубрикам включая адресацию запросов.**

## Состав и структура

**Варианты таблиц УДК – это система таблиц различной степени детализации (полное издание, среднее, сокращенное, отраслевые и рабочие таблицы классификации)**

- 1. Введение**
- 2. Основные таблицы**
- 3. Вспомогательные таблицы**
- 4. АПУ**

- 1. Введение**
- 2. Основные таблицы**
- 3. Вспомогательные таблицы**
- 4. АПУ**

**Система таблиц ББК в постсоветский период:**

- 1. Эталонный, базовый вариант – максимально полные таблицы.**
- 2. Средние таблицы**
- 3. Сокращенные таблицы.**

- 1. Введение**
- 2. Основные таблицы**
- 3. АПУ**

## Авторское право и обновление

**Авторским правом на УДК до 1992 г. обладала Международная Федерация по документации (МФД). Сейчас авторские права на УДК перешли к Консорциуму УДК. В России ВИНТИ имеет исключительное право на распространение документов консорциума на русском языке**

**Авторским правом на ББК располагают 3 национальные библиотеки – соавторы ББК: РГБ, РНБ и БАН.**

**1 раз в 5 лет**

## Область распространения

**УДК является международной системой классификации. Это развитый ИПЯ, объем словаря более 250 тыс. л. е., развитая грамматика, развитая система символов, существует в книжной и в электронной форме.**

**ББК наиболее распространена в библиотеках РФ. ББК – общенациональная классификация и общегосударственное лингвистическое средство. Применяется в библиотеках СНГ и зарубежных стран.**

**ГРНТИ – общегосударственное лингвистическое средство, которое используется одновременно с УДК, ББК и МКИ. Основная сфера применения – органы НТИ. Вариантов не имеет.**

## Особенности

### Различие УДК и ДКД:

- 1. Введение в УДК дополнительных знаков и символов, включая "+", ":", "Г".**
- 2. Основное деление в УДК сохранено, но обязательные 3-х-значные индексы заменены усеченной формой (54, а не 540).**
- 3. Разработана система определителей и вспомогательных таблиц.**
- 4. Расширилась семантика знака ".".**

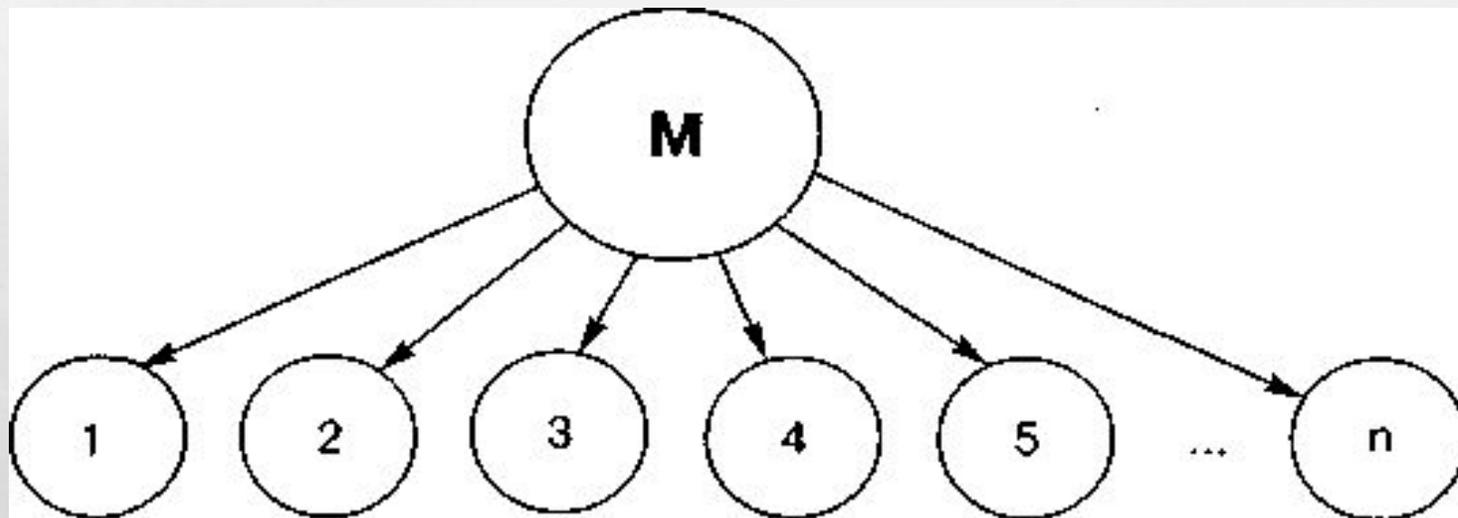
### Идеологизированность ББК.

**В настоящий момент переведена в машиночитаемую форму.**

**Компактность (3 уровня иерархии). ГРНТИ создавался специально для автоматизированной системы НТИ. Отличительные особенности 4-го издания ГРНТИ:**

- 1. Введение новых разделов и понятий.**
- 2. Устранение элементов идеологизации в классификационной схеме.**
- 3. Каждой рубрике ГРНТИ поставлен в соответствие индекс УДК.**
- 4. АПУ подготовлен автоматизированным способом и позволяет найти понятие по любому из терминов, элементов входящих в его состав.**

**ИПЯ фасетного типа** основаны на принципах многоаспектной классификации, в которой каждый конкретный класс строится при индексировании по определенным правилам из предварительно заданных категориальных классов — фасетов. В системах фасетной классификации не ставится задача перечислить все сложные классы. Такие системы предлагают составные элементы, из которых по фасетной формуле составляется индекс.



**Алфавитно - предметная классификация** – это система классов, соответствующих определенной теме и расположенных в алфавитном порядке имен этих классов.

### Основные понятия, характеризующие алфавитно-предметную классификацию



**АЛФАВИТ И ЛЕКСИКА ИПЯ.**

**Основными элементами ИПЯ являются алфавит, лексика и грамматика.**

**Алфавит** ИПЯ — система знаков, используемых для записи слов и выражений ИПЯ.

**Лексика**, или словарный состав, ИПЯ — совокупность слов, словосочетаний и выражений, используемых для построения текстов ИПЯ.

**Грамматика** ИПЯ — совокупность средств и способов построения, изменения и сочетания лексических единиц. Грамматика включает морфологию и синтаксис.

**Парадигматические отношения**— это отношения, обусловленные наличием логических связей между предметами и явлениями, обозначенными данными словами.

**Синтагматические отношения** — отношения слов при соединении их в словосочетания и фразы. Для уточнения смысла документа или запроса, помимо ключевых слов, часто необходимо указывать, в каких синтагматических отношениях эти слова находятся. Так, фраза «защита окружающей среды от человека» и фраза «защита человека от окружающей среды» имеют совершенно разный смысл, хотя и состоят из одних и тех же ключевых слов.

Основными структурными составляющими любого ИПЯ являются алфавит, лексика, грамматика, парадигматические и синтагматические отношения.

Алфавит ИПЯ может включать в свой состав буквы алфавита какого-либо естественного языка (латинского, кириллического и др.), арабские и римские цифры, знаки пунктуации (точка, запятая, двоеточие, кавычки, дефис, круглые и квадратные скобки, апостроф), знаки арифметических операций (плюс, минус, косая черта, знак процента), знаки отношений (больше, меньше, равно), операторы булевой алгебры (знаки конъюнкции, дизъюнкции, отрицания), специальные знаки и символы, например: стрелки, двойное двоеточие, квадрат, знак конгруэнтности и т. д.

Алфавит ИПЯ может быть представлен только цифрами, например УДК, Рубрикатор ГАСНТИ, или только буквами, базируясь на алфавите какого-либо естественного языка, например, язык предметных рубрик, дескрипторный язык, язык библиографических описаний и некоторые другие ИПЯ. Кроме того, алфавит ИПЯ может быть смешанным, буквенно-цифровым. На буквенно-цифровом алфавите основаны фасетная классификация, язык КХ-кодов, СИНТОЛ, семантический код Перри – Кента и др.

Лексические единицы в каждом ИПЯ имеют разные названия:  
в дескрипторном языке - дескрипторы,  
в языке предметных рубрик (алфавитно-предметной классификации) —  
предметные рубрики,  
в фасетной классификации — изолаты,  
в УДК, ББК, рубрикаторах — индексы и соответствующие им словесные формулировки.

В отличие от естественного языка, в качестве лексической единицы ИПЯ может выступать не только отдельное слово, но и устойчивое словосочетание, а также фраза, предложение. В соответствии с этим различают уровни интеграции.

К первому уровню интеграции относятся ИПЯ, лексическими единицами которых являются преимущественно единичные слова и устойчивые словосочетания, например: бетон, индикаторы, квантовая механика, каменный уголь. К числу ИПЯ, характеризующихся лексикой первого уровня интеграции, относятся языки типа «Унитерм», дескрипторные ИПЯ, фасетные классификации.

Вторым уровнем интеграции лексики характеризуются те ИПЯ, лексическими единицами в которых выступают не только отдельные слова и словосочетания, но и фразы, имеющие сложную синтаксическую структуру. На этом уровне находится лексика ИПЯ иерархических классификаций, таких, как УДК, ББК, Рубрикатор ГАСНТИ, Международная классификация изобретений (МКИ), и некоторых других.

Построение лексических единиц ИПЯ может осуществляться двумя способами:

- 1.использованием в качестве лексических единиц специальных кодов, например в УДК, ББК, Рубрикаторе ГАСНТИ и др.;
- 2.использованием в качестве лексических единиц слов и словосочетаний естественного языка, например: язык типа «Унитерм», дескрипторные ИПЯ, язык библиографического описания, язык предметных рубрик.

В соответствии с этим различают словарные и кодированные ИПЯ

**Словарные ИПЯ**— это языки, в которых план содержания совпадает с планом выражения естественного языка. Поэтому для лексических единиц этого типа ИПЯ не требуется перевод на естественный язык. Таковы язык предметных рубрик и дескрипторный язык.

**Кодированные ИПЯ** — это языки, в которых план выражения не совпадает с планом содержания. В этих случаях создаются специальные двуязычные словари, позволяющие коды ИПЯ переводить на естественный язык и обратно. Примером таких языков служат ББК, УДК и другие классификации, рубрикаторы.

**Термины – основа лексики ИПЯ.** Независимо от способа контроля лексики, уровня ее интеграции и способа построения лексических единиц в основе лексики любого ИПЯ лежат термины, так как они являются основными носителями научно-технической информации в текстах документов и запросов. Термины представляют собой наиболее существенный «строительный материал», при помощи которого создается ИПЯ.

Все термины в зависимости от языковых средств выражения можно разделить на две группы: языковые и неязыковые.

**Языковые термины** включают в свой состав термины-слова, термины-словосочетания и термины-предложения.

Термины-слова подразделяются на непроизводные: газ, лак, информация; производные: газовый, лаковый, информационный; сложные: газоносный, лако-красочный, информационнопоисковый; аббревиатуры: ДНК, МБА, ЭВМ.

Термины-словосочетания включают свободные словосочетания, в которых каждый из компонентов — термин и может вступать в двустороннюю связь: квантовая химия, физическая химия, квантовая биология, библиотечная система, библиотечная статистика; несвободные словосочетания, в которых компоненты, взятые изолированно, могут быть и не терминами, и лишь в сочетании друг с другом они образуют термин: библиотечный почерк, авторский лист, белый стих, читальный зал, черный ящик; термины- фразеологизмы, т. е. устойчивые, неделимые словосочетания: роза ветров, роза волнений, кошачьи лапки (морские термины); кошачье золото, кошачье серебро (геологические термины); кармашек книги, ключ к систематическому каталогу, источник информации, информационный шум (библиотечно-информационные термины).

Неязыковые термины представлены специальными знаками, графическими символами. Главная особенность неязыковых средств выражения терминов заключается в их абсолютной условности, искусственности, сознательной закреплённости за понятием.

**В качестве неязыковых средств часто используются буквы греческого алфавита, например в терминологии физики элементарных частиц:  $\alpha$ -частица,  $\alpha$ -распад,  $\rho$ - радиоактивность и т. п.**

**Грамматическая структура терминов.** Экспериментальные исследования терминологии показывают, что наиболее типичными моделями образования терминов являются следующие:

- 1) прилагательное+существительное (библиотечные каталоги, жаропрочные сплавы);
- 2) существительное+существительное (комплектование фондов, катушки индуктивности);
- 3) прилагательное+прилагательное-1-существительное (интегральные информационные системы, дуговые пламенные лампы).

## **Грамматические средства ББК**

**Знак плюс (+) – равнозначен союзу или**

**Косая черта (/) – знак распространения , равнозначно « от и до»**

**Предшествующий знак (-) дефис служит для обозначения специальных типовых делений**

**Запятая (,) – служит для обозначения планов расположения**

**Двоеточие (:) – знак отношений**

# **ТАБЛИЦЫ КЛАССИФИКАЦИИ**

**НА ПРИМЕРЕ ББК**

**Таблицы классификации** – пособия, предназначенные для определения классификационных индексов в целях перевода содержания документов и запросов с естественного языка на язык информационно-поисковой системы для последующего поиска информации.

Таблицы классификации являются материальными представлениями систем классификации, отражающими их структуру, содержание, индексацию.

**Таблицы классификации состоят из основных и вспомогательных таблиц.**

**Основная таблица классификации** - часть классификационной таблицы, включающая классификационные записи, которые в совокупности исчерпывают предметную область.

**Вспомогательная таблица классификации** - часть классификационной таблицы, выделенная из состава основной классификационной таблицы и включающая классификационные записи, которые используются, главным образом, для детализации классов основной классификационной таблицы

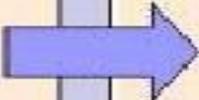
# Индексы ББК

## БИБЛИОТЕЧНО-БИБЛИОГРАФИЧЕСКАЯ КЛАССИФИКАЦИЯ

### Основные деления

- 1
- 2 **Естественные науки +**
- 3 **Техника. Технические науки.**
- 4 **Сельское и лесное хозяйство. Сельскохозяйственные и лесохозяйственные науки.**
- 5 **Здравоохранение. Медицинские науки.**
- 6/8 **Общественные и гуманитарные науки.**
- 9 **Литература универсального содержания.**

- 
- 20 **Естественные науки**
  - 22 **Физико-математические науки.**
  - 24 **Химические науки.**
  - 26 **Науки о Земле (геодезические, геофизические, геологические и географические)**
  - 28 **Биологические науки**

- 
- 26.1 **Геодезические науки. Картография**
  - 26.2 **Геофизические науки.**
  - 26.3 **Геологические науки.**
  - 26.8 **Географические науки**