

**Гендина Н. И., доктор пед. наук, профессор
Кемеровского государственного университета
культуры и искусств**

Структура и назначение информационно-поисковых тезаурусов (ИПТ)

План

1. Тезаурус как способ систематизированного представления знаний и разновидность идеографического словаря
2. ИПТ: назначение и функции
3. Структура ИПТ
4. Порядок разработки, экспертизы, регистрации и ведения ИПТ

Список литературы

1. Большаков И. А. Многофункциональный словарь-тезаурус для автоматизированной подготовки русских текстов / И. А. Большаков // НТИ. Сер.2. – 1994. – №1 – С. 11–23.
2. ГОСТ 7.24–2007. Тезаурус информационно-поисковый многоязычный. Состав, структура и основные требования к построению. – Взамен ГОСТ 7.24-90 ; введ. 2008-07-01 / Межгос. совет по стандартизации, метрологии и сертификации. – Москва : Стандартинформ, 2008. – 7 с.
3. ГОСТ 7.25–2001. Тезаурус информационно-поисковый одноязычный. Правила разработки, структура, состав и форма представления. – Взамен ГОСТ 7.25-80 ; введ. 2002-07-01. – Москва : ИПК Изд-во стандартов, 2001. – 16 с.
4. Жмайло С. В. К вопросу об определении тезауруса / С. В. Жмайло // НТИ. Сер. 1.–2003. – №12. – С.20 – 25.
5. Жмайло С. В. К разработке современных информационно-поисковых тезаурусов / С. В. Жмайло // НТИ. Сер. 1.– 2004. – №1. – С.23 – 31.

Многозначность термина «тезаурус»

1. Тезаурус как идеальный объект – мир личности («тезаурус ребенка», «тезаурус взрослого» и т. п.).
2. Тезаурус как материально существующий объект – словарь.

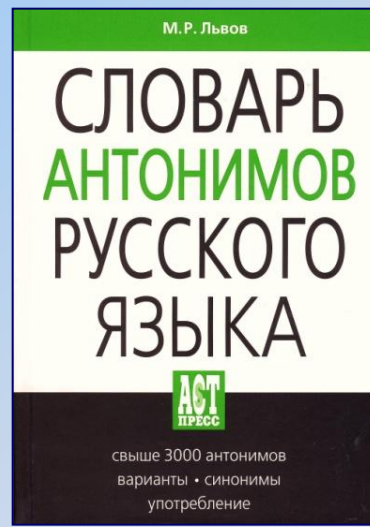
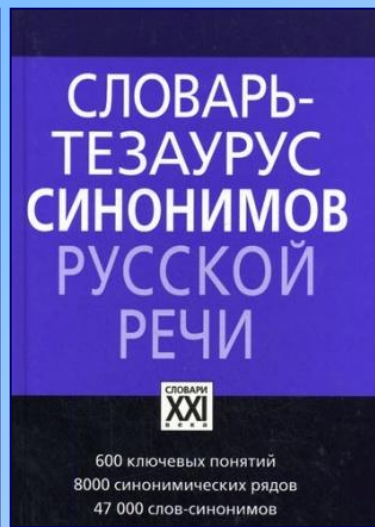
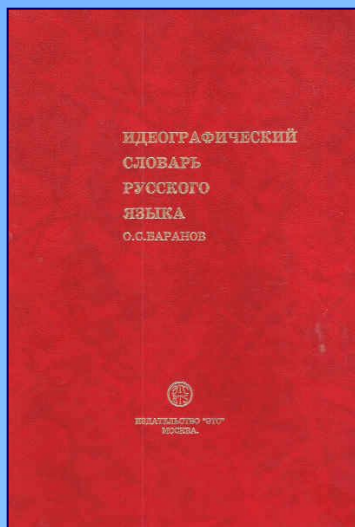
Тезаурус: определение понятия

Тезаурус (от греч. *θησαυρός* – сокровище, клад) – это словарь особого типа (идеографический словарь), в котором слова располагаются не по алфавиту, а по степени смысловой близости.



Тезаурус как идеографический словарь

Тезаурус – особая разновидность словарей общей или специальной лексики, в которых указаны семантические отношения (синонимы, антонимы, паронимы, гипонимы, гиперонимы и т. п.) между лексическими единицами.



Пример идеографического словаря

Лексическая основа русского языка: Комплексный учебный словарь/ под ред. В. В. Морковкина. – Москва : Рус. язык, 1984.

- 1168 с.

(Фрагмент)

1.0 Абстрактные отношения и формы существования материи

1.1. Существование, бытие

1.2. Пространство

1.3. Форма

1.4. Время

.....

2.0. Материальный мир

2.1. Материальный мир: общие понятия

2.2. Неорганический мир

2.3. Органический мир

2.3.4. Человек как живое существо

2.3.4.1. Человек как разумное существо

2.3.4.3.2. Ощущение и восприятие

2.3.4.3.3. Эмоциональные, волевые, интеллектуальные действия и состояния

2.3.4.3.3.1. Чувства и эмоции

2.3.4.3.3.2. Душевный склад человека

2.3.4.3.3.3. Поведение человека

2.3.5. Человек как общественное существо

Пример идеографического словаря

<http://www.glossary.ru/>

Информационное общество – концепция постиндустриального общества; новая историческая фаза развития цивилизации, в которой главными продуктами производства являются информация и знания. Отличительными чертами информационного общества являются:






- увеличение роли информации и знаний в жизни общества;
- возрастание доли информационных коммуникаций, продуктов и услуг в валовом внутреннем продукте;
- создание глобального информационного пространства, обеспечивающего (а) эффективное информационное взаимодействие людей, (б) их доступ к мировым информационным ресурсам и (в) удовлетворение их потребностей в информационных продуктах и услугах.

Новая функция: Формула глоссария

Пример идеографического словаря

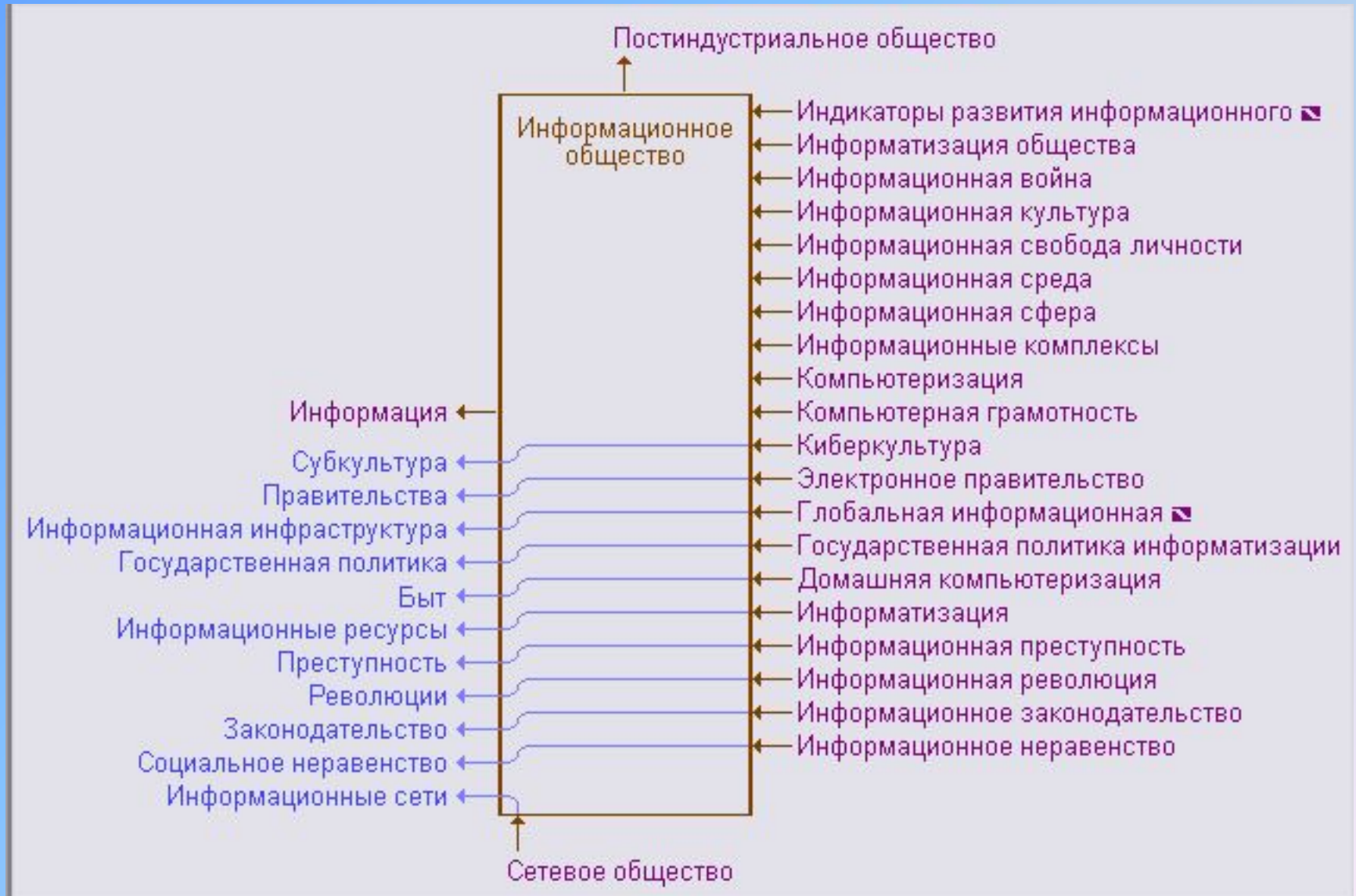
<http://www.glossary.ru/>

Глоссарий / понятие: **Информационное общество**

- представленный прямоугольником **120 x 346**
-  → включенный в **1** родовое понятие
-  → включающий **1** подвид
-  → оперирующий **20** терминами, в т.ч. **10** собственными
-  → задействованный в **1** понятии
-  → пересекающийся с **11** понятиями
- отвечающий формуле **$1 + 1 + 20_{10} + 1 + 11$**

Пример идеографического словаря

<http://www.glossary.ru/>



Пример идеографического визуального словаря (<http://vslovar.org.ru>)



[О проекте](#) | [Помощь](#)

Найти

в словарях в Интернет

[Энциклопедия](#)

[Компьютеры](#)

[Финансы](#)

[Психология](#)

[Право](#)

[Философия](#)

[Культура](#)

[Медицина](#)

[А](#) [Б](#) [В](#) [Г](#) [Д](#) [Е](#) [Ж](#) [З](#) [И](#) [Й](#) [К](#) [Л](#) [М](#) [Н](#) [О](#) [П](#) [Р](#) [С](#) [Т](#) [У](#) [Ф](#) [Х](#) [Ц](#) [Ч](#) [Ш](#) [Щ](#) [Ъ](#) [Ы](#) [Ь](#) [Э](#) [Ю](#) [Я](#)

На правах рекламы: Недвижимость у Черного моря -квартира Сочи . Сам-плюс: упаковочные пакеты - лучшие цены Профессиональные переводчики в бюро переводов . Обширный словарный запас! Профилактика диабет Ваш дом в поселке Яхромские горки Электростанции, бензогенераторы Распродажа мягкая мебель угловые диваны от фабрики Электростанции, бензогенераторы, дизель-генераторы мебель детская комната от фабрики Большой выбор ремонт кпк королев в Москве бизнес тренинг линкомаулия Болезнь, мор и бремя, иго автомойка требования к печати Sex интернет-казино, виртуальное казино, азартные игры, рулетка бассейны спа экскурсии по москве бюро переводов готовые предприятия недвижимость в Испании

Пример идеографического визуального словаря (<http://vslovar.org.ru>)

СЛОВАРИК визуальный

О проекте | Помощь

Энциклопедия | Компьютеры | Финансы | Психология | Право | Философия | Культура | Медицина

А Б В Г Д Е Ж З И Й К Л М Н О П Р С Т У Ф Х Ц Ч Ш Щ Ъ Ы Ь Э Ю Я

Ба Бб Бв Бг Бд Бе Бж Бз Би Бй Бк Бл Бм Бн Бо Бп Бр Бс Бт Бу Бф Бх Бц Бч Бш Бщ Бь Бы Бю Бя

БАЗА ДАННЫХ

БАЗА ДАННЫХ (БД) (database (DB)). Один или несколько специальным образом организованных файлов, хранящих систематизированную информацию, для доступа к которой используются программные средства системы управления базой данных (СУБД). Б. д. представляет собой хранилище связанных между собой данных, не зависящих от программ, которые к ним обращаются. Например, в Б. д. предприятия хранятся конструкционные, технологические, производственные, эксплуатационные, экономические и т. п. характеристики изделий предприятия и их аналогов, выпускаемых конкурирующими фирмами. Такой Б. д. могут пользоваться специалисты из отдела маркетинга (для прогноза конкурентоспособности изделий на рынке), плановики (для определения необходимых объемов выпуска продукции), конструкторы (для анализа конструктивных особенностей изделий конкурентов) и т. д. Б. д. должна быть достаточно полной, правильно организованной, всегда соответствовать текущему моменту времени и удобной для использования. Содержащиеся в ней данные не должны противоречить друг другу. Редактирование, пополнение и исключение данных, их поиск и сортировка осуществляются под управлением СУБД. Б. д. могут быть персональными и коллективного пользования. Сопровождение крупных Б. д. коллективного пользования осуществляет администратор базы данных. Ср. банк данных

Идеографический словарь для термина "БАЗА ДАННЫХ". Центральная концепция "БАЗА ДАННЫХ" связана с "ПРОГРАММНЫМИ СРЕДСТВАМИ", "СУБД", "СОРТИРОВКА", "РЕДАКТИРОВАНИЕ" и "БАНК ДАННЫХ".

- ПРОГРАММНЫЕ СРЕДСТВА** связаны с **ДОКУМЕНТ** и **ДАННЫЕ**.
- СУБД** связан с **СОРТИРОВКА**.
- СОРТИРОВКА** связан с **СОРТИРОВКА ДАННЫХ**.
- РЕДАКТИРОВАНИЕ** связан с **ТЕКСТ**, **ФОРМАТ** и **УДАЛЕНИЕ**.
- БАНК ДАННЫХ** связан с **ФАЙЛ** и **ДОСТУП**.

Информационно-поисковый тезаурус: определение понятия

Информационно-поисковый тезаурус (ИПТ) – нормативный словарь дескрипторного ИПЯ с зафиксированными в нем парадигматическими отношениями (*ГОСТ 7.74-96. Информационно-поисковые языки. Термины и определения*)

Способы представления лексики в ИПТ

ИПТ по технологии машиностроения

ТЕЛЬФЕРЫ 1202

в Механизмы грузоподъемные
Оборудование подъемно-
транспортное периодического
действия

н Тельферы электрические

а Тали электрические

ИПТ ИНИОН по философии

Дескриптор: Абдукция

Вышестоящий: Логический
метод исследования;
Умозаключение

Ассоциативный: Гипотеза;
Индукция; Логика открытия

Рубрика: A0221; A023121

ББК: 87; 87.4

Дескрипторная и аскрипторная статьи в ИПТ

Дескрипторная статья

ТОПЛИВО 1802

с Горючее

н Топливо газообразное

Топливо жидкое

Топливо искусственное

Топливо твердое

а Мазут

Масла

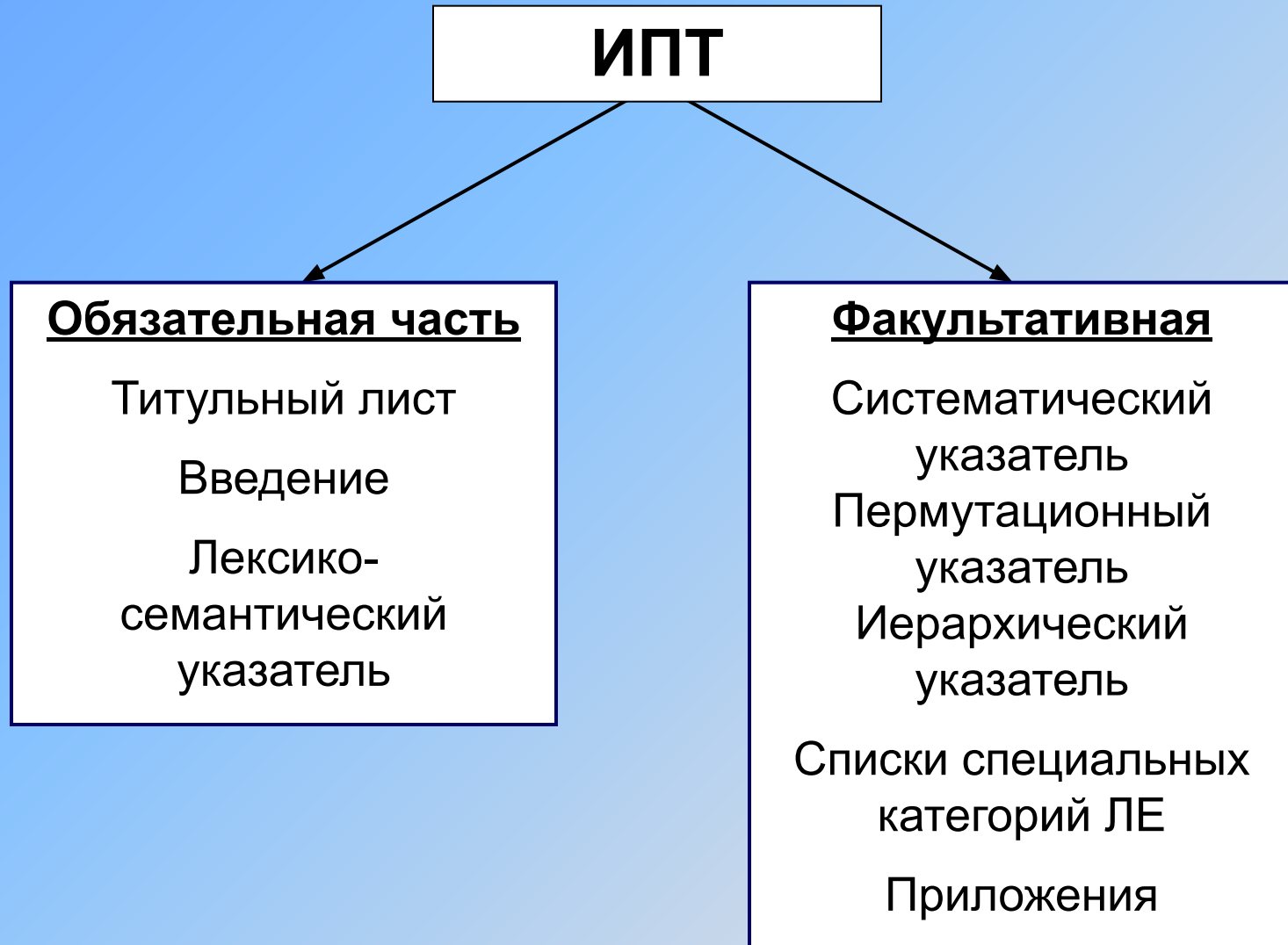
Материалы горюче-смазочные

Аскрипторная статья

Горючее

См. Топливо

Структура ИПТ



Лексико-семантический указатель (фрагмент ИПТ по технологии машиностроения)

ГОРНЫ 2703, 2709

- в Оборудование термическое
- н Горны высокие
 - Горны кузнечные
 - Горны низкие
 - Горны плавильные
- а Вагранки
 - Печи (оборудование)
 - Топки

ГОРНЫ ВЫСОКИЕ 2709

- в Горны

Горючее
см. Топливо

ГОСТ
См.Стандарт

ГОФРИРОВАНИЕ 2703

- в Обработка механическая
- а Накатка
 - Рифление

ГРАВИРОВАНИЕ 2709

- с Гравировка
- а Нанесение рисунка
 - Обработка электроискровая
 - Обработка электрохимическая
 - Чеканка

Гравировка
См. Гравирование

Структура систематического указателя комплекса ИПТ Кемеровского ЦНТИ

Систематический указатель дескрипторов



Дескрипторные области



Дескрипторные области и группы



Перечень дескрипторов по группам

ДЕСКРИПТОРНЫЕ ОБЛАСТИ (на примере ИПТ по технологии машиностроения)

02 Безопасность и аварии

07 Исследования, разработки, испытания

09 Материалы и вещества

11 Общие понятия

12 Промышленное оборудование и технологические процессы

14 Общественные науки

15 Строительство

19 Физика

20 Химия

22 Хранение и тара

25 Электротехника и радиоэлектроника

26 Математика

27 Технология машиностроения

ДЕСКРИПТОРНЫЕ ОБЛАСТИ И ГРУППЫ **(на примере ИПТ по технологии** **машиностроения)**

02 Безопасность и аварии

0201 Безопасность и аварии

0204 Средства обеспечения безопасности

09 Материалы и вещества

0902 Материалы

0904 Металлы

0905 Пластмассы

0907 Сплавы

20 Химия

2002 Химические соединения

2003 Химические элементы

25 Электротехника и радиоэлектроника

2530 Электрооборудование. Электронная техника и оборудование.

ПЕРЕЧЕНЬ ДЕСКРИПТОРОВ ПО ГРУППАМ (на примере ИПТ по технологии машиностроения)

25 ЭЛЕКТРОТЕХНИКА И РАДИОЭЛЕКТРОНИКА

2530 Электрооборудование. Электронная техника и оборудование.

Автотрансформаторы

Аккумуляторы электрические

Блоки питания

Выпрямители

Выпрямители полупроводниковые

Генераторы

Емкость электрическая

Зуммеры

Кабели

Кабели воздушные

Кабели высоковольтные

Взаимосвязь лексико-семантического и систематического указателей ИПТ

Лексико-семантический указатель

**БАББИТЫ НА МЕДНОЙ
ОСНОВЕ 0907**
в Баббиты

ИНГИБИТОРЫ 2002
а Реакции химические

КАБЕЛИ ВОЗДУШНЫЕ 2530
в Кабели

Систематический указатель

ДЕСКРИПТОРНЫЕ ОБЛАСТИ И ГРУППЫ

02 Безопасность и аварии

0201 Безопасность и аварии

0204 Средства обеспечения
безопасности

09 Материалы и вещества

0902 Материалы

0904 Металлы

0907 Сплавы

20 Химия

2002 Химические соединения

2003 Химические элементы

25 Электротехника и радиоэлектроника

2530 Электрооборудование

Электронная техника и оборудование.

Иерархический указатель (на примере ИПТ по этике)

ТЕОРИЯ МОРАЛИ

- Мораль
 - - Моральная оценка
 - - Моральное сознание
 - - Моральные качества
 - - - Морально-волевые качества
 - - - - Безволие
 - - - - Властность
 - - - - Дисциплинированность
 - - - - Мужество
 - - - - Настойчивость

Пермутационный указатель (на примере ИПТ по информатике)

БАЗА

Администраторы базы данных

Базы данных

Базы данных прямого доступа

Базы знаний

Библиографические базы данных

Ведение баз данных

ДОСТУП

Базы данных прямого доступа

Диалоговый доступ

Дистанционный доступ

Пример ИПТ

Денисова А. А.

Тезаурус терминологии гендерных исследований

<http://www.owl.ru/gender/thesaurus.htm>

Состав:

- Алфавитный список
- Лексико-семантический указатель
- Пермутационный список
- Список англоязычных терминов

Денисова А. А.

**Тезаурус терминологии гендерных исследований
АЛФАВИТНЫЙ СПИСОК (фрагмент)**

А

адаптивная фаза гендерной социализации
активизм
амазонки
анархо-феминизм
анатоми-политики тела
андрогиния
андрогинность (а)
андроцентризм
андроцентризм языка
антирасистский феминизм (а)
антропология
антропология женщин
антропоцентризм
ассоциативный образ женщины
ассоциативный образ мужчины

Б

бедность
биодетерминизм
биологический детерминизм (а)
биомедицинские политики тела
биополитики народонаселения
бисексуальность
брак

Денисова А. А.

**Тезаурус терминологии гендерных исследований
ЛЕКСИКО-СЕМАНТИЧЕСКИЙ УКАЗАТЕЛЬ (фрагмент)**

(D) Амазонки

SN В греческой мифологии племя женщин-воительниц. Исторической основой мифа об амазонках является противоборство с воинственными племенами. Сведения об амазонках приводит Геродот в IV книге своей "Истории". Вопрос о том, являлись ли амазонки классическим матриархатным обществом, остается открытым. Археологические открытия свидетельствуют, что в Древнем мире женщины-воительницы существовали в действительности.

RT матриархат, женщины-воины, разделение труда по признаку пола, археологические открытия

LE *Amazons, Amazon women warriors*

Денисова А. А.

Тезаурус терминологии гендерных исследований
ЛЕКСИКО-СЕМАНТИЧЕСКИЙ УКАЗАТЕЛЬ (фрагмент)

(D) Куртуазная любовь

SN "Утонченная любовь", форма внебрачных отношений между мужчиной и женщиной, предполагающая утонченность ухаживания и поведения. Куртуазная любовь известна по литературным памятникам примерно с XI века, расцвет ее приходится на XII-XIII вв. Этот термин ввел Гастон Парис, французский филолог; название указывает на определенный социальный круг (*courtois* (франц.) - учтивый, рыцарский, придворный*).

BT любовь

RT куртуазная литература, куртуазность, Прекрасная Дама, рыцарь

LE *Courtly love*

Денисова А. А.

**Тезаурус терминологии гендерных исследований
ПЕРМУТАЦИОННЫЙ СПИСОК (фрагмент)**

андроцентризм

андроцентризм языка (языковой андроцентризм)

степень андроцентризма языка

антропология

антропология женщин

гендерная антропология

феминистская антропология

антропоцентризм

бедность

феминизация бедности

Денисова А. А.

Тезаурус терминологии гендерных исследований СПИСОК АНГЛОЯЗЫЧНЫХ ТЕРМИНОВ (фрагмент)

А

Activism - активизм
Adaptation phase of gender socialization -
адаптивная фаза гендерной социализации
Advancement of women - защита интересов
женщин
Affirmative actions - позитивные действия
Affirmative discrimination - позитивная
дискриминация
Affirmative discrimination - позитивная
дискриминация
Amazons, Amazon women warriors - амазонки
Anarcha-feminism - анархо-феминизм
Androcentrism - андроцентризм,
фаллологоцентризм
Androcentrism in language - андроцентризм
языка, сексизм языка
Androgyny - андрогиния, андрогинность
Anthropology - антропология
Anthropology of women - антропология женщин

В

Bio-determinism -
биодетерминизм, биологический
детерминизм
Bio-medical politics of body -
биомедицинские политики тела
Bio-politics of population -
биополитики народонаселения
Bisexuality - бисексуальность
Black feminism - "черный"
феминизм, антирасистский
феминизм
Body politics - политики тела

Назначение ИПТ

- Индексирование документов и запросов в АИС, включая избыточное индексирование;
- Обеспечение контроля и нормализации лексики конкретной отрасли знания для формализованного представления информации в АИС;
- Уточнение информационных запросов пользователей за счет явного выражения парадигматических отношений в структуре ИПТ.

Нормативная база разработки, экспертизы, регистрации и ведения ИПТ

- ГОСТ 7.25-2001 «Тезаурус информационно-поисковый одноязычный. Правила разработки, структура, состав и форма представления»
- ГОСТ 7.24-90 «Тезаурус информационно-поисковый многоязычный. Состав структура и основные требования к построению»

Центры экспертизы и регистрации ИПТ

Депозитарные фонды

```
graph TD; A[Депозитарные фонды] --> B[Национальный, включающий ИПТ на русском языке (ВИНИТИ, г. Москва)]; A --> C[Международные]; C --> D[Депозитарный фонд ИПТ на английском языке (библиотека факультета информационных наук Университета в Торонто, Канада)]; C --> E[Депозитарный фонд ИПТ на всех языках, кроме английского (Институт научной и технико-экономической информации в Варшаве, Польша)];
```

Национальный, включающий ИПТ на русском языке (ВИНИТИ, г. Москва)

Международные

Депозитарный фонд ИПТ на английском языке (библиотека факультета информационных наук Университета в Торонто, Канада)

Депозитарный фонд ИПТ на всех языках, кроме английского (Институт научной и технико-экономической информации в Варшаве, Польша)