

Вопросы систематизации терминологии

М.Г.Мальковский,

ВМК

С.Ю.Соловьев

МГУ

(39 слайдов)

Терминология

Друзы – кристаллические включения в виде сростков кристаллов, накапливаемых в вакуолях растительных клеток.

Латы крестца – элемент конской брони, в виде большой попоны, покрывающий круп лошади. Латы крестца соединяются с боевым седлом и боковыми прикрытиями.

Терминоведение

разделы

- теоретическое
- прикладное

+ терминография

- общее | частное
- типологическое
- сопоставительное
- семасиологическое
- ономасиологическое
- историческое
- функциональное
- когнитивное

организации

- КНТ РАН
 - Комитет научной терминологии в области фундаментальных наук
 - бывш. КНТТ РАН
- ВНИИКИ Госстандарта
 - Всероссийский научно-исследовательский институт классификации, терминологии и информации по стандартизации и качеству

Терминоведение-2



Источники терминологии

- авторская
- признанная
 - публикации
 - законы
 - ГОСТ-ы

ГОСТ Р 51303-99 Торговля.
Термины и определения.
(148 терминов)



Термин – элемент терминосистемы

Терминосистема – совокупность единиц специальной номинации некоторой области деятельности, изоморфная системе ее понятий и обслуживающая ее коммуникативные потребности.

Термин – элемент терминосистемы; слово или словосочетание, именующее понятие некоторой области деятельности.

Определение термина – истолкование термина посредством выделения его основных свойств.

Типы определений

1. $\langle A \rangle$ есть $\langle B \rangle +$ особенности $\langle C \rangle$

2. $\langle A \rangle$ есть $\langle B_1 \rangle | \langle B_2 \rangle | \dots | \langle B_n \rangle$

3. $\langle A \rangle$ есть совокупность $\langle B_1 \rangle, \dots \langle B_n \rangle$

*. $\langle A \rangle$ есть

1 | 2 | 3

Различают (подклассы A)

Обычно (типичная ситуация)

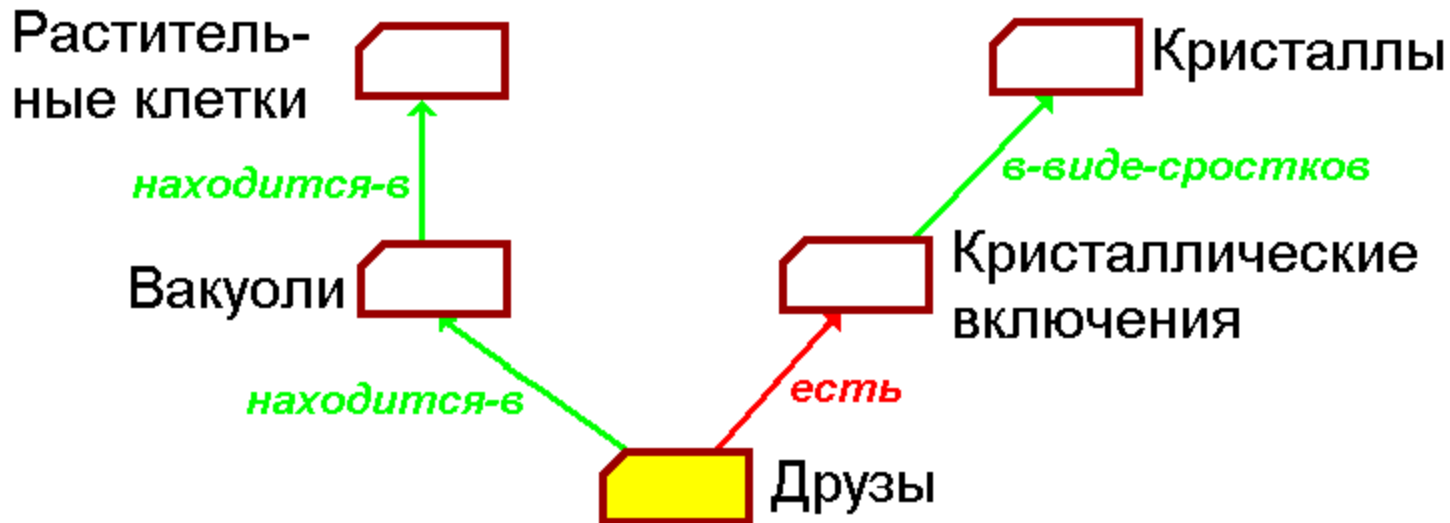
... для (прагматика)

$\langle A \rangle$ в им.падеже; ед. или мн. число 7

Статья (тип 1)

<А> есть <В> + особенности <С>

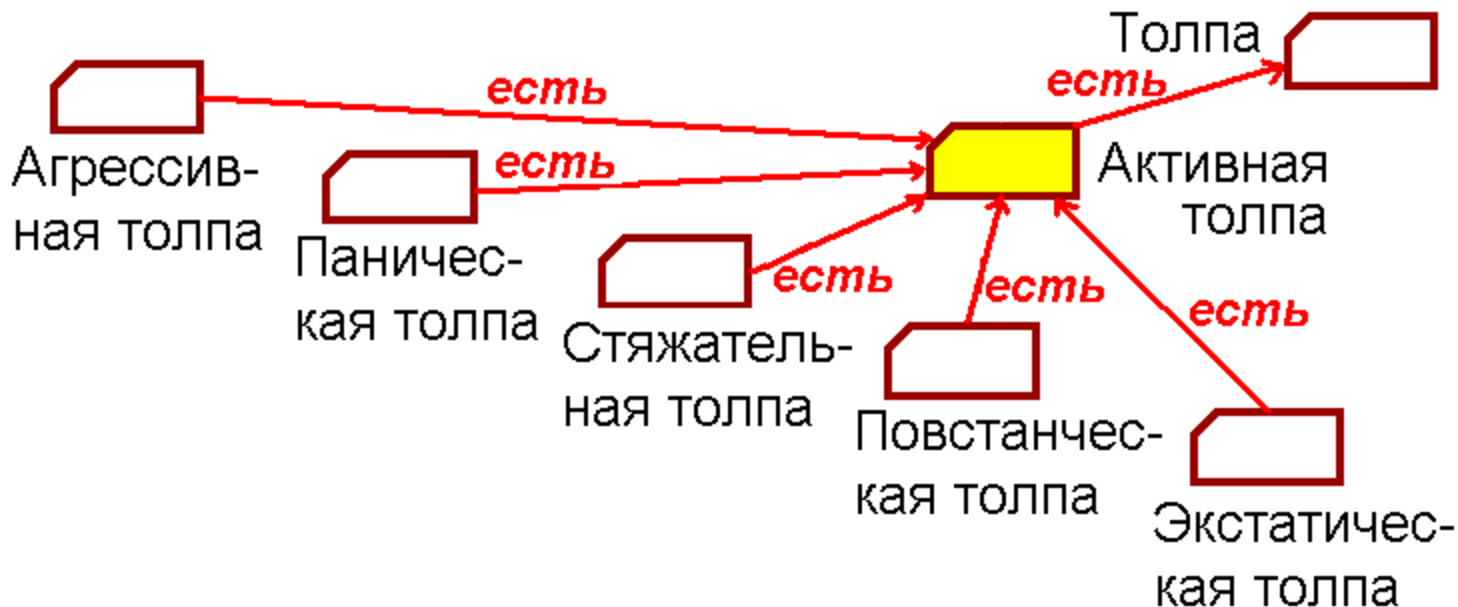
Друзы - кристаллические включения в виде сростков кристаллов, накапливаемых в вакуолях растительных клеток.



Статья (тип 2)

$\langle A \rangle$ есть $\langle B_1 \rangle$ | $\langle B_2 \rangle$ | ... | $\langle B_n \rangle$

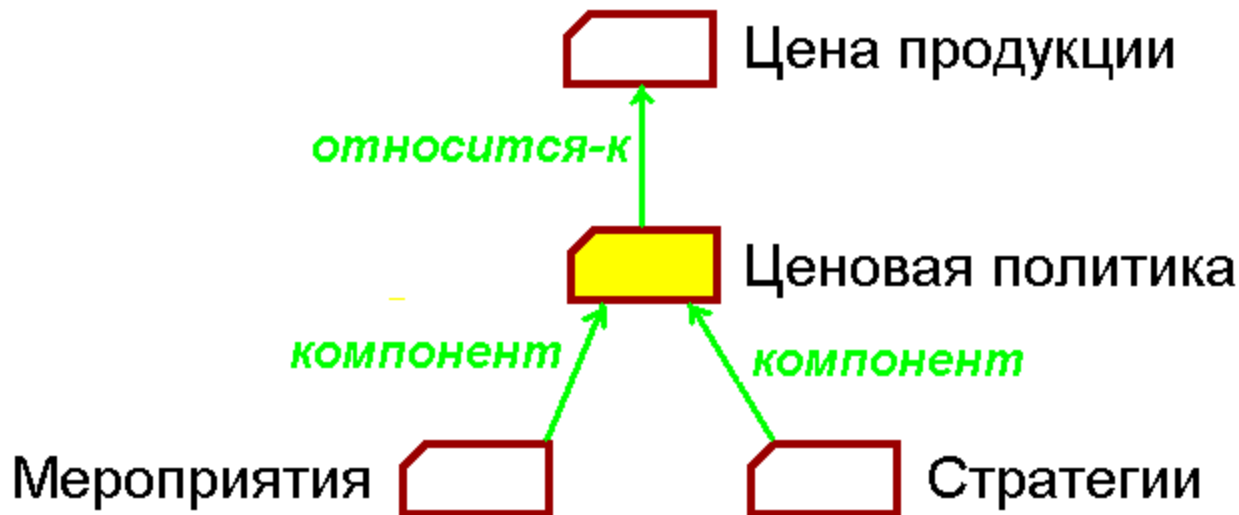
Активная толпа – агрессивная, паническая, стяжательная, повстанческая или экстатическая толпа. (n=5)



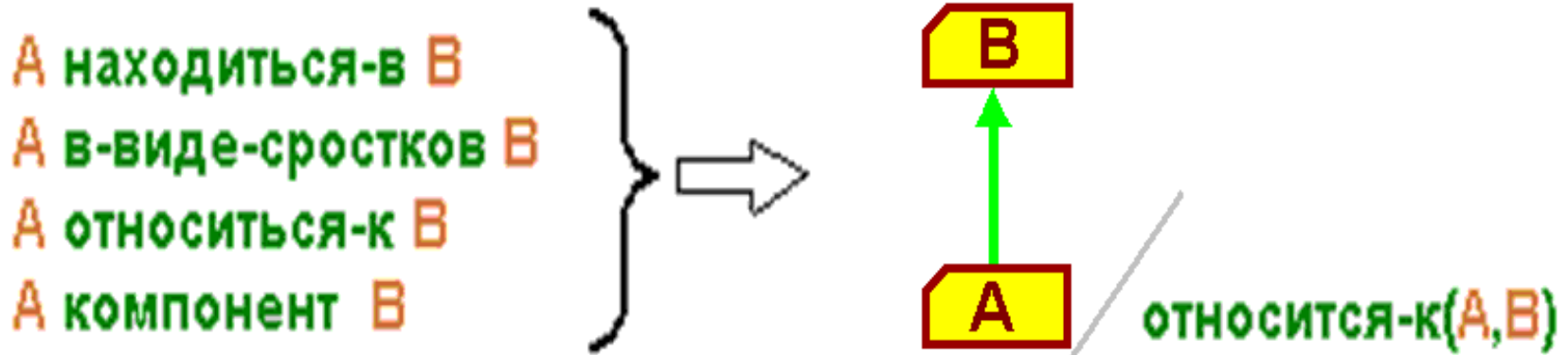
Статья (тип 3)

$\langle A \rangle$ есть совокупность $\langle B_1 \rangle, \langle B_2 \rangle \dots \langle B_n \rangle$

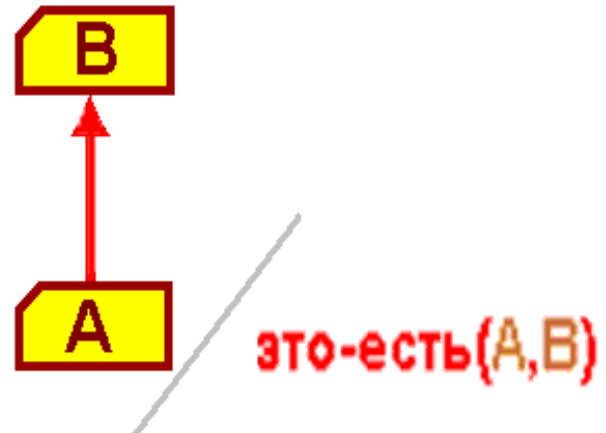
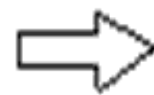
Ценовая политика – совокупность мероприятий и стратегий, которые использует предприятие при установлении цены. (n=2)



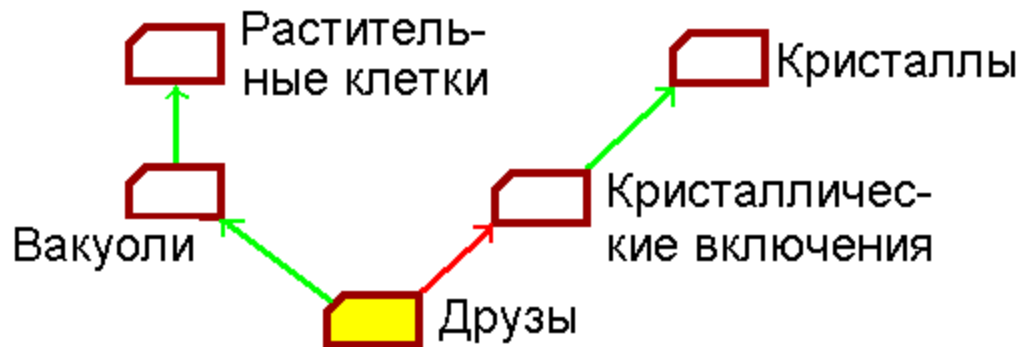
Соглашения



A есть B

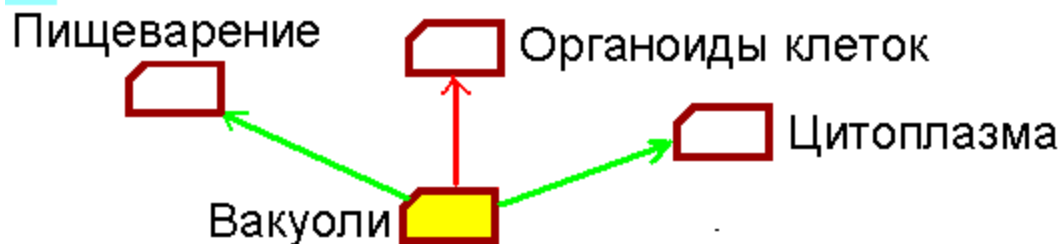


Связывание



<< Было

+



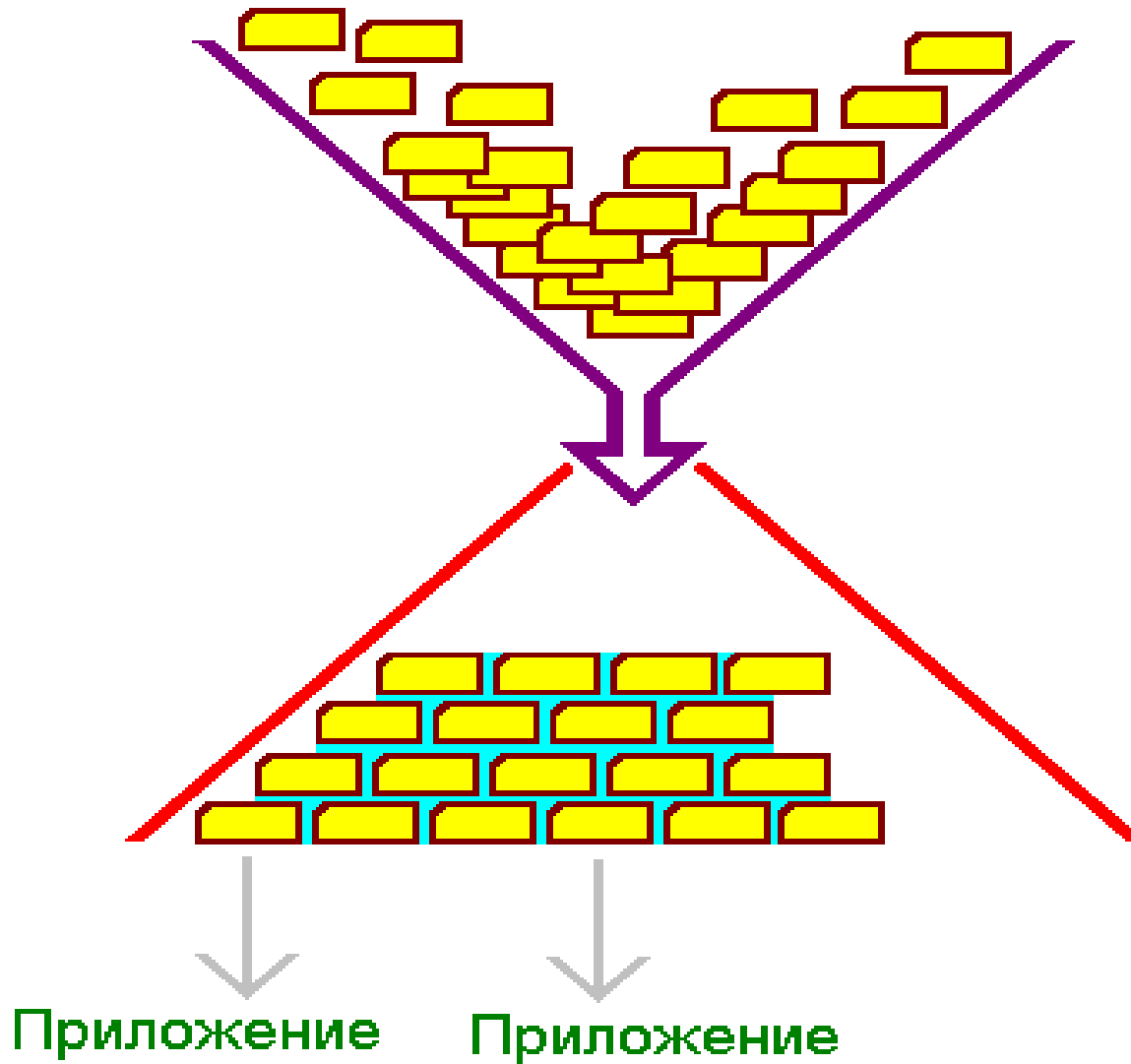
<< Новая статья
+ Редактор +

=

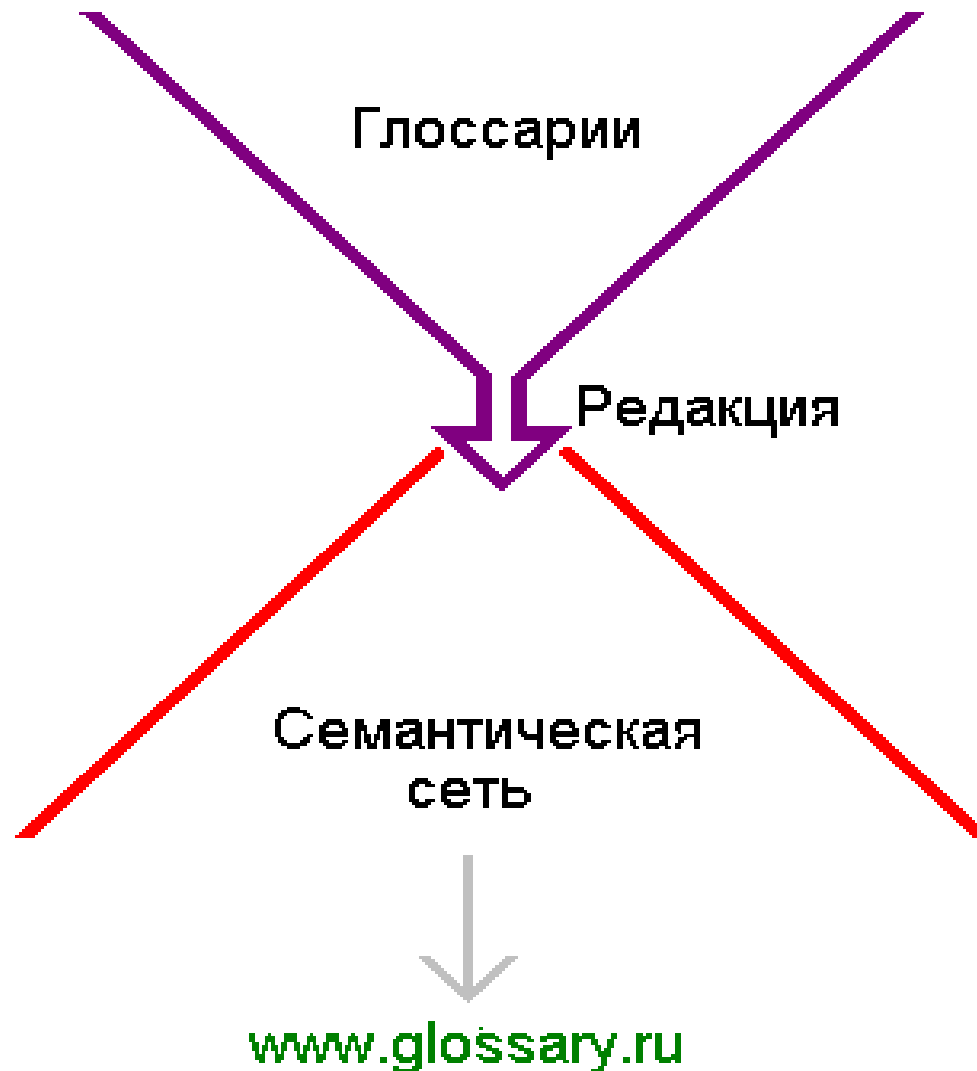


>> Стало

Идея проекта



Идея проекта (уточнение)



Редакция

- Юридическое
лицо
- Штат
научных
редакторов
- Специальное
программное
обеспечение



+ Редактор +

+ Редактор-программа:

Надстройка: поиск и выбор статей фрагмента

Текстовый редактор:

модификация фрагмента

Исполнитель фрагмента

+ Редактор-человек: 50 статей в сутки

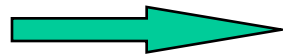
эффект обучения

Понятийные вершины

Понятийная вершина – формально - вершина, в которую входит одно или несколько ребер.

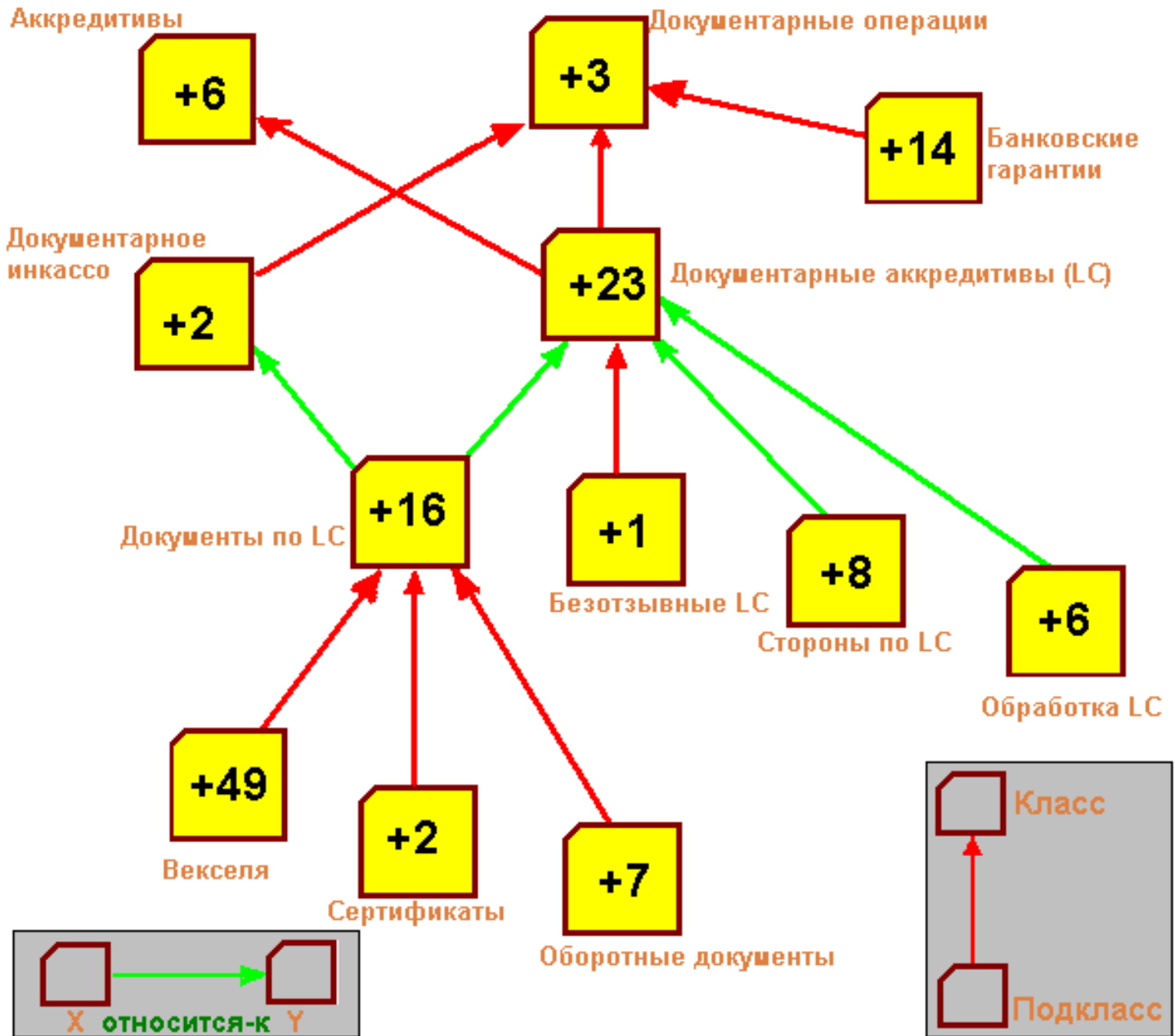
Неформально: Понятийной вершине соответствует класс объектов.

Статья *Кредит*



понятие *Кредиты*

Ф
Р
А
Г
М
Е
Н
Т



Состав сети

Банковское дело Биология Бухгалтерский учет

География Гидрология Информатика Искусство

Культура Психология Страхование Перевозки

Образование Полезные ископаемые Вселенная

Социология Государство Физика Экономика

Объем сети

Количество статей = 53' 477 шт.

в т.ч. «понятий» = 9' 044 шт.

Количество связей = 75' 569 шт.

в т.ч. «это-есть» = 34' 566 шт.

«относится-к» = 41' 003 шт.

Простые свойства сети

1. 1 статья - 1.4 связей

75569 / 53477

2. 6 статей - 1 понятие

53477 / 9044

3. Число 19

4. СВЯЗНОСТЬ

5. Циклы

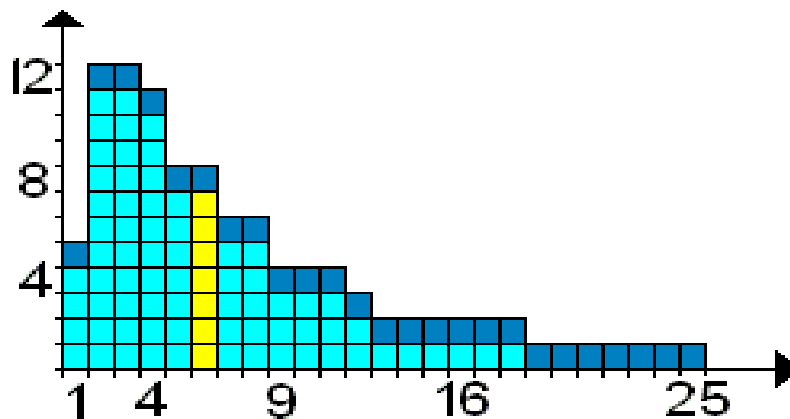
6. Ошибки

7. Национальные особенности: вексеся, эрозия.

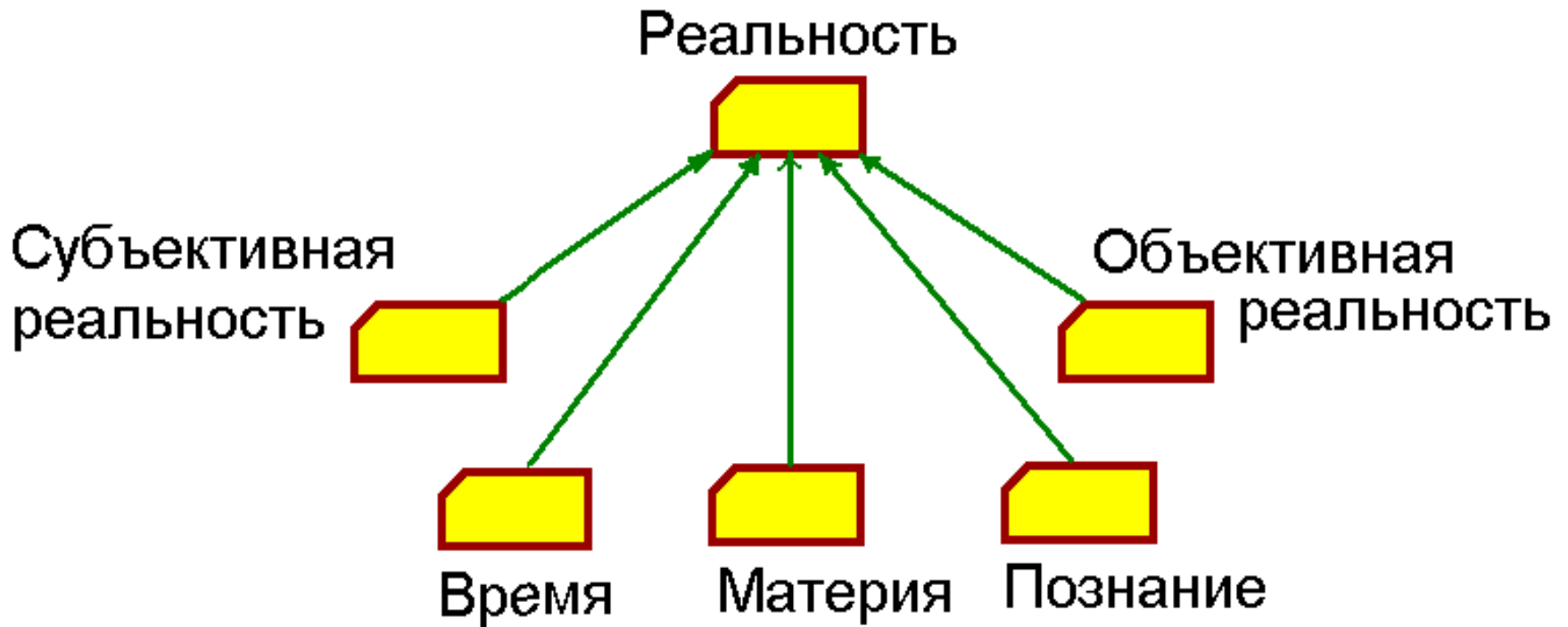
8. Субъективизм (5%)

9. Краевые эффекты: смысловые, структурные

10. 92% уникальной терминологии



Материализм



Прагматика

1. Сеть → Интернет: www.glossary.ru

- задача No.1 Единица показа?

- задача No.2 Навигация?

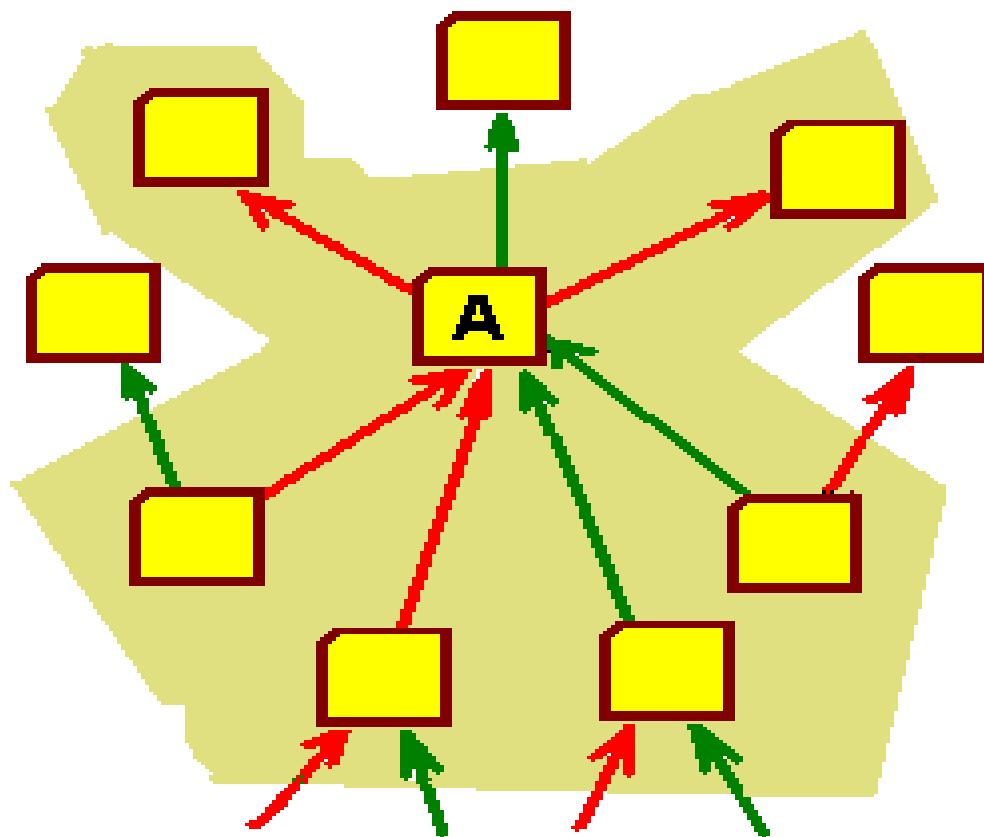
- задача No.3 Интеллектуальные функции?

2. Другие

1. Единица показа: Глоссарий

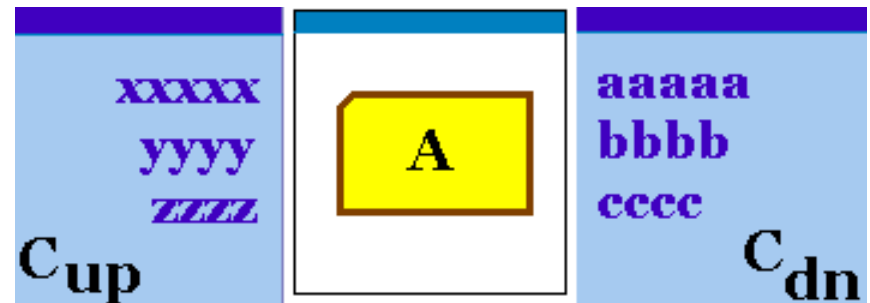
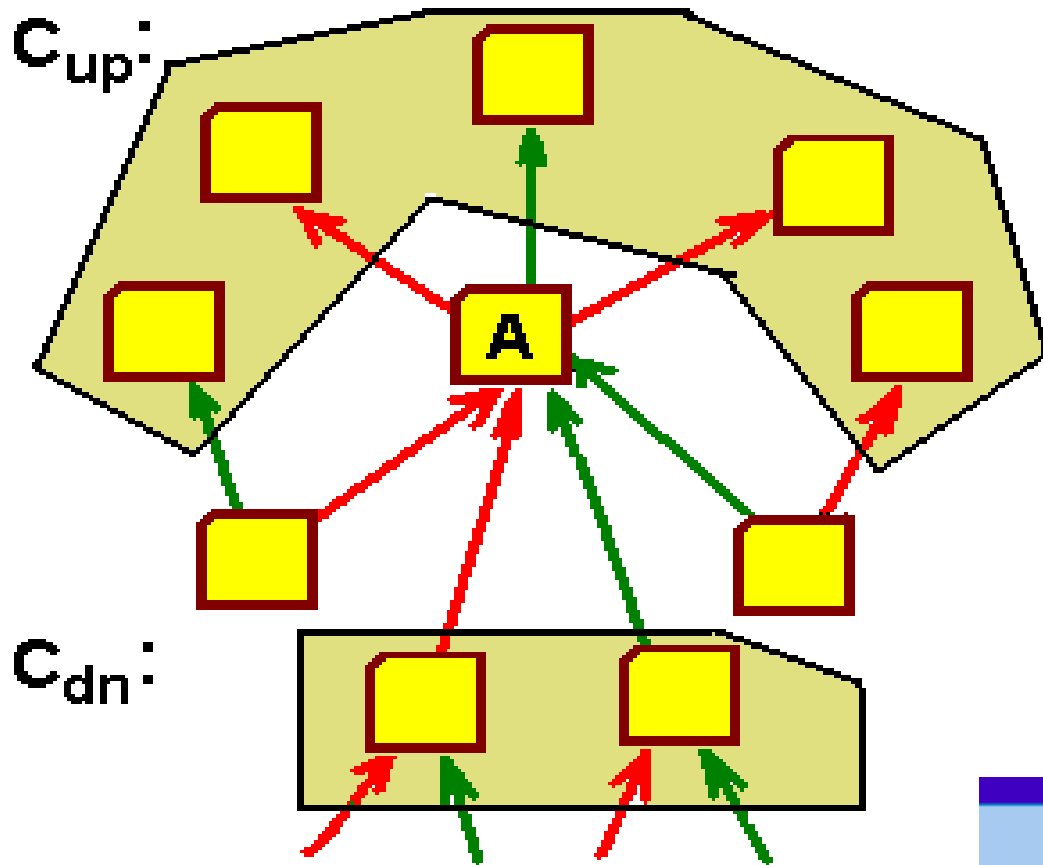
Пусть A – понятийная вершина.

$G(A)$ – множество статей глоссария



2. Навигация: Смежные глоссарии

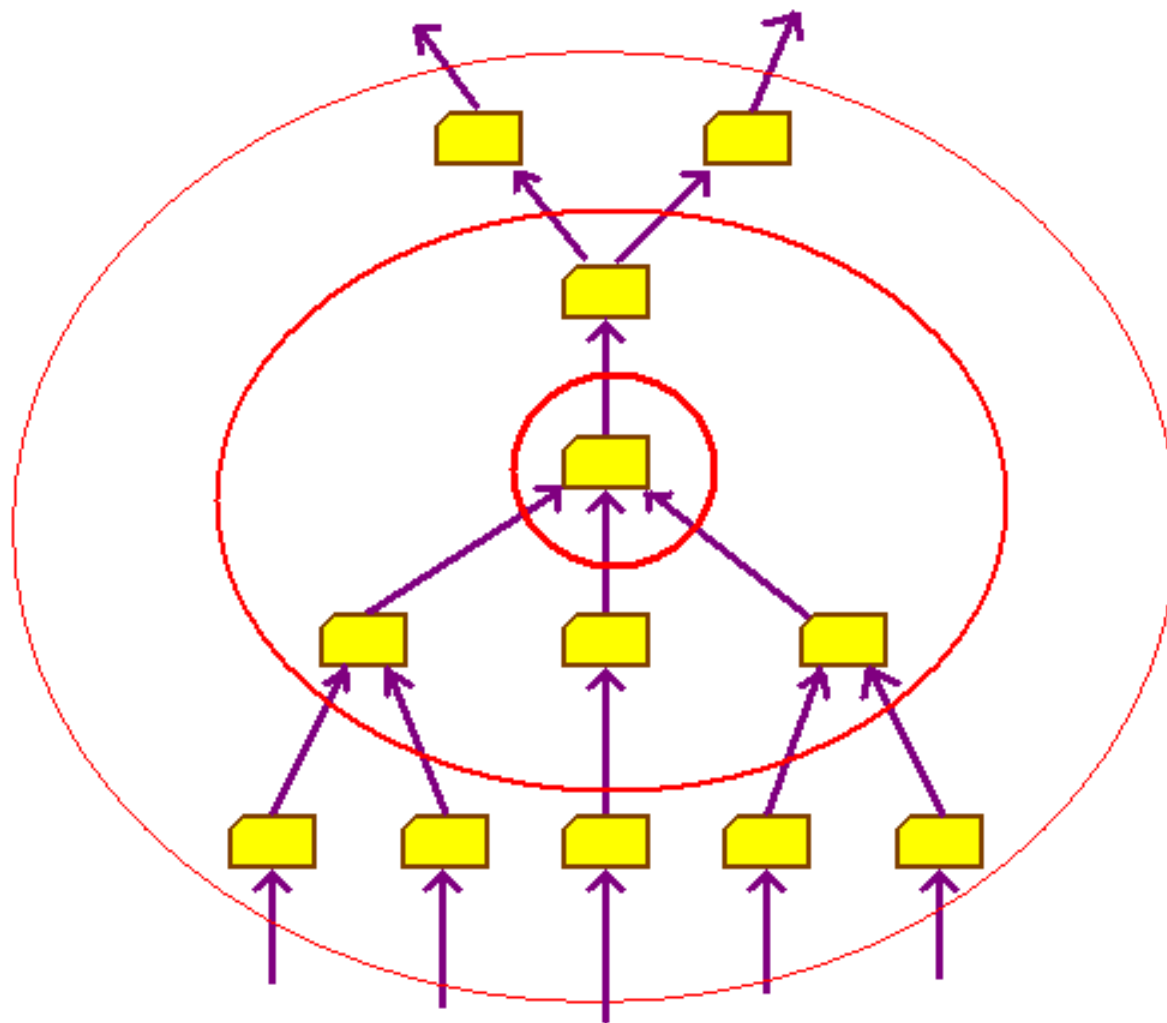
Множества понятий $C_{up}(A)$ и $C_{dn}(A)$



3. Интеллектуальные функции:

1. Тематическая группировка
2. Метод расширения (наследование)
3. Конструктор поисковых запросов к ИПС
4. Веб-карты
5. Визуализация фрагмента сети

3.1 Тематическая группировка

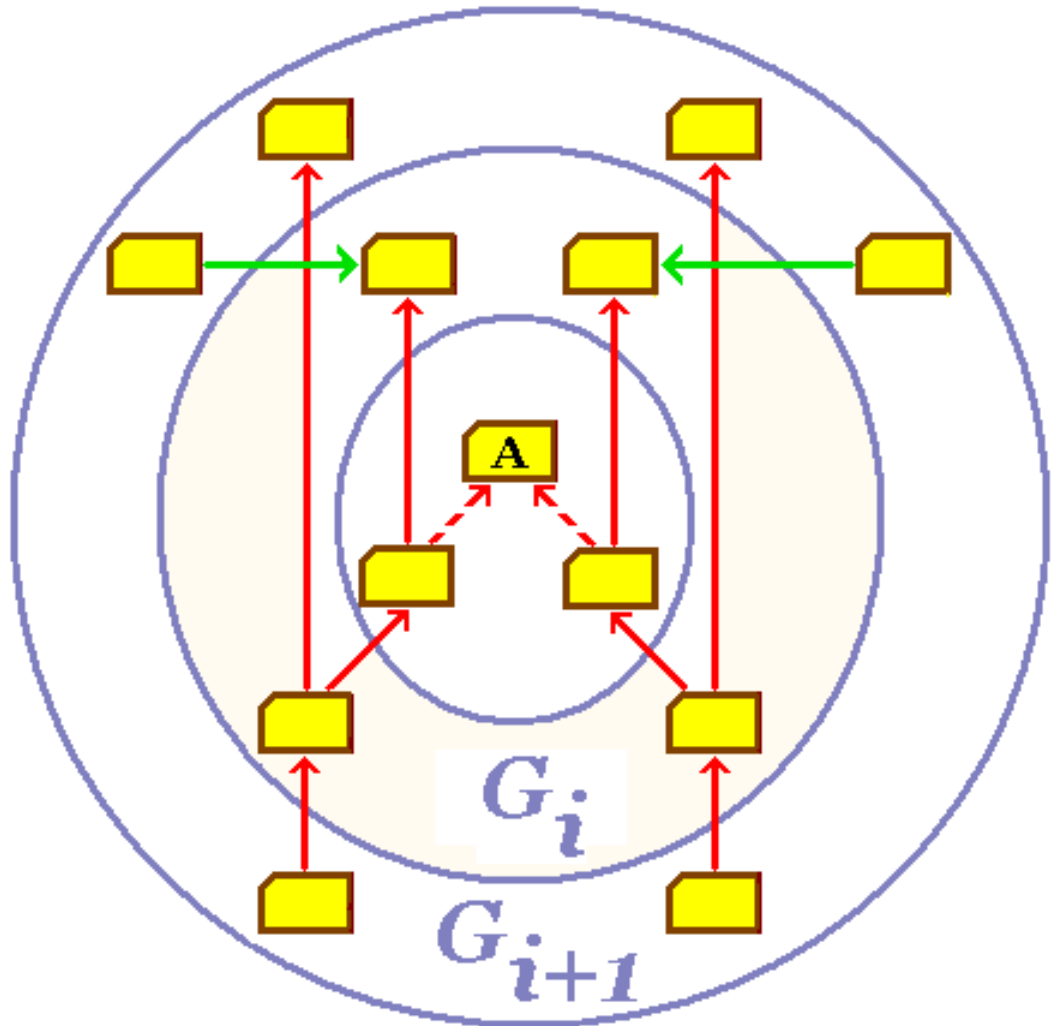


3.2. Метод расширения

$$G_0(A) \subseteq G_1(A) \subseteq \dots \subseteq G_n(A)$$

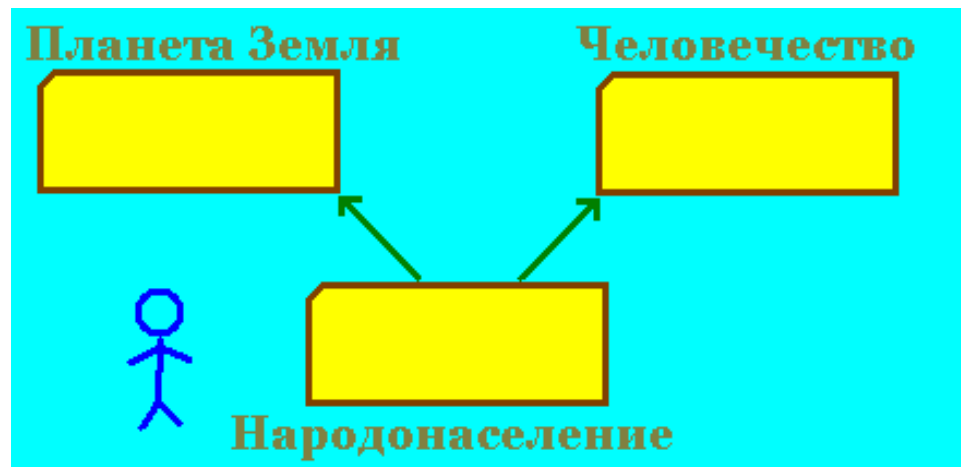
$$G_0(A) = \{ A \}$$

$$G_1(A) = G(A)$$



Наследование

3.3 Запросы к ИПС - 1



Планета Земля

Человечество

Народонаселение

Население
Population

Народонаселение - совокупность людей, живущих на Земном шаре (человечество) или в пределах конкретной территории - континента, страны, области и т.д. Народонаселение непрерывно возобновляется в ходе воспроизводства.

Развернуть глоссарий

Конструктор поисковых запросов к ИПС Яндекс

Я ИЩУ ДОКУМЕНТ, КОТОРЫЙ

содержит не содержит содержит с сохранением порядка слов

В ОДНОМ ПРЕДЛОЖЕНИИ УПОМИНАНИЕ:

народонаселение

ИЛИ ЕГО СИНОНИМЫ:

население
 population

содержат упоминание планета земля и его синонимы

игнорируют упоминание человечество и его синонимы

Найти в Интернет

3.3 Запросы к ИПС - 2

Планета Земля 

Человечество 

Народонаселение

Население
Population

Народонаселение - совокупность людей, живущих на Земном шаре (человечество) или в пределах конкретной территории - континента, страны, области и т.д. Народонаселение непрерывно возобновляется в ходе воспроизводства.

Развернуть
гlossарий

Конструктор

поисковых запросов
к ИПС Яндекс

я ищу ДОКУМЕНТ, который

содержит не содержит содержит с сохранением порядка слов

В ОДНОМ ПРЕДЛОЖЕНИИ УПОМИНАНИЕ:

народонаселение

ИЛИ ЕГО СИНОНИМЫ:


- население
 population

ПОИСК ОГРАНИЧИТЬ
ДОКУМЕНТАМИ,
которые

содержат упоминание планета земля и его синонимы

игнорируют упоминание человечество и его синонимы

Найти в Интернет

 <http://yandex.ru/yandsearch?text=%28%D0%F0%E4%EE%ED%F1%E5%EB%E5%ED%E8>

Поиск [Почта](#) [Новости](#) [Маркет](#) [Карты](#) [Словари](#) [Блоги](#) [Картинки](#) ещё ▾

Яндекс

Нашёлся
1 млн страниц

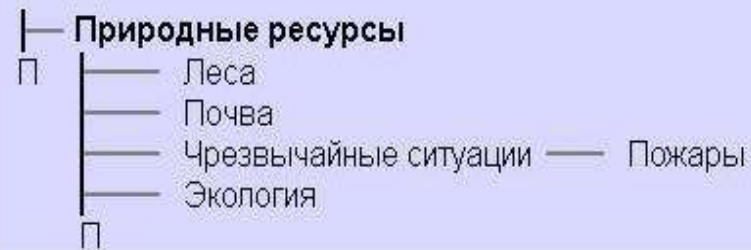
(народонаселение | население) && ((планета земля) | earth)

Найти

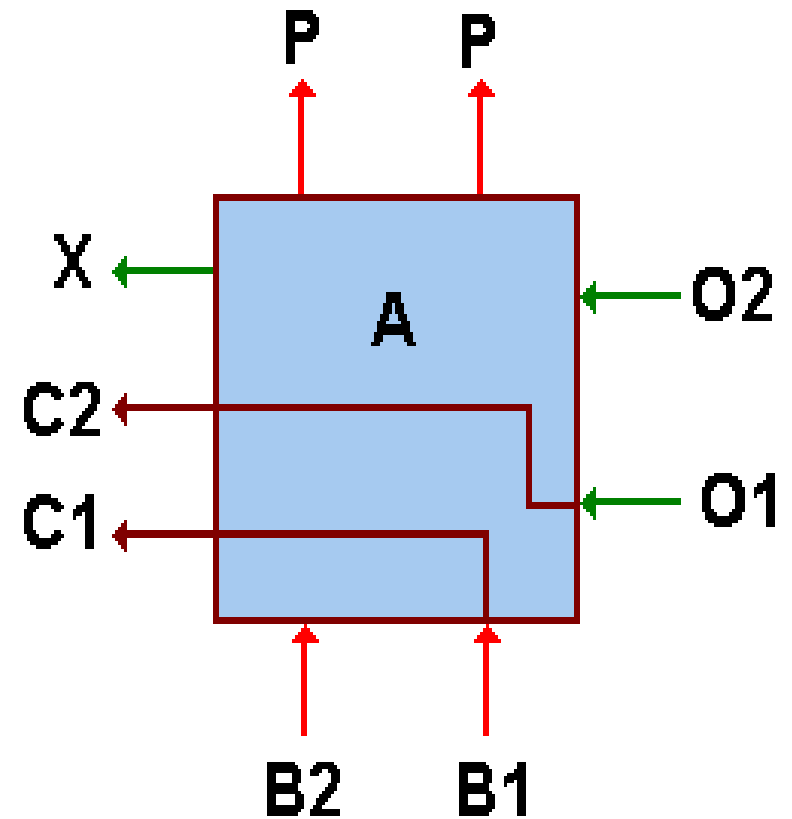
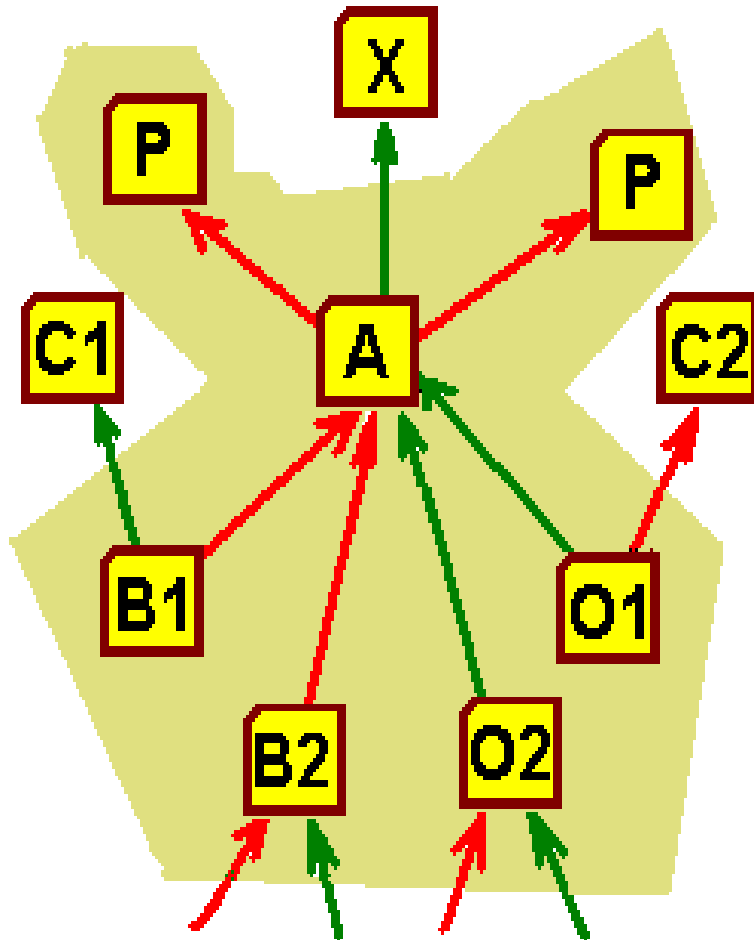
в найденном в Москве

[расширенный поиск](#)

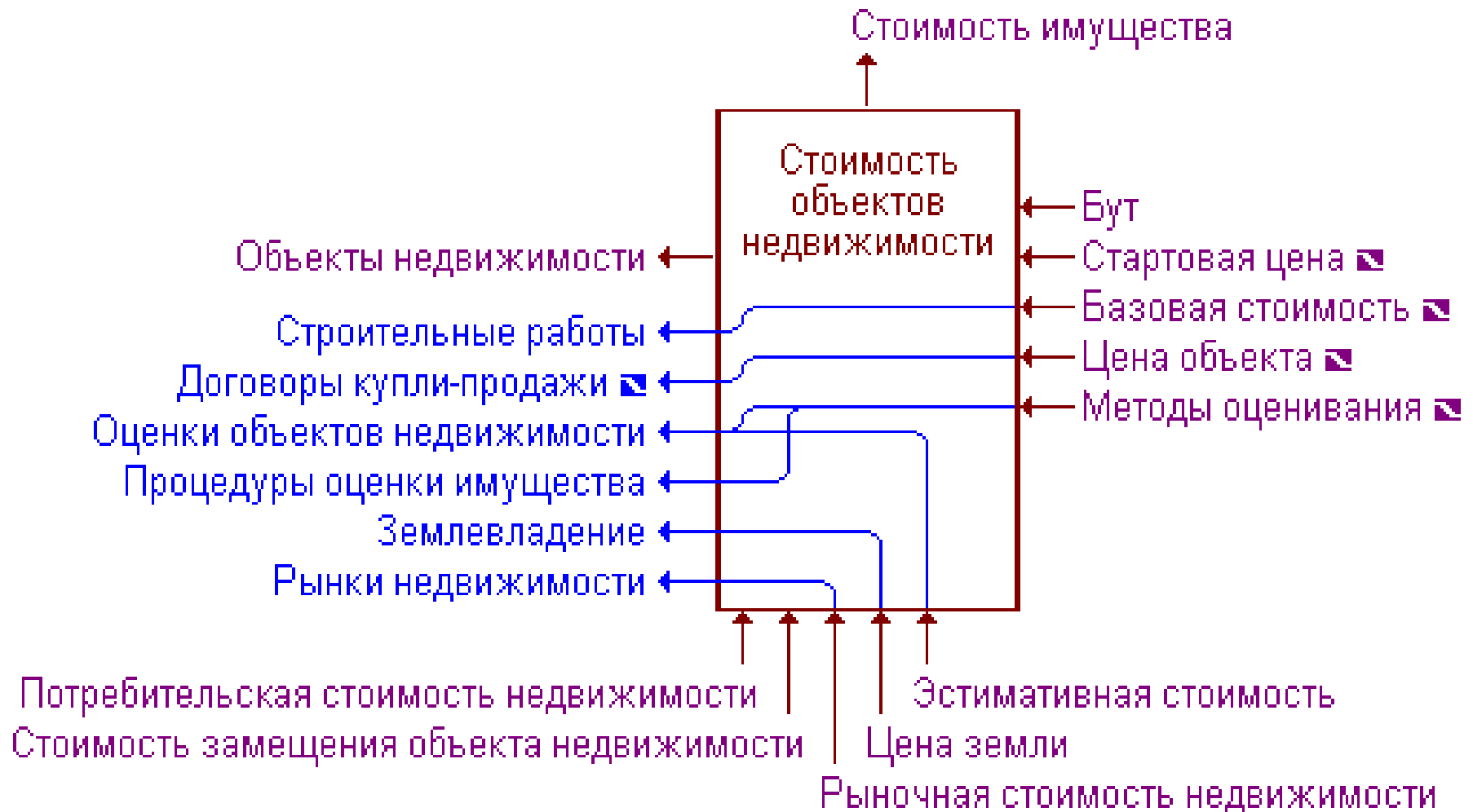
3.4 Веб-карта



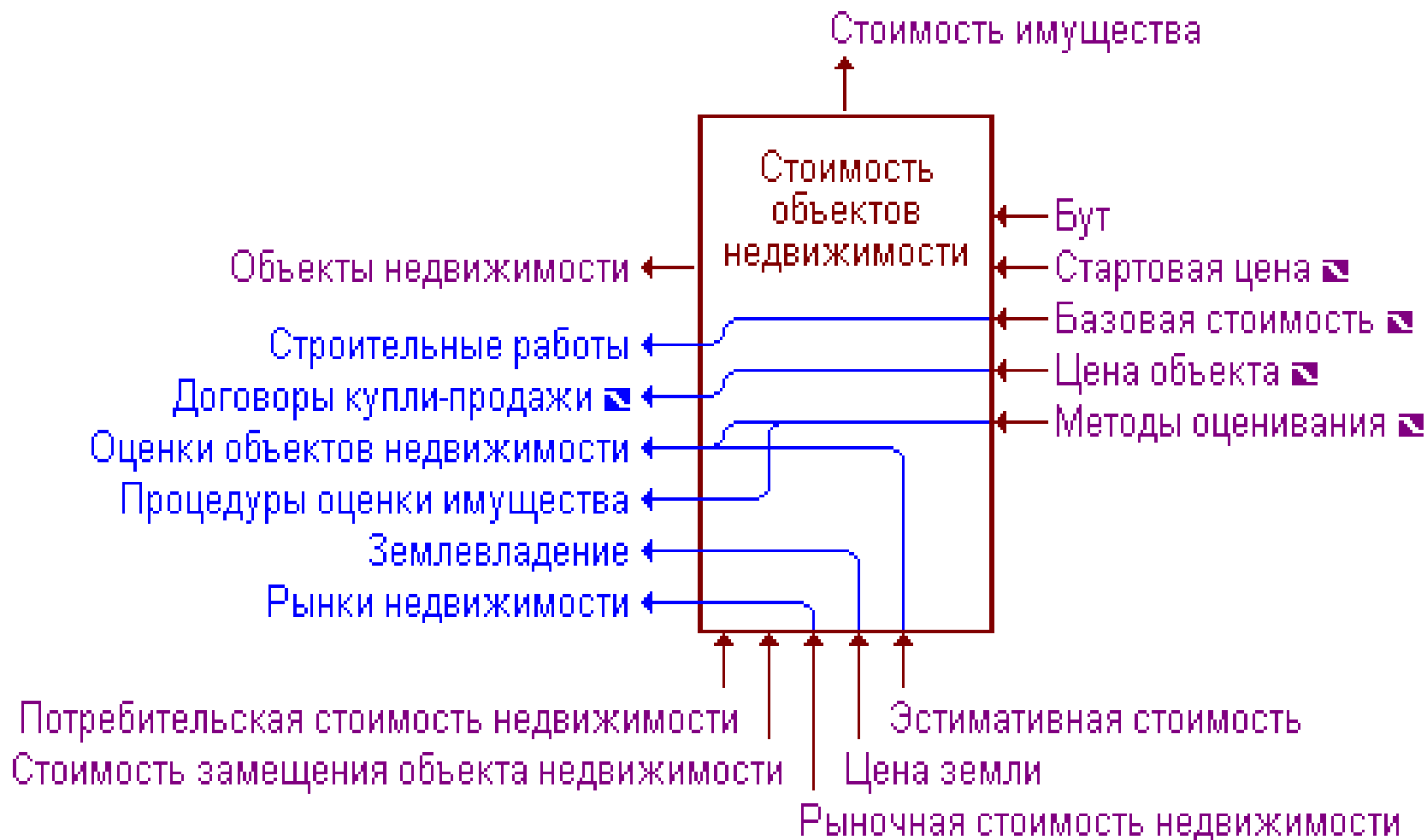
3.5 Принцип визуализации



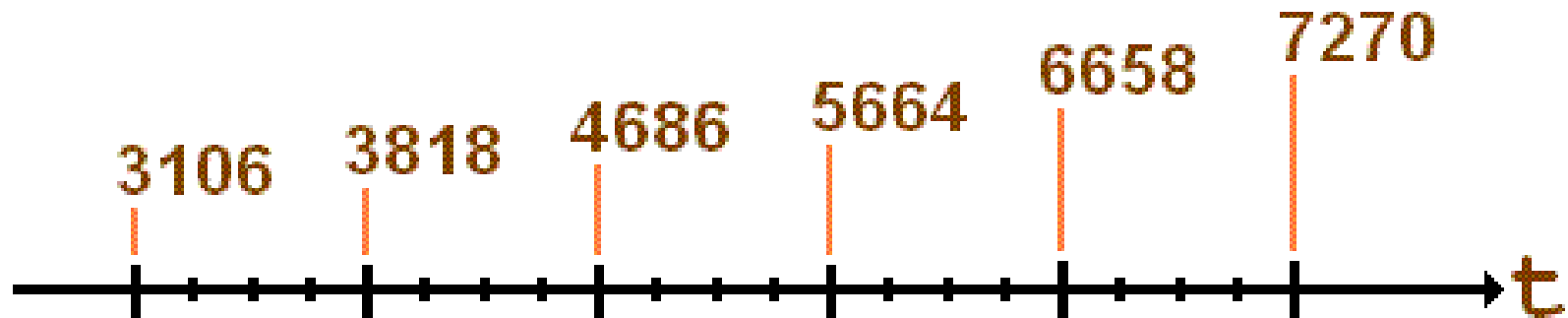
Результат визуализации



Формула глоссария $1 + 5/2 + 5/2 + 1 + 6$

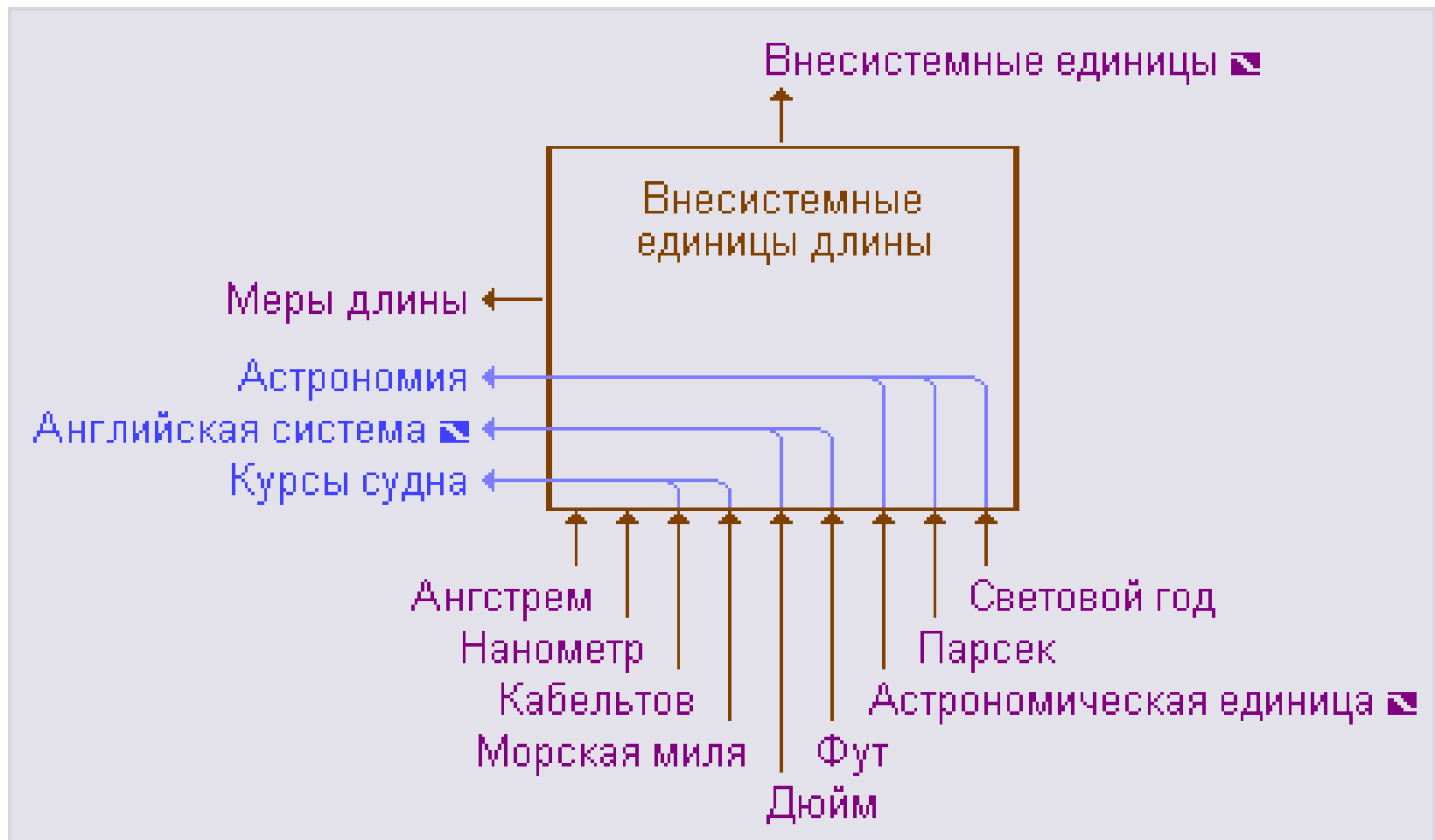


Статистические закономерности

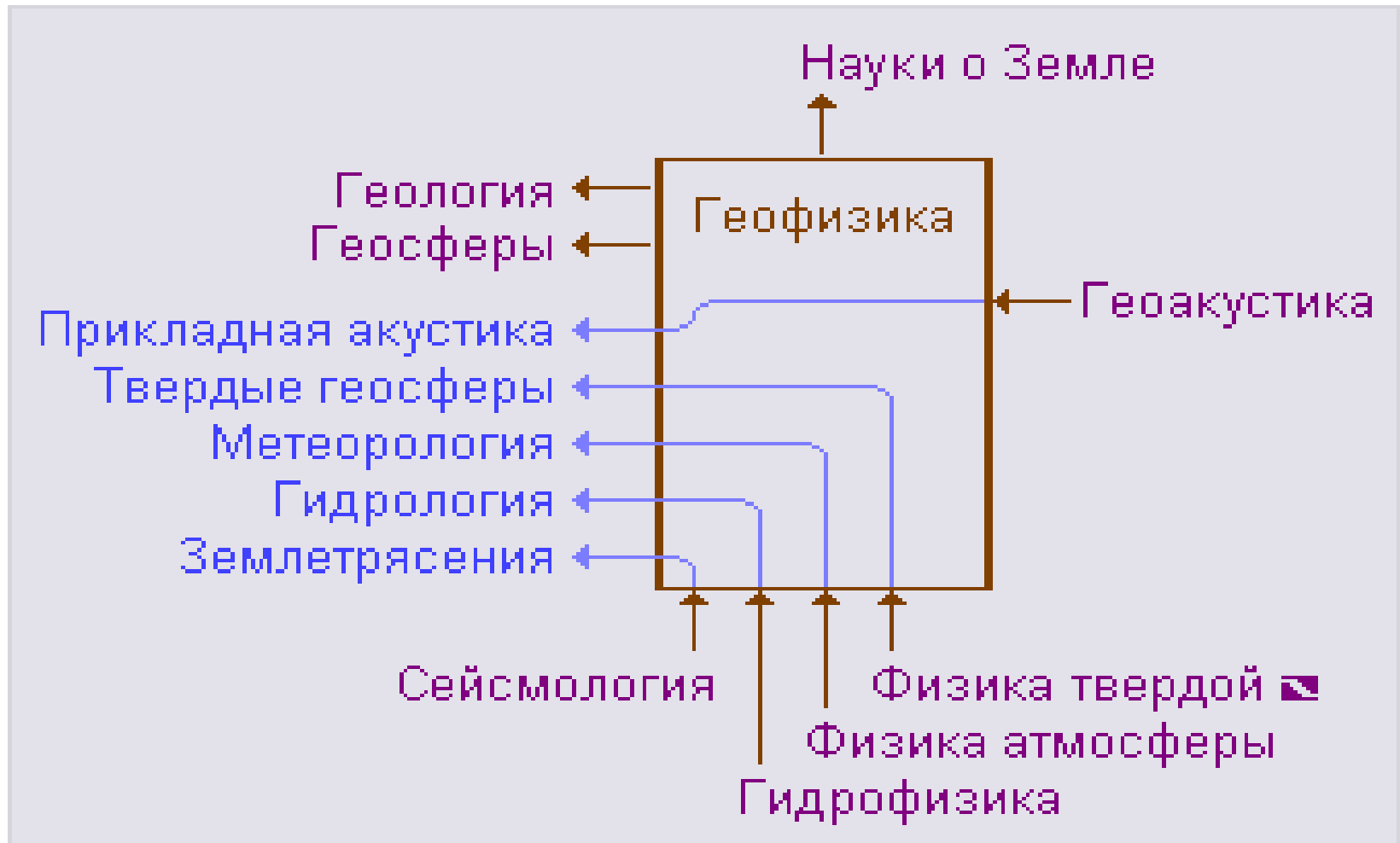


93% понятий содержат некоторое количество собственных подвидов или собственных аксессуаров.

93%



7%



Нерешенные задачи

1. **Эволюция и старение терминологии**
2. **Коллекции терминов**
3. **Кластеризация понятий и навигация**
4. Словарное представление фрагмента сети
5. Порядок предъявления статей
6. Алгоритм формирования запросов к ИПС (часть 2)
7. Автоматизация процессов формирования сети
8. Применение сети в ИПС

Презентация предоставлена
автором для размещения на
сайте www.commonmind.ru.

Вопросы?

С.Ю.Соловьев

soloviev@jscc.ru