

Расширение запроса при поиске-2

Маннинг и др. Введение в
информационный поиск, гл.9

Методы расширения запроса

- Несовпадение слова запроса:
 - самолет – лайнер
- Методы расширения запроса:
 - Глобальные методы
 - Ручные тезаурусы
 - Автоматически порождаемый тезаурус
 - Локальные методы
 - Relevance feedback (обратная связь по релевантности)
 - Pseudo Relevance feedback (обратная связь по псевдорелевантности)

Представление знаний: ручные тезаурусы

- Тезаурус – словарь, в котором
 - Единицы соответствуют близким по смыслу словам (понятия, дескрипторы и др.)
 - Между единицами установлены формализованные отношения
- Типы тезаурусов
 - Информационно-поисковые тезаурусы
 - Тезаурусы типа WordNet
 - Тезаурус Роже
 - Ассоциативные тезаурусы
 - Поиск с использованием доп. знаний - семантический информационный поиск

Тезаурус

- Разные виды
- Словарь с формализованными связями между словами

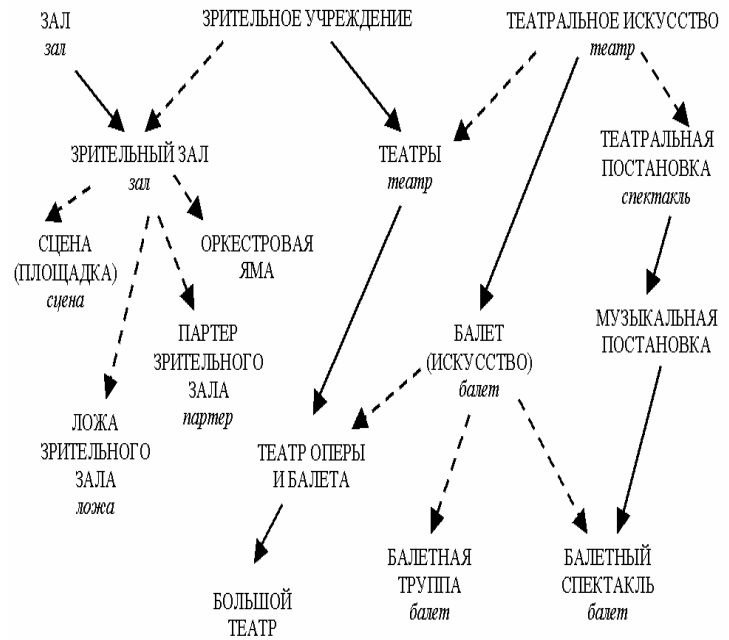
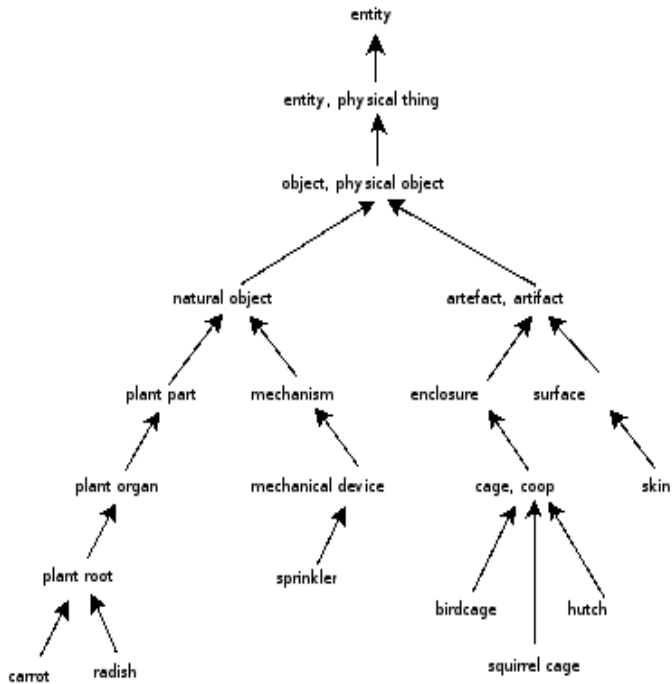


Figure 1. "is a" relation example

Исторические этапы в расширении запросов

- Булевский поиск: короткие тексты (1965)
 - Информационно-поисковые тезаурусы, ручное индексирование
- Векторные модели (1991-1992)
 - Локальное расширение запросов: relevance feedback
- Публикация тезауруса WordNet (1995)
 - Комбинирование с векторными моделями
- Автоматически порождаемые тезаурусы (2010)
 - Google, Yandex
- Рост значимости поведенческих факторов (2015)
- Использование тезаурусов в предметно-ориентированных системах

Информационно-поисковые тезаурусы

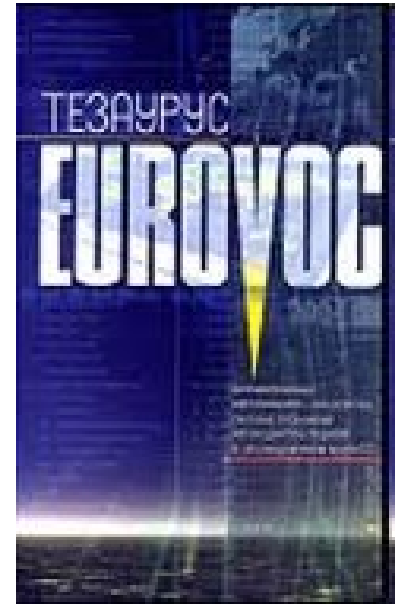
- Информационно-поисковый тезаурус – нормативный словарь терминов предметной области, создаваемый для улучшения качества информационного поиска в данной предметной области
 - Нормативные ключевые слова
 - Национальные и международные стандарты
 - Используются в ряде международных организаций и парламентский организаций
 - Европейский парламент – EUROVOC
 - ООН – UNBIS Thesaurus

Цели и использование

- Цели:
 - Перевод языка авторов на нормативный язык, используемый для индексации и поиска
 - Обеспечение последовательности в присваивании индексных терминов
 - Обозначение отношений между терминами
 - Облегчение информационного поиска
- Использование
 - Особенно популярны в 60-80х годах 20 века
 - Повышение интереса в последнее время за счет роста интереса к предметно-ориентированному и корпоративному поиску

Примеры тезаурусов

- Тезаурус ООН – UNBIS Thesaurus
- Тезаурус Европейского союза – EuroVoc
- Тезаурус Исследовательской службы Конгресса США – LIV
- СССР
 - Правовой тезаурус
 - ИНИОН
 - Шемакин «Технический тезаурус»
- Стандарты ISO, ГОСТы



Традиционные информационно-поисковые тезаурусы для ручного индексирования: структура

- Основные понятия ПО – дескрипторы
- Условные синонимы – аскрипторы –
 - Отношения эквивалентности
 - аскриптор – дескриптор
- Отношения между дескрипторами

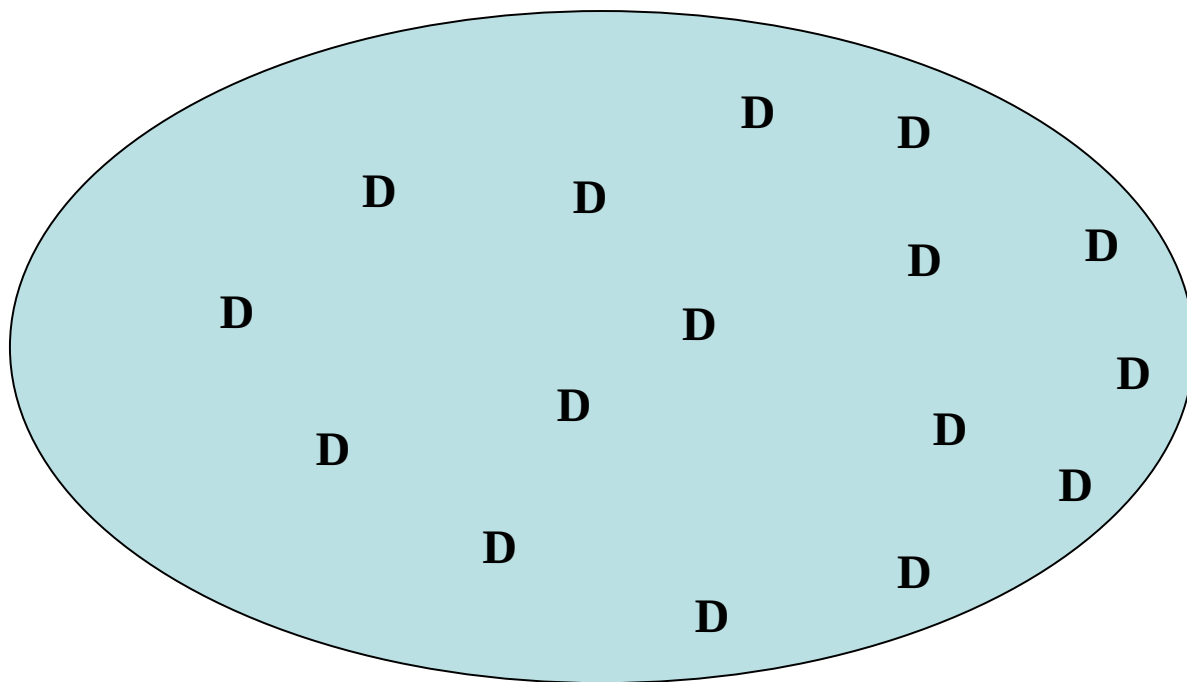
Отношения в ИПТ

- Родовидовые отношения (выше – ниже)
 - BT (broader term) – NT (narrower term)
- Отношение ассоциации
 - RT (related term)
 - Транспорт – путешествие
 - Неврология – нервная система
 - Переплетные работы - книга
- Часть –целое: части должны всегда относиться к своему целому (рекомендация стандарта Z39.19):
 - органы тела, географические объекты, дисциплины, иерархические структуры (*полк – батальон – рота*)

Информационно-поисковые тезаурусы: использование для индексирования

- Процедура индексирования по тезаурусу
 - Индексатор читает документ, формулирует его основное содержание
 - Затем ему нужно подобрать наиболее подходящий набор дескрипторов для описания содержания документа
 - Отношения используются для уточнения набора дескрипторов
- Контекст использования
 - 60-80е годы – в информационных системах не было полных документов
 - Вид библиотечной обработки в парламентах, международных организациях

Дескрипторы в предметной области



Дескрипторы похожи на теги (ключевые слова), но тщательно отобраны профессионалами для обеспечения равномерного покрытия предметной области

Тезаурус исследовательской службы Конгресса США (LIV)

Transportation

Scope Note *For writings on transportation in a specific city or metropolitan area, use
Urban transportation.*

Narrower Term [Choice of transportation](#)

[Commercial aviation](#)

[Energy transportation](#)

[Highspeed ground transportation](#)

[Inland water transportation](#)

[Intermodal transportation](#)

[Marine transportation](#)

[Military transportation](#)

[Motor transportation](#)

[Railroads](#)


[Student transportation](#)

[Terminals \(Transportation\)](#)

[Transportation and the aged](#)

[Transportation and the disabled](#)

Тезаурус EuroVoc

This site is part of 

[Useful links](#) | [About EuroVoc](#) | [Sitemap](#) | [Help](#) | [Contact](#) | [Legal notice](#)



EuroVoc

Multilingual Thesaurus of the European Union

Interface language: (en) English

[Europa](#) > [EuroVoc homepage](#) > [Domains and MT](#) > [rail transport](#)

Content language:

(en) English

Simple search

■ [Advanced search](#)

Browse

■ [Browse the subject-oriented version](#)

Download

- [By domain](#)
- [Permuted alphabetical](#)
- [Multilingual list](#)
- [Alphabetical index](#)
- [EuroVoc SKOS/RDF](#)
- [EuroVoc XML](#)

Your proposals

- [Contribute](#)
- [New approved concepts](#)

rail transport

UF *rail connection*
rail traffic
railway
transport by railway

48 TRANSPORT

MT 4816 land transport
BT1 land transport
NT1 CIV Convention
NT1 rail network
NT2 railway station
NT1 rolling stock
NT1 vehicle on rails

RT *air-cushion vehicle* [4811]
European Railway Agency [1006]
high-speed transport [4811]
railway industry [6821]
railway tariff [4806]
transport staff [4811]

URI <http://eurovoc.europa.eu/4514>

Has Exact Match

Rail transport (AGROVOC)
Rail transport (ECLAS)
rail transport (GEMET)
railway transport (UNBIS)

Has Related Match

Rail traffic (ECLAS)
railway (GEMET)
railway traffic (UNBIS)
railways (UNBIS)

Has Close Match

Railways (ECLAS)

LANGUAGE EQUIVALENTS

BG	железопътен транспорт
ES	transporte ferroviario
CS	železniční doprava
DA	jernbanetransport
DE	Schienentransport
ET	raudteetransport
EL	σιδηροδρομικές μεταφορές
EN	rail transport
FR	transport ferroviaire
HR	željeznički prijevoz
IT	trasporto ferroviario
LV	dzelzceļa transports
LT	geležinkelių transportas
HU	vasúti közlekedés
MT	trasport ferrovjarju
NL	vervoer per spoor
PL	transport kolejowy
PT	transporte ferroviário
RO	transport feroviar
SK	železničná doprava
SL	železniški promet
FI	rautatiekuljetus
SV	järnvägstransport
SR	железнички превоз
MK	железнички сообраќај
SQ	transport hekurudhor

Традиционные ИПТ: автоматическое расширение запроса

- Объем информационно-поисковых тезаурусов относительно небольшой
 - Многие слова отсутствуют
- Если использовать описанные отношения для автоматического расширения запроса,
 - то полнота поиска увеличивается (находим больше релевантных документов),
 - но точность поиска резко падает (появляется много нерелевантных документов).

В среднем качество поиска падает.

- объем относительно небольшой, нет большого количество нужных слов
- расширение запроса по отношениям иногда сильно вредит

Тезаурус WordNet

- Реляционное описание лексики английского языка
 - Иерархическая сеть понятий (synset)
 - Каждое слово относится к одному или нескольким понятиям
 - Отдельная иерархическая сеть для различных частей речи – психолингвистическое обоснование
- Автор: George Miller
(50-е годы статья «Магическое число 7»)
- Версия 1.6:
 - 95 тысяч понятий, около 130 тысяч слов
- Версия 3.0:
 - 117 тыс. синсетов, около 155 тысяч разных слов и словосочетаний, 200 тысяч отношений слово-синсет

Синсеты

- Синсет (набор синонимов) – единица словаря => понятие
- Синонимы имеют одинаковое толкование
- Семантические отношения между синсетами
 - Гипонимы, гиперонимы, часть, целое и др.
- Многозначные слова: слово входит в несколько синсетов

Единицы WordNet - синсеты

Word to search for:

Display Options:

Key: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations

Display options for sense: (gloss) "an example sentence"

Noun

- **S: (n) cat, true cat** (feline mammal usually having thick soft fur and no ability to roar: domestic cats; wildcats)
 - [direct hyponym](#) / [full hyponym](#)
 - [direct hypernym](#) / [inherited hypernym](#) / [sister term](#)
- **S: (n) guy, cat, hombre, bozo, sod** (an informal term for a youth or man) *"a nice guy"; "the guy's only doing it for some doll"; "the poor sod couldn't even buy a drink"*
- **S: (n) cat** (a spiteful woman gossip) *"what a cat she is!"*
- **S: (n) kat, khat, qat, quat, cat, Arabian tea, African tea** (the leaves of the shrub *Catha edulis* which are chewed like tobacco or used to make tea; has the effect of a euphoric stimulant) *"in Yemen kat is used daily by 85% of adults"*
- **S: (n) cat-o'-nine-tails, cat** (a whip with nine knotted cords) *"British sailors feared the cat"*
- **S: (n) Caterpillar, cat** (a large tracked vehicle that is propelled by two endless metal belts: frequently used for moving earth in construction and farm work)

Отношения гипоним-гипероним: Лексическая иерархия

- Отношение гипоним-гипероним
 - Грач – перелетная птица семейства вороновых
 - Птица – покрытое пухом и перьями позвоночное животное
 - Животное – всякое живое существо, исключая растения
 - Грач – птица – животное – живое существо
- Свойства отношения
 - Транзитивность
 - Наследование свойств от гиперонимов к гипонимам

WordNet Search - 3.1

- [WordNet home page](#) - [Glossary](#) - [Help](#)

Word to search for:

Display Options:

Key: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations

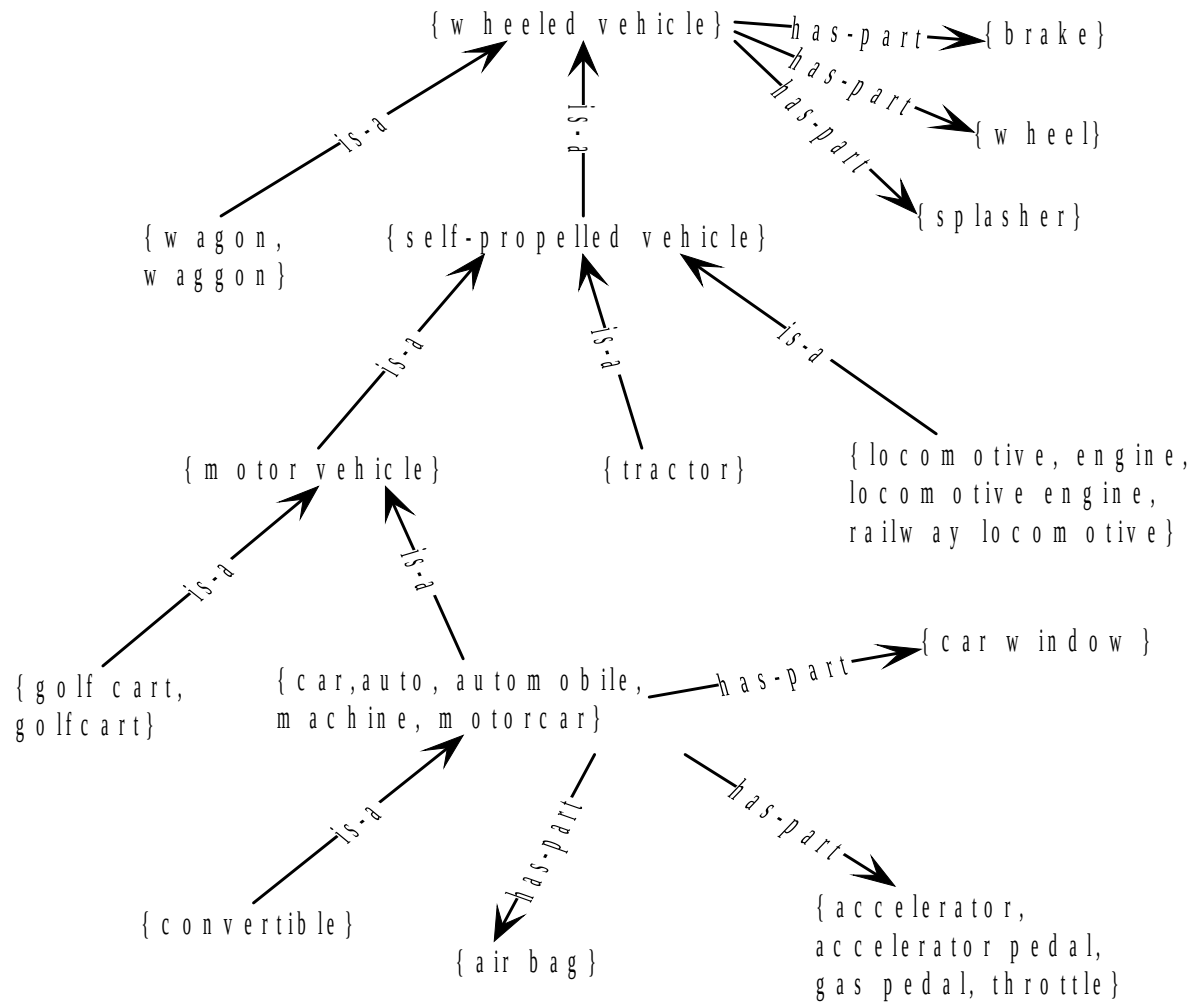
Display options for sense: (gloss) "an example sentence"

Синсеты и значения

Noun

- [S:](#) (n) **nature** (the essential qualities or characteristics by which something is recognized) *"it is the nature of fire to burn"; "the true nature of jealousy"*
- [S:](#) (n) **nature** (a causal agent creating and controlling things in the universe) *"the laws of nature"; "nature has seen to it that men are stronger than women"*
- [S:](#) (n) **nature** (the natural physical world including plants and animals and landscapes etc.) *"they tried to preserve nature as they found it"*
- [S:](#) (n) **nature** (the complex of emotional and intellectual attributes that determine person's characteristic actions and reactions) *"it is his nature to help others"*
- [S:](#) (n) **nature** (a particular type of thing) *"problems of this type are very difficult to solve"; "he's interested in trains and things of that nature"; "matters of a personal nature"*

WordNet (Miller et al. 1990; Fellbaum, 1998)





EuroVoc

Multilingual Thesaurus of the European Union

Int

EuroVoc > EuroVoc homepage > Search > nature

Content language:

(en) English ▼

Simple search

nature

■ Advanced search

Browse

■ Browse the subject-oriented version

Download

- By domain
- Permuted alphabetical
- Multilingual list
- Alphabetical index
- EuroVoc SKOS/RDF
- EuroVoc XML

Your proposals

- Contribute
- New approved concepts

Search results (3 concepts)

MT 2436 public finance and budget policy

budget policy

budget policy

UF

yearly **nature** of the budget

MT 5206 environmental policy

environmental protection

environmental protection

UF

conservation of **nature**

environmental protection

UF

nature protection

environmental protection

UF

protection of **nature**

nature reserve

UF

nature park

nature reserve

UF

nature preserve

Using WordNet for Text Retrieval

- Ellen M. Voorhees (SIGIR - 1995)
- Разрешение многозначности в запросе и текстах
- Hood – область в сети WordNet, где слово однозначно – наибольший подграф, включающий S (синсет для одного из значений), и не включающий другие синсеты с S

Расширение запроса

1. Расширение запроса с ручным разрешением многозначности
2. Расширение запроса с автоматическим разрешением многозначности

50 Запросов из конференции TREC

TREC: пример запроса

Domain: Science and Technology

Topic: Aftermath of Chernobyl

Description: document cites measures taken by European countries to record and contain the results of the Chernobyl accident.

Narrative: a document will cite an action undertaken by a European government such as testing food supplies, testing water, measuring fallout, banning new reactor construction ...

Concepts:

1. Chernobyl, nuclear accident, radiation, contamination
2. Consequences, evacuation, health fears, cancer
3. Banning foodstuffs, propaganda campaign, testing soil...

Автоматическое расширение запроса

- Запросы TREC: description
 - Слишком частотные слова в коллекции не расширяются ($\text{freq} < N$)
 - Для остальных берутся все понятия-соседи (шаг 1, 2)
 - Если есть пересечение синсетов от разных слов запроса, то элементы синсета добавляются к запросу
- Оценивались:
 - Различные N – 5%, 10% коллекции
 - Различные веса на расширение 0.3, 0.5, 0.8.
 - Максимальное улучшение 0.7%
 - Этот результат не является статистически значимым

Основная проблема при разрешении многозначности в контексте информационного поиска

- Выводы в работе Vorhees (1995)
 - В запросе происходит такой выбор значения, которого нет нигде в текстах.
 - Запрос: separation anxiety in infants and preschool children
 - 8 значений слова separation
 - Всего 8 релевантных документов, но находится только 1.
- The inability to automatically resolve word senses prevented any improvements from being realized.

Расширение запроса, основанное на тезаурусных знаниях

- Для каждого термина t в запросе происходит расширение синонимичными словами или близкими по смыслу (связанными отношениями с исходным словом)
 - из тезауруса
 - *feline* → *feline cat*
- Как расширять:
 - Можно добавлять в вектор запроса (с более низкими весами и в зависимости от типа отношения к слову запроса)
 - Можно вставлять в булевское выражение
 - *Налог* → (*НАЛОГ* или *НАЛОГОВЫЙ*)
- Используется в предметно-ориентированных системах
 - Современные тезаурусы, встроенные в ПО поисковые системы, могут иметь другие формы, чем описано в стандартах, например, только список синонимов и вариантов

Расширение запроса, основанное на тезаурусных знаниях-2

- Увеличивает полноту поиска
- Обычно снижает точность поиска, обычно для многозначных слов
 - “interest rate” → “interest rate fascinate evaluate”
 - Можно вводить в тезаурус многословные термины «interest rate», но запросы все равно разнообразнее
- Сложность создания и обновления тезаурусов
- Поэтому в интернет-поиске
 - Долгое время не было расширения запросов
 - Затем стали расширять на однокоренные слова
 - Сейчас для расширения запроса используются статистически насчитанные «синонимы»

Тезаурусные отношения при автоматич. расширении запросов

- Синонимы
 - хорошо работает для однозначных слов (выражений)
- Родовидовые отношения (выше-ниже)
 - Хорошо работает, если запрос совпадает с термином тезауруса
 - В длинном запросе может приводить к снижению точности
 - Города Сибири -> город столица Сибири

Методы расширения запроса

- Несовпадение слова запроса:
 - самолет – лайнер
- Методы расширения запроса:
 - Глобальные методы
 - Информационно-поисковый тезаурус
 - Автоматически порождаемый тезаурус
 - Локальные методы
 - Relevance feedback (обратная связь по релевантности)
 - Pseudo Relevance feedback (обратная связь по псевдорелевантности)

Слайды доклада Расширение поисковых запросов

А. Сокирко

Е. Соловьев (Яндекс)

[http://romip.ru/russir2010/slides/
yandex_lecture.pdf](http://romip.ru/russir2010/slides/yandex_lecture.pdf)

Типы синонимов для расширения запроса

- С соответствиями между внутренними элементами:
 - Словообразование: **Москва** – **москов**ский; компиляция - компилирование
 - Аббревиатуры: **МГУ** - **Московский** **г**осударственный **у**ниверситет
 - Транслиты: **Гугл** – **Google**
 - Слитно – раздельно: **ватер-поло** – **ватерполо**
 - Орфоварианты: **colour** - **color**
- Без поддержки внутренних элементов
 - Переводы (стол – table)
 - Синонимы: бегемот – гипопотам
 - Подвиды: фильм - биопик

Алгоритм составления базы

- Получение списка гипотез



~ 200 миллионов гипотез

- Машинное обучение



~ 150 миллионов гипотез

- Отсечение результатов – отсекается первый миллион и объявляется словарем

Учет словообразования

- Лингвистические модели
 - Memorize->memorization
- Много ошибочных образований: sense – sensibility
- Обобщенные модели:
 - Memorize->memorization (e->ation)
 - Induce->induction (e->tion)
 - Publish -> publication (sh->cation)

Признаки синонимов: Скобочное написание

Скобочное написание – это набор n-gram, которые встречаются в текстах рунета в контексте скобок:

Московский государственный
университет (МГУ)

Владимир Путин (Vladimir Putin)

Признаки синонимов: Открытые словари

Русская Википедия содержит около миллиона жестких перенаправлений, типа:

Абрикос сибирский	---	Даурсат
Авачинская бухта губа	---	Авачинская

Извлечение синонимов для автоматического расширения запросов

- Признаки для извлечения синонимов
 - Совместная встречаемость в одном документе (странице)
 - Совместная встречаемость в тексте ссылки (анкор)
 - Встречаемость в документе и в тексте ссылки
 - Как часто пользователь в запросах заменяет одно на другое
 - Клики пользователя на страницу, содержащую S2, при запросе, содержащем S1
 - сходство контекстов употребления S1 и S2 (запросы, документы) и др.

Оценка качества

- 1) Оценка пары синонимов без контекста
- **2) Оценка пары синонимов в контексте запроса**
- 3) Оценка качества поисковой выдачи

Меры качества

- F-measure for the dictionary is around 87% (Metric 1).
- F-measure for query expanding by derivation pairs is 65% (Metric 2).

[Australian population] (*Australian*=>*Australia* +)

[Australian gold] (*Australian* => *Australia* -)

[milk diet] (diet => dietary +)

[The Diet of the German Empire] (diet=>dietary -)
(a kind of Parliament)

Лексические расширения с использованием автоматического тезауруса

- Те же проблемы, что и в ручном:
 - Многозначность запроса или расширения
 - Проблемы с расширением устойчивых словосочетаний
 - Влияние контекста запроса и/или документа

Примеры расширения запросов в Яндексе (декабрь 2010 – февраль 2011)

Запрос — документ

- | | | |
|------------------------|---|--------------------------------------|
| речное судно | — | морское судно |
| речной порт | — | морской порт (Находка) |
| присуждение имущества | — | передача имущества |
| ледяная горка | — | холодная гора |
| расширение отверстия | — | расширение канала
(сети интернет) |
| договор поручительства | — | договор поручения |
| аварийное отключение | — | знак аварийной остановки |
| замкнутая граница | — | закрыть границу |

Яндекс

Нашлось
504 тыс.
ответов

Поиск Почта Карты Маркет Новости Словари Блоги Видео Картинки ещё ▼

речное-судно

Найти

☐ в найденном ☐ в Троицке

расширенный поиск

Мои находки


Настройка

Регион: Троицк


Войти


Помощь


2011 год

- 1  [Как устроены морские суда - Заглавная страница](#)
Надстройки и Рубки. Штевни морского судна. Кронштейны, выкружки и коридор гребного вала. Люки и шахты на морском судне. Фундаменты судовых машин. Фальшборт и леерное ограждение.
www.seaships.ru [копия](#) [ещё](#)

- 2 [Российские речные суда](#)
© "Российские речные суда", 1998-2005. © "Российский речной флот и туризм INFOFLOT.RU"
www.riverships.ru > [Российские речные суда](#) [копия](#) [ещё](#)

- 3  [приобретение в собственность морского/речного судна](#)
Annuitet Пишет: - дык у меня аренда, а не ипотека, и речное судно. > (плавучая гостиница), а не морское - поэтому вроде.
forum.garant.ru/read.php?7,1262078 [копия](#) [ещё](#)

- 4  [оценка морского судна, оценка корабля, оценка речных судов, оценка...](#)
Все речные и морские суда, которые необходимо, согласно существующему законодательству регистрировать, причисляются к недвижимому имуществу.
www.proflexpert.ru/ocenka-transporta/morskoj-i-... [копия](#)

- 5  [РЕЧНОЕ СУДНО-АВТОМОБИЛЕВОЗ. - СУДА - ДОСКА ОБЪЯВЛЕНИЙ](#)
Разместить объявление. Речное судно-автомобилевоз, Судно катамаранного типа для перевозки автомобилей, проект Р-19. для перевозки авто, негабаритов, яхт.
www.korabel.ru > [Доска объявлений](#) > [detail/34122.html](#) [копия](#) [ещё](#)

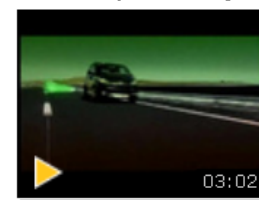
[Разместить объявление по запросу «речное-судно»](#)

[«речное-судно» в картинках](#)



[Все картинки](#)

[Видео «речное-судно»](#)



Восстановление Шелкового пути в современном формате
[Все видеоролики](#)

расширение-отверстия — Яндекс: Нашлось 207 тыс. ответов - Mozilla Firefox

Файл Правка Вид Журнал Закладки Инструменты Справка

http://yandex.ru/yandsearch?text=расширение-отверстия&lr=20674

расширение-отверстия — Янд...

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

Яндекс

Нашлось 207 тыс. ответов

расширение-отверстия

☐ в найденном ☐ в Троицке

Найти

расширенный поиск

[Мои находки](#)

[Настройка](#)


[Регион: Троицк](#)

2011г.

- [Новости сети](#)
Сегодня произведено плановое **расширение канала** связи Интернет. Дата размещения: 27.03.2009.
www.oz-web.ru/news/index.php?news=34 [Орехово-Зуево](#) [копия](#) [ещё](#)
- [Роль предварительного расширения каналов - Медицинская библиотека...](#)
Определение апикального сужения в изогнутых корнях нижнечелюстных моляров - предварительно **расширенные каналы** в сравнении с нерасширенными.
www.medlinks.ru/article.php?sid=27150 [копия](#) [ещё](#)
- [Largal ultra – жидкость для химического расширения каналов. Septodont...](#)
С помощью пипетки ввести ЛАРГАЛЬ УЛЬТРА в полость зуба, а затем в **каналы**, но уже используя для этого корневую иглу. Сразу после этого можно начать механическое **расширение канала**.
www.konsort.ru/index.php?productID=934 [Москва](#) [копия](#) [ещё](#)
- [ОАО ГОТТЦ Гарант - Расширение канала Интернет Гродно-Минск](#)
Для удобства пользователей в ближайшее время канал Гродно-Минск для выхода в сеть Интернет будет **расширен** дополнительно на 6 Мегабит/с...
www.garant.by/internet/news/items/rasshirenje... [копия](#) [ещё](#)
- [Новости](#)
Расширение канала по д. Волжский. Указанные абоненты на прошлой неделе был


[Разместить объявление по запросу «расширение-отверстия»](#)

«расширение-отверстия» в картинках



[Все картинки](#)

Видео «расширение-отверстия»



04:01

БИА микрохирургическое **расширение канала**

[Все видеоролики](#)

Готово

работа-на-большой-высоте — Яндекс: Нашлось 30 тыс. ответов - Mozilla Firefox

Файл Правка Вид Журнал Закладки Инструменты Справка

http://yandex.ru/yandsearch?text=работа-на-большой-высоте&lr=20674

Яндекс

Поиск Почта Карты Маркет Новости Словари Блоги Видео Картинки ещё ▼

Яндекс
Нашлось
30 тыс. ответов





работа-на-большой-высоте

☐ в найденном ☐ в Троицке

Найти

расширенный поиск

Мои находки
Настройка
Регион: Троицк

-  [Страховка при работе на большой высоте / Энергосфера, Пермь / Статьи](#)
Профессия монтажников связана с повышенной опасностью, т.к. **большая** часть их работы проходит на **большой высоте**. ... Но работа на большой высоте сопряжена с огромным риском. Поэтому, для того, чтобы монтажник чувствовал себя защищенным, его необходимо...
[energosever.ru > stati_strahovka...na_visote.html](#) [копия](#) [ещё](#)
-  [Работа на больших высотах., цена, купить в Полтаве — Prom.ua](#)
Подробная информация о товаре и поставщике с возможностью онлайн-заказа. ... **Работа на больших высотах.**, Полтава. Поставщик: ЧП Конюшенко. **Работа на большой высоте.**
[Poltava.prom.ua > p374540-rabota-na-bolshih...](#) [копия](#) [ещё](#)
-  [НОВЫЕ ТЕХНИЧЕСКИЕ СРЕДСТВА ДЛЯ РАБОТЫ НА БОЛЬШИХ...](#)
Таким образом, современные малогабаритные АОПА становятся эффективным средством выполнения широкого спектра **работ на больших глубинах**, являясь серьезной альтернативой применяемым в настоящее время телеуправляемым аппаратам рабочего и ...
[korabel.ru > news/comments/novie...na...glubinah...](#) [копия](#) [ещё](#)
-  [Кого привлекают сегодня для выполнения работ на большой высоте?](#)
Для того, чтобы установить кондиционер **на большой высоте**, пользуются услугами

Яндекс.Директ

[Зарплата 439\\$ в неделю.](#)
Работа дома, через интернет. Свободны график. Вот это настоящая **работа!**
[psyforex.ru](#)

[Требуются работники](#)
Вахтовая работа: требуются специалисты люди без опыта! Зп 95000 руб
[www.vahtajob.ru](#)

[Работа: вакансии и резюме!](#)
Поиск **работы** на нашем сайте! **Большая** база вакансий! Обновление каждый день
[bestrobotazwc.moy.su](#)

[Высокооплачиваемая работа!](#)
Актуальные вакансии от серьезных работодателей! Очень **большая** база!
[elitejobbai.do.am](#)

[Новая работа! Свежие вакансии!](#)
Найди **работу** уже сегодня! Быстро! Легко! Здесь серьезные работодатели!
[elitevacansyebi.do.am](#)

[Все объявления](#)

[Разместить объявление по запросу](#)

Готово

2015 год

Поиск

Картинки

Видео

Карты

Маркет

Ещё

**Расширение - отверстие** - Большая Энциклопедия Нефти...[ngpedia.ru > id383768p1.html](#) ▼**Расширение отверстия** определяется по среднеобъемной температуре нагрева детали.Размест
«расши**Зенкеры цилиндрические для расширения отверстий**[info.instrumentmr.ru > instrum_otverst5.shtml](#) ▼Цилиндрические зенкеры для **расширения отверстий** наиболее широко распространены в промышленности.**Как расширить отверстие** в стеклянных бусинах. МК...[filosofyfree.ru > post281904285/](#) ▼берем нужную бусину где нужно **расширить отверстие** и начинаем процесс **расширения**. ВАЖНО! смачивать сверло, что бы оно не нагревалось...**Устройство для расширения отверстий** | Банк патентов[bankpatentov.ru > node/172640](#) ▼1 изображен общий вид устройства для **расширения отверстий**; на фиг. 2 - устройство 7 в **отверстии** негабарита**Расширение межпозвоночного отверстия**[VIM-clinic.ru > services/Back-surgery-...](#) ▼**Расширение межпозвоночного отверстия** (Foraminotomy). Операции по **расширению** межпозвоночного **отверстия**, или фораминотомия — операция, при которой...**Сверление отверстий** в щитах [Архив] - Форум "Город...[mastercity.ru > Archive > 404040.html](#)

Наша специализация

- Лингвистические ресурсы в информационно-аналитических системах
 - Тезаурус РуТез
 - Общественно-политический тезаурус
 - Тезаурус по безопасности

Тезаурус по безопасности: понятия и их текстовые входы

Отношения на концептах

ЕÑЕÀÈÑÒ

Название концепта
ИСЛАМИЗАЦИЯ
ИСЛАМИЗМ
▶ ИСЛАМИСТ
ИСЛАМИСТСКАЯ ОРГАНИЗАЦИЯ
ИСЛАМИСТСКИЙ ОТРЯД
ИСЛАМОВЕД
ИСЛАМОВЕДЕНИЕ

Добавить

Изменить

Удалить

1 + -

ISLAMITE RADICAL

Фильтр

Текстовый вход
▶ ИСЛАМИСТ
ИСЛАМИСТКА
ИСЛАМИСТСКИЙ
ИСЛАМИСТСКИЙ ЭЛЕМЕНТ
ИСЛАМСКИЙ РАДИКАЛ
ИСЛАМСКИЙ ЭКСТРЕМИСТ

Перейти к синонимам

Фрагменты текстов

Отношение	Название концепта
ЦЕЛОЕ	ИСЛАМИЗАЦИЯ
НИЖЕ	БОКОХАРАМОВЕЦ
НИЖЕ	ВАХХАБИТ
НИЖЕ	ИСЛАМСКИЙ ТЕРРОРИСТ
НИЖЕ	ТАЛИБЫ (религиозное дви
НИЖЕ	ИСЛАМОФАШИСТ
НИЖЕ	МОДЖАХЕД

Добавить

Изменить

Перейти

Удалить

1 + -

ISLAMIZATION

Текстовый вход

Текстовый вход
▶ ИСЛАМИЗАЦИЯ
ИСЛАМИЗИРОВАТЬ
ИСЛАМИЗМ
МУСЛИМИЗАЦИЯ
ПРОИСЛАМСКИЙ
ПРОМУСУЛЬМАНСКИЙ

Добавить

Изменить

Удалить

Заккрыть

Тезаурус по безопасности: текстовые входы и отношения

Отношения на концептах

ОБЩАЯ УЯЗВИМОСТЬ

Название концепта
АТАКА НА ПЛАТЕЖНУЮ СИСТЕМУ
АТАКА НА ПОИСКОВУЮ СИСТЕМУ
АТАКА НА ПОСОЛЬСТВО
▶ АТАКА НА УЯЗВИМОСТЬ
АТАКА НА ФАЙЛООБМЕННУЮ СЕ
АТАКА НУЛЕВОГО ДНЯ
АТАКА ОТКАЗА ОБСЛУЖИВАНИЯ

Добавить

Изменить

Удалить

512 + -

Отношение	А	Название концепта
ВЫШЕ		КОМПЬЮТЕРНАЯ АТАКА
▶ НИЖЕ		АТАКА НУЛЕВОГО ДНЯ
АССОЦИАЦИЯ	1	УЯЗВИМОСТЬ В ПРОГРАМ

Добавить

Изменить

Перейти

Удалить

512 + -

Фильтр

Текстовый вход
▶ АТАКА НА УЯЗВИМОСТЬ
АТАКОВАТЬ УЯЗВИМОСТЬ
УЯЗВИМОСТЬ ЭКСПЛУАТИРУЕТСЯ
ЭКСПЛУАТАЦИЯ УЯЗВИМОСТИ
ЭКСПЛУАТИРОВАТЬ УЯЗВИМОСТЬ
ЭКСПЛУАТИРУЕТСЯ УЯЗВИМОСТЬ

Перейти к синонимам

Фрагменты текстов

Заккрыть

Текстовый вход

▶ 0-DAY АТАКА
0DAY АТАКА
0DAY УГРОЗА
ZERO-DAY АТАКА
ZERO-DAY УГРОЗА
АТАКА НУЛЕВОГО ДНЯ

Добавить

Изменить

Удалить

Тезаурус по безопасности: Отношения онтологической зависимости

Отношения на концептах

ΕΙΣΑΓΩΓΗ

Название концепта
ИМПОРТ ВООРУЖЕНИЙ
ИМПОРТ ДАННЫХ
ИМПОРТЕР
ИМПОРТНАЯ ЗАВИСИМОСТЬ
ИМПОРТНАЯ КВОТА
ИМПОРТНАЯ ЛИЦЕНЗИЯ
ИМПОРТНАЯ ПОШЛИНА

DEPENDENCE ON IMPORT

Фильтр

Текстовый вход
ЗАВИСИМОСТЬ ОТ ИМПОРТА
ЗАВИСИМОСТЬ ОТ ИМПОРТНЫХ ТОВАРОВ
ИМПОРТНАЯ ЗАВИСИМОСТЬ

Перейти к синонимам

Фрагменты текстов

Добавить

Изменить

Удалить

1

+

-

Отношение	Название концепта
ВЫШЕ	ЗАВИСИМОСТЬ ЭКОНОМИКИ
НИЖЕ	ЭНЕРГОЗАВИСИМОСТЬ
ACCOL 2	ЗАМЕЩЕНИЕ ИМПОРТА
ACCOL 1	ИМПОРТ

IMPORT SUBSTITUTION

Добавить

Изменить

Перейти

Удалить

1

+

-

Текстовый вход
ЗАМЕЩЕНИЕ ИМПОРТА
ИМПОРТЗАМЕЩЕНИЕ
ИМПОРТОЗАМЕЩАЮЩИЙ
ИМПОРТОЗАМЕЩЕНИЕ
ОТКАЗ ОТ ИМПОРТА

Добавить

Изменить

Удалить

Заккрыть

Наши работы: Использование тезауруса в информационных системах

/ТЕРМИН_РАСШ="ВРЕДОНОСНОЕ ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ"

ИНФОРМЕР: 16

Найти [Ctrl + Enter]

ранжировать по релевантности ↑

<ПО ВСЕМ>

Выбрать

Показано документов: 20 из 21798 Время поиска: 00:00:00.083 (инфо)

Ранг/ИД	Местоположение / Снимплет (<input type="checkbox"/> еще поля)
<div><div></div><div>0.95</div></div> <div>1 533329654</div>	<div>Схвачены хакеры, заразившие миллион телефонов</div> <div>Действия преступной группы, распространяющей вредоносные программы под названием «Cron» (Крон) для ОС Android, были зафиксированы системой киберразведки Group-IB весной 2015 года. Попадая на телефон жертвы, вредоносная программа злоумышленников получала возможность осуществлять переводы с ее банковского счета на счета, подконтрольные...</div> <div><div><div><div>Описание: Попадая на телефон жертвы, вредоносная программа злоумышленников получала возможность осуществлять переводы с ее банковского счета на счета, подконтрольные злоумышленникам. Программа Cron могла отправлять SMS-сообщения на указанные преступниками телефонные номера, пересылать текст получаемых жертвой SMS-сообщений на удаленные сервера, а также скрывать поступающие по SMS уведомления от банка. Специалисты советуют не рутовать свои мобильные устройства, своевременно обновлять прошивку, Операция по задержанию большей части преступной группы состоялась 22 ноября 2016 года – тогда на территории 6 субъектов Российской Федерации были задержаны сразу 16 ее участников.</div></div><div>Вести.Финанс 22.05.2017 17:11:56</div></div></div>
<div><div></div><div>0.94</div></div> <div>2 533000223</div>	<div>ЦБ начинает публиковать сообщения о выявленных киберрисках</div> <div>Первое сообщение посвящено массовой рассылке вредоносного программного обеспечения Wannacry (он же WannaCrypt, WanaCrypt0r 2.0). Специалисты ФинЦЕРТ Банка России разъясняют особенности Wannacry и подробно излагают рекомендации, которые были даны участникам финансового рынка для противодействия данному вредоносному программному обеспечению. ФинЦЕРТ...</div> <div><div><div>Описание: Новый формат прорабатывается Центром мониторинга и реагирования на компьютерные атаки в кредитно-финансовой сфере (ФинЦЕРТ) Банка России с участниками информационного обмена. Специалисты ФинЦЕРТ Банка России разъясняют особенности Wannacry и подробно излагают рекомендации, которые были даны участникам финансового рынка для противодействия данному вредоносному программному обеспечению. ФинЦЕРТ Банка России в апреле 2017 года разослал участникам информационного обмена информацию о</div></div></div>

0.94

3 531469110

[Пора проверить свой Mac на заражение OSX.Snake](#)

заражение OSX.Snake К сожалению, нам приходится все чаще сообщать о новых примерах **вредоносного ПО**, найденных пользователями macOS и OS X. Новая **вредоносная программа** изначально использовала уязвимости операционной системы Windows. Теперь она начала свою работу и на компьютерах Mac, маскируясь под установочный файл Adobe Flash Player и прячась в...

Описание: Пора проверить свой Mac на заражение OSX.Snake <cirpar num="1">К сожалению, нам приходится все чаще сообщать о новых примерах вредоносного ПО, найденных пользователями macOS и OS X. Новая вредоносная программа изначально использовала уязвимости операционной системы Windows. Теперь она начала свою работу и на компьютерах Mac, маскируясь под установочный файл Adobe Flash Player и прячась в уже существующие папки. Это уже вторая вредоносная программа, о которой стало известно за последнюю неделю, использующая действующий сертификат. </V

Appleinsider

06.05.2017 15:13:09

0.94

4 531082028

[Преступления в сфере IT становятся более профессиональными – эксперт](#)

эксперт 3 мая 2017, 13:38 , Фото: university.innopolis.ru Основные тренды в сфере **вредоносного кода** — «профессионализация» и «монетизация», отмечает специалист по IT-безопасности Сергей Петренко. В Казани меморандум о согласии подписали iCarnegie Learning University и ИТ-университет (Казань, 3 мая, «Татар-информ», Рустам Кильсинбаев). Успешные...

Описание: По глобальной сети в каждый момент времени распространяются сотни уникальных, достаточно проработанных, свежих (не определяемых сигнатурными методами) экземпляров вредоносного кода, формируя массивы зараженных рабочих станций и серверов (ботнеты). Подавляющее большинство из них умеет распространяться одновременно по нескольким различным схемам, временно нейтрализовать антивирусное ПО и подключаться к центрам контроля за инструкциями и "свежим" программным кодом», — комментирует специалист. Сергей Петренко отмечает, что с течением времени и увеличением пропускной способности каналов связи пропорционально растут масштабы распространения вредоносного кода.

ИА «Татар-информ»

03.05.2017 14:03:06

0.94

5 524896338

[38 смартфонов производства Samsung, Lenovo, LG, ASUS, Vivo, Oppo, ZTE и Xiaomi "комплек](#)

производителей смартфонов на платформе Android были обнаружены серьезные вредоносные **вирусы**, внедренные вручную во время прохождения через цепочку поставщиков. Хотя **вредоносное ПО** обычно ассоциируется со случайной загрузкой вредоносных файлов пользователем, похоже, что компрометирующие приложения были добавлены на Android-смартфоны во время их...

Описание: Хотя вредоносное ПО обычно ассоциируется со случайной загрузкой вредоносных файлов

Пример выдачи информационно-аналитической системы

Version: 9

Закладки: 0 документов

Коллекция: Progress 2017

Запрос (?)

док. 50

Точность: точно (0.00)

зона <Основная>

☐ отчет ☐ искать в найденном ☐ XML-граф ☐ отчет по sentimentу ☐ расширить

по кластерам

демографические угрозы

Найти [Ctrl + Enter]

ранжировать по релевантности ↑

<ПО ВСЕМ>

Показано документов: 50 из 805

Время поиска: 00:00:00.073 (инфо)

Ранг/ИД	Местоположение / Снимок (<input type="checkbox"/> еще поля)
<div><input type="checkbox"/> 0.85</div> <div>1 550815034</div>	<div>Заголовки утра. 20 ноября</div> <div>« Иммиграционный зов »: эксперты ЦСР назвали главные демографические «народные» облигации банков, рассказывается в материале «Купоны вместо поддержки продажи алкоголя в интернете ». С нового года в сеть попадет разведут по-новому »:...</div> <div><div>Описание: Канцлер британского казначейства Филип Хаммонд допустил финансовые предложения, связанные с отсоединением от ЕС, еще до гл. The Guardian: В Британии нет безработных, неуместно заявил Хаммонд, says Hammond in TV gaffe). Reuters: Американский ядерный генерал заявил, что он воспротивился бы "нелегальному" приказу Трампа нанести удар » (U.S. nuclear general says would resist 'illegal' Trump strike order).</div><div>Полит.ру</div><div>20.11.2017 09:02:35</div></div>
<div><input type="checkbox"/> 0.85</div> <div>2 550795021</div>	<div>Иммиграционный зов</div> <div>Иммиграционный зов Иммиграционный зов Эксперты ЦСР назвали главные демографические угрозы России Эксперты Центра стратегических разработок Алексея Кудрина сформулировали основные демографические вызовы в России. Чтобы преодолеть их, члены экспертного совета предлагают привлекать еще больше мигрантов и давать им гражданство Фото: Владимир...</div> <div><div>Описание: Иммиграционный зов Иммиграционный зов Эксперты ЦСР назвали главные демографические угрозы России Эксперты Центра стратегических разработок Алексея Кудрина сформулировали основные демографические вызовы в России. Чтобы решить их, нужно привлечь больше мигрантов, а россиянам придется повысить порог социальной терпимости, заключили аналитики. При этом авторы доклада отмечают, что прекращение роста населения является</div></div>

ТЕРМИН

+	-	+t	-t	ДЕМОГРАФИЧЕСКАЯ ОБСТАНОВКА
+	-	+t	-t	УБЫЛЬ НАСЕЛЕНИЯ
+	-	+t	-t	АБОРТ
+	-	+t	-t	РОЖДАЕМОСТЬ
+	-	+t	-t	ПРИРОСТ НАСЕЛЕНИЯ
+	-	+t	-t	ДЕМОГРАФИЧЕСКИЙ ПОКАЗАТЕЛЬ
+	-	+t	-t	ДЕМОГРАФИЧЕСКИЙ ПРОГНОЗ
+	-	+t	-t	СОКРАЩЕНИЕ РОЖДАЕМОСТИ
+	-	+t	-t	ЧИСЛЕННОСТЬ НАСЕЛЕНИЯ
+	-	+t	-t	ВОЕННАЯ КОНФРОНТАЦИЯ
+	-	+t	-t	МИЗУЛИНА, ЕЛЕНА БОРИСОВНА
+	-	+t	-t	ДЕМОГРАФ
+	-	+t	-t	ГЕОПОЛИТИКА
+	-	+t	-t	НАЦИОНАЛЬНАЯ БЕЗОПАСНОСТЬ
+	-	+t	-t	ЕСТЕСТВЕННАЯ УБЫЛЬ НАСЕЛЕНИЯ
+	-	+t	-t	ИНСТИТУТ СЕМЬИ

ИНФОРМЕР: 16

В настоящее время

- Глобальные поисковики
 - Автоматически порождаемые тезаурусы
 - Рост поведенческих факторов
 - Обогащение представления содержания документов
 - Нейронные сети для создания векторных представлений
- Поиск в предметной области
 - Используются по крайней мере списки синонимов
 - Правовая область:
 - налог на добавленную стоимость – НДС
 - Больничный лист- медицинский бюллетень – листок нетрудоспособности

Средняя точность AP

- Было сказано, что это комбинированная мера, учитывает точность и полноту.
- Вопрос: как учитывается полнота

$$AP = \frac{\frac{1}{1} + \frac{2}{3} + \frac{3}{4} + \frac{4}{6} + \frac{5}{10} + \dots}{10}$$