Система двуязычного (русско-английский, англо-русский) поиска в массиве научных публикаций с разрешением многозначности запросов

Функционал системы 1

Система предполагает следующие функции:

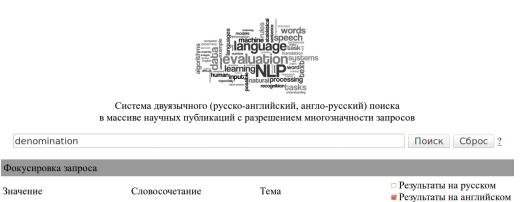
- формулирование запроса в виде фразы или ключевых слов на одном языке (русском, английском);
- уточнение значения ядра запроса за счет следующих механизмов (выбора из списка значений, выбора сопутствующих словосочетаний или терминов, выбора тематической области);
- поиск на двух языках (русском, английском);
- ранжирование и отображение результатов поиска.

Интерфейс 2

Значение

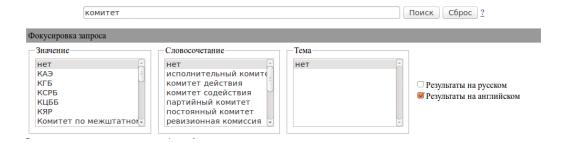
2.1Форма запроса

Базовая часть интерфейса пользователя – это экранная форма поискового запроса. Запрос может быть задан на любом из языков (русский, английский), система определяет язык автоматически.



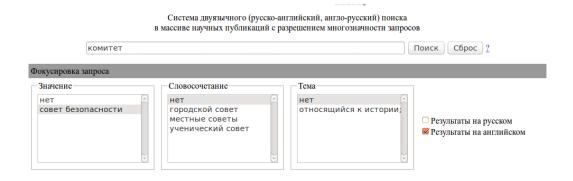
Выбор языковой принадлежности документов, возвращаемых на запрос системой осуществляется с помощью установки флагов «Результаты на русском» и «Результаты на английском» на форме запроса. По умолчанию обе опции включены. При отключении обоих флагов результат всегда будет пустым.

Уточнение запросов для разрешения многозначности осуществляется с помощью списков выбора группы «Фокусировка запроса» формы поискового запроса.

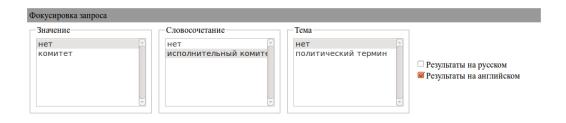


В системе используются три сценарии разрешения многозначности при фокусировке запроса:

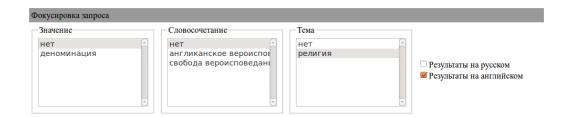
• уточнение значения путем выбора из предложенного списка;



• уточнение значения путем выбора из списка словосочетаний;



• уточнение значения путем задания темы.



Возможно сочетание механизмов фокусировке в любом порядке. Сделанный выбор по одному из механизмов уточнения запроса может сократить варианты выбора по другим механизмам. Например, при выбранном значении останутся только те словосочетания, которые соответствуют данному значению.

2.2 Форма вывода результатов

Результаты выводятся системой в соответствии с ранжированием по метрике tf-idf. Если результатов более 10, то они разбиваются на страницы по 10 результатов с кнопками навигации по этим страницам. Перед списком результатов указывается конкретный запрос (на двух языках) заданный системе и количество возвращенных документов.

Результат включает в себя: название статьи, ссылку на документ, список авторов и поисковый сниппет с отрывком документа, в котором подсвечен искомый текст. < Предыдущая Страница 1 Следующая >

1. Training the Next Generation of Astronomers

Peter K. G. Williams, Eric Huff, Holly Maness, Maryam Modjaz, Kristen L. Shapiro, Jeffrey M. Silverman, Linda Strubbe, Betsey Adams, Katherine Alatalo, Kuenley Chiu, Mark Claire, Bethany Cobb, Kelle Cruz, Louis-Benoit Desroches, Melissa Enoch, Chat Hull, Hannah Jang-Condell, Casey Law, Nicholas McConnell, Rowin Meijerink, Stella Offner, John K. Parejko, Jonathan Pober, Klaus Pontoppidan, Dovi Poznanski, Anil Seth, Steven Stahler, Lucianne Walkowicz, Andrew A. West, Andrew Wetzel, David Whysong, 2009 of UC Berkeley faculty indicates that they spend approximately 25% of their work time on research (excluding student and postdoc interaction), 25% on teaching, 19% on administrative duties (including committee participation and large-scale project management), 14% on advising students and postdocs, 12% on securing funding (for personal research as well as observatories/organizations/departments as a whole), and 5% on public outreach\footnote {Data available in ASCII format online at \url

2. Lattice Multiverse Models

S. Gill Williamson, 2010

 $\begin{array}{l} \overline{\text{N}^{\prime}} \text{k}, let \Phi^{D}_{r} = \{F^{x}_{r}((y_{1}, n_{1}), (y_{2}, n_{2}), \ldots (y_{r}, n_{r})) \mid y_{1}, \ldots, y_{r} \in G^{x}_{D}, \ r \geq 1, \ F^{x}_{r} \in \mathcal{F}_{G}\} \\ \text{bethesetofde fined values of } F_{r}' = \text{n.r. } If \text{Phi^{D}}_{r} \text{ wise a lempty set, } \\ lets_{D}(x) \text{bethem in imumover Phi^{D}}_{r} \text{ is the in the ``committee'} \text{ labeling} \\ \text{substitutes called this the ``committee'} \text{ labeling} \\ \end{array}$

3. Opportunities in Delivery of Preventive Services in Retail Settings

Yaneer Bar-Yam, Dion Harmon, Keith Nesbitt, May Lim, Suzanne Smith, Bradley A. Perkins, 2012 preventive services; and what reductions in time and cost would be obtained through efficiency in a retail context. \section {Methods} \label{sec:methods} We estimated the required number of repetitions (from target population and frequency), time and cost in the traditional healthcare setting for 35 of the services recommended by United States Preventive Services Task Force (USPSTF) [4], the Advisory Committee on Immunization Practices (ACIP) [17], and the Centers for Medicare and Medicaid Services

3 Список обрабатываемых запросов

Система кросс-поиска работает в демо-режиме. В настоящий момент проработаны двуязычные запросы по ключевым словам, относящимся к тематике организаций.

3.1 Русскоязычные запросы

авиакрыло
авиалиния
авиапредприятие
авиация
агенство
академия
альянс
антанта
армия
артиллерия
ассоциация
ассоциация ветеранов
банда

баскетбольная команда

батальон

баттэри

бейсбольная команда

бизнес

благотворительная организация

благотворительность

блок

больница

братство

бригада

бюро

ведомство

вера

ветвь

взвод

власти

военнослужащие

военный резерв

воинство

войска

войско

враг

газета

гарнизон

гильдия

главная лига

Голливуд

государство

двор

делегация

деноминация

департамент

депутация

держава

деташемент

джаз-банд

диван

дивизион

дивизия

дилер

дипломатическая миссия

дистрибьютор

дом

домашние

домочадцы

железные дороги

женский клуб

журнал

жюри

закон

защита

звено

земля

иерархия

изготовитель

империя

институт

кавалерия

картель

квартет

квинтет

клика

клиника

клуб

коалиция

колледж

коллектив

команда

командование

комбинат

комиссия

комитет

коммерческая концессия

компания

конгломерат

конница

контингент

конфедерация

конференция

концерн

кооператив

кооперативное общество

корпорация

корпус

культ

легион

лига

личный состав

лобби

ложа

машина

министерство

миссия

младшая лига

монорельс

музыкальный коллектив

мушкетеры

наряд

нация

оборона

общество

объединение

объединенный резерв

октет

ополчение английских горожан

кирикоппо

орган

организация

орден

ОРИТ

оркестр

оркестровая группа

ось

отдел

отделение

партия

патруль

пехота

подкомиссия

подпольная организация

подразделение

покупатель

политбюро

полицейские силы

полиция

полк

поразделение

посольство

правительство

правление

предприятие

принимающая сторона

притон

производитель

промышленность

профсоюз

пул

разведка

режим

рейдер

религия

рота

рынок

секстет

семинария

семья

септет

сеть

сила

синдикат

система жирорасчетов

служба

смена

совет

содружество

сообщество

сопротивление

союз

союз государств

союзник

союзники

сторона

страна

судьи

супердержава

товарищество

торговое судно

торговый банк

трест

трио

труппа

ударные

университет

управление

учреждение

фаланга

федерация

фирма

флотилия

фонд

франшиза

футбольная команда

хедж-фонд

xop

церковь

цирк

член

шайка

школа

штаб

штат

эквайер

экипаж

экспедиция

эмитент

энтерпрайз

эскадра

эскадрилья

эшелон

юрисдикция викария

ячейка

3.2 Англоязычные запросы

academy acquirer agency airline airway alignment alinement alliance Allies ally argosy arm army artillery association authorities authority aviation axis band bank battalion battery bloc board branch brigade brotherhoodbureau business cartel cavalry cell chain charity chorus church circus clinic

club

coalition

collective

college

combination

combine

combo

command

commission

committee

commonwealth

community

company

complement

concern

confederacy

confederation

conference

conglomerate

constabulary

contingent

cooperative

corp

corporation

corps

council

country

court

crew

cult

dealer

dealership

defence

 ${\tt defense}$

delegacy

delegation

 ${\tt denomination}$

department

deputation

detachment

detail

distributor

divan

division

diwan

duo

echelon

eleven

embassy

empire

enemy

entente

enterprise

establishment

expedition

faith

family

 ${\tt federation}$

fellowship

firm

five

fleet

flight

foot

force

foundation

franchise

frat

 ${\tt fraternity}$

fund

gang

garrison

gild

giro

 ${\tt government}$

guild

 ${\tt headquarters}$

hedgefund

hierarchy

Hollywood

home

horse

hospital

host

house

household

ICU

industry

infantry

institute

institution

intelligence

issuer

jury

land

law

league

legation

legion

lobby

lodge

machine

magazine

majors

maker

manufacturer

market

member

menage

minors

mission

mob

monorail

musketry

nation

nest

newspaper

nine

octet

octette

office

opposition

orchestra

order

organ

pack

panel

paper

party

patrol

percussion

phalanx

platoon

police

politburo

pool

power

quartet

quartette

quintet

quintette

raider

railroad

railway

regime

regiment

religion

reserve

resistance

ring

school

section

seminary

septet

septette

service

sestet

sextet

sextette

shift

side

society

soldiery

sorority

soviet squad squadron staff state subcommittee subdivision superpower syndicate trainband trio troop troops troupe trust underground union unit university vicariate vicarship wing