****

**МИНОБРНАУКИ РОССИИ**

**федеральное государственное бюджетное образовательное учреждение**

**высшего образования**

**«Московский государственный технологический университет «СТАНКИН»**

**(ФГБОУ ВО «МГТУ «СТАНКИН»)**

|  |  |
| --- | --- |
| **ИНСТИТУТ**  информационных технологий и систем | **Кафедра**  информационных систем |

**Отчет по лабораторной работе №4**

по дисциплине «**Мировые информационные ресурсы и сети**»

Направление **09.03.02 Информационные системы и технологии**

**Руководитель,** **Солдатов А.В.**

преподаватель

**Студент,**

группа ИДБ–14–13 **Шуленина А.В.**

«\_\_\_» \_\_\_\_\_\_\_\_ 2018 г.

Москва 2018 г.

В ходе лабораторной работы были составлены 10 поисковых запросов, с использованием операторов языка поисковых запросов. Поиск был осуществлен в поисковых система Google, Yandex, Yahoo.

Запросы и результаты по ним представлены в таблице 1.

Таблица 1

Поисковые запросы

|  |  |  |  |
| --- | --- | --- | --- |
| Описание поискового запроса | Синтаксис запроса в Google и количество найденных документов | Синтаксис запроса в Яндекс и количество найденных документов | Синтаксис запроса в Yahoo! и количество найденных документов |
| Страницы, содержащие слова «котики» или «енотики», или оба слова | котики OR енотики  ~135 тыс. | котики | енотики  ~106 млн. | котики OR енотики  ~1,04 млн. |
| Страницы, содержащие цитату «Недопустимы доказательства» | "Недопустимые доказательства"  ~23,6 тыс. | "Недопустимые доказательства"  ~5 млн. | "Недопустимые доказательства"  ~50,9 тыс. |
| Страницы, содержащие фразу «с переподвыподвертом» в том виде, в котором она представлена | "с переподвыподвертом"  ~8,69 тыс. | !с переподвыподвертом  ~53 млн. | "с переподвыподвертом"  ~20,3 тыс. |
| Страницы обязательно содержащие слова «площадь», «Лондон» и возможно содержащие слово «Кинг» | Кинг +площадь +Лондон  ~472 тыс. | Кинг +площадь +Лондон  ~15млн. | Кинг +площадь +Лондон  ~30,7 тыс. |
| Поиск страниц с цитатой «К нам на… рассол», содержащей пропущенное слово | К нам на \* рассол  ~10 | "К нам на \* рассол"  =258 | К нам на \* рассол  ~93,5 тыс. |
| Поиск страниц, содержащих слово «безумие» и не содержащих слово «радужное» | безумие –радужное  ~6,01 млн | безумие –радужное  ~130 млн. | безумие –радужное  ~517 тыс. |
| Поиск страниц в домене рф | финальная стадия site:.рф  ~23,4 тыс. | финальная стадия domain:рф  ~19 млн. | финальная стадия site:.рф  ~101 тыс. |
| Поиск страниц, обязательно содержащих «штрафы», не содержащих «Недвижимость» и опционально содержащих «налоги» в домене рф | -Недвижимость + штрафы налоги site:.рф  =108000 | -Недвижимость +штрафы налоги domain:рф  ~191000 | -Недвижимость +штрафы налоги domain:ru  ~131000 |
| Страницы, обязательно содержащие «Грудинин» и одно из «счета» или «налоги», или оба и «счета» и «налоги», не содержащие «Швейцария» | +Грудинин +(счета OR налоги) –Швейцария  ~2 | +Грудинин +(счета | налоги) –Швейцария  ~29 млн. | +Грудинин +(счета OR налоги) –Швейцария  =5 |
| Поиск страниц содержащих фразу «Улица … фонарей» с пропущенным словом и обязательно содержащих точную фразу «сезон 1» | Улица \* фонарей +"сезон 1"  ~3,47 млн. | "Улица \* фонарей + сезон 1"  ~3 тыс. | Улица \* фонарей +"сезон 1"  ~31,8 тыс. |

Для расчетов значения точности выдачи и информационного шума поисковой системы был выбран следующий запрос:

Запрос -Недвижимость +штрафы налоги site:.рф

В поисковой системе Google

множество релевантных запросов - а=107982,

множество нерелевантных запросов - b=18,

множество релевантных, но не выводимых системой - c=207.

Точность выдачи:

TB=(107982/(107982+207))\*100%=99,8%

Информационный шум:

ИШ=(18/(107982+18))\*100%=0,02%

В поисковой системе Yandex

множество релевантных запросов - a=190932,

множество нерелевантных запросов b=68,

множество релевантных, но не выводимых системой - c=311.

Точность выдачи:

TB=(190932/(190932+311))\*100%=99,83%\

Информационный шум:

ИШ=(68/(190932+68))\*100%=0,03%

В поисковой системе Yahoo

множество релевантных запросов - а=130968,

множество нерелевантных запросов b=32,

множество релевантных, но не выводимых системой - c=520.

Точность выдачи:

ТВ=(130968/(130968+520))\*100%=99,6

Информационный шум:

ИШ=(32/(130968+32))\*100%=0.024%

Вывод: Поисковые системы Google, Yandex, Yahoo выводят релевантные запросы примерно с вероятность в 98-99% с учетом использования специальных операторов. Поисковая система Yandex выдает наибольшее количество запросов по сравнению с Google и Yahoo, но многие страницы повторяются.