## Министерство образования Республики Беларусь

# Учреждение образования «Белорусский государственный университет информатики и радиоэлектроники»

# Факультет Информационных Технологий и Управления Кафедра ИТАС

# Отчет По лабораторной работе №1 «Технология поиска информации»

Выполнил: ст. гр. 820601 Шведов А.Р.

Проверил:

Ярмолик В.И.

## 1 Цель работы

Необходимо практически освоить технологии эффективного поиска информации, изучить основные операторы и конструкции языка запросов, проанализировать полученные результаты.

## 2 Рыночная доля поисковых систем

Согласно ресурсу *StateCounter* следующие поисковики занимают существенные доли на белорусском рынке: *Google*, *Yandex*, *Mail.ru*. Их доли на рынке хорошо отображает монополию *Google*.

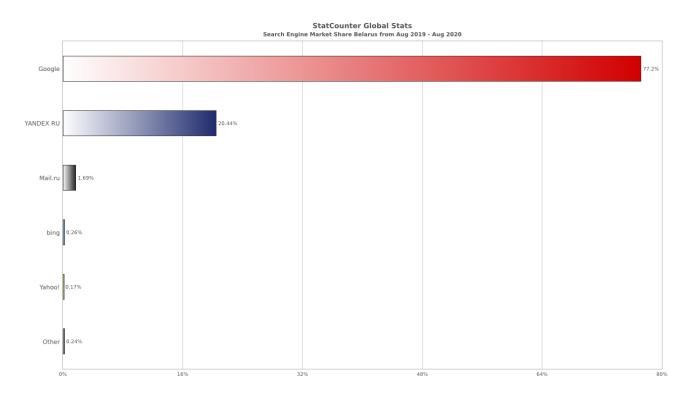


Рисунок 2.1 - Рыночная доля поисковых систем.

#### 3 Характеристика операторов

Не все поисковые системы поддерживают одинаковый набор операторов, однако есть определенные универсальные символы, которые позволяют составлять более сложные поисковые запросы.

### **Оператор** *AND*

Оператор *AND* (логическое **И**; сокращенное обозначение «**&**») образуют сложный запрос, по которому будут найдены только те документы, которые одновременно содержат оба аргумента оператора. Например, по запросу: «**информационная** *AND* **технология**» будут найдены только те документы, которые содержат и слово «**информационная**», и слово «**технология**». Аналогичный результат будет получен и при таком запросе: «**информационная & технология**».

## Оператор ОК

Оператор OR (ИЛИ; сокращенное обозначение «|») образуют запрос, по которому будут найдены все документы, удовлетворяющие хотя бы одному из аргументов оператора. По запросу: «информационная OR технология» будут найдены документы, которые содержат хотя бы одно из слов «информационная» или «технология», либо оба эти слова.

#### Кавычки

Для поиска цитат или слов в заданной форме можно использовать двойные кавычки. Слова по той части запроса, которая заключена в двойные кавычки, ищутся в документах именно так, как они в этих кавычках представлены. Например, запросу «информационная технология "применяться"» удовлетворяет документ, который содержащий текст «...информационная технология будет применяться...», но не удовлетворяет документ, содержащий «...информационная технология применяется...». А на поисковый запрос «"информационная технология применяется"» будут выданы только документы, содержащие с точностью до буквы именно такую

комбинацию слов, идущих подряд, при этом количество выбранных документов резко сократится.

#### **Оператор** *SITE*:

Позволяет использовать поисковую систему для поиска информации на конкретном сайте. Например, запрос «заявление site: bsuir.by» выведет результаты только с сайта bsuir.by, содержащие подходящее слово.

### 4 Результаты поисков

## 4.1 Оператор AND

Яндекс выдал 10 млн результатов (рисунок 4.1.1).

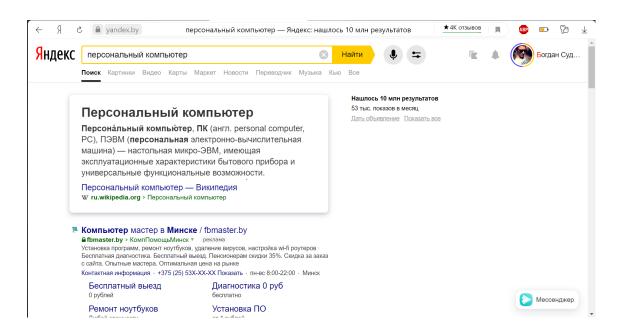


Рисунок 4.1.1 – Результаты поиска Яндекс.

Google нашел примерно 39 300 000 результатов (рисунок 4.1.2).

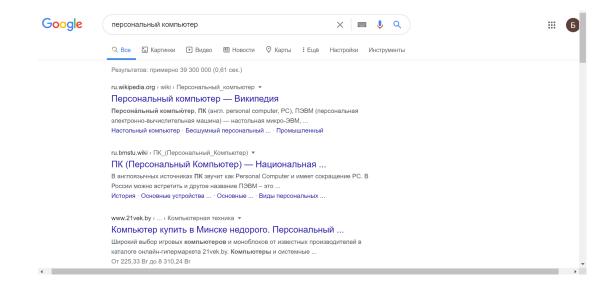
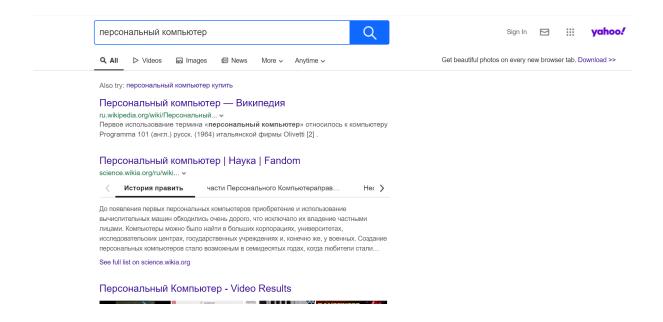


Рисунок 4.1.2 – результаты поиска *Google*.

*Yahoo!* предоставил 464 000 результатов, что значительно лучше, чем при поиске в Рамблере или *Aport.ru* (рисунок 4.1.3).



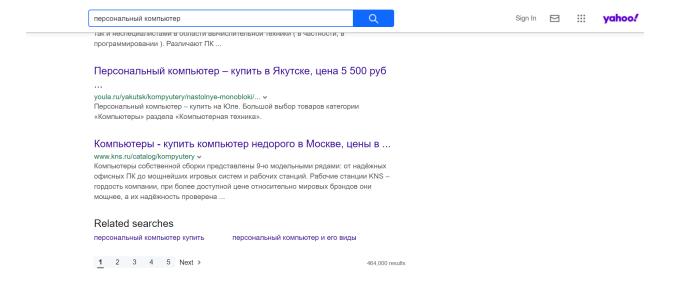


Рисунок 4.1.3 – результаты поиска *Yahoo!* 

Рамблер  $-310\ 000$  результатов (рисунок 4.1.4).

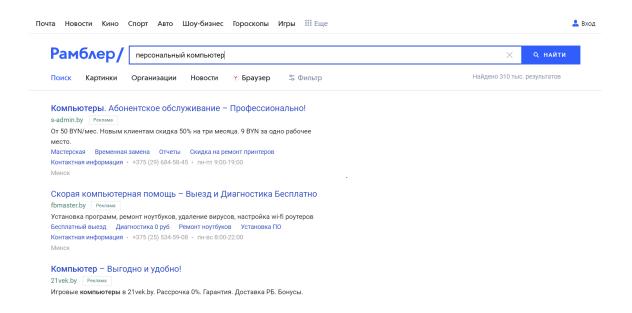


Рисунок 4.1.4 – результаты поиска Рамблер.

## **4.2** Оператор *OR*

Google нашел 214 000 000 результатов (рисунок 4.2.1).

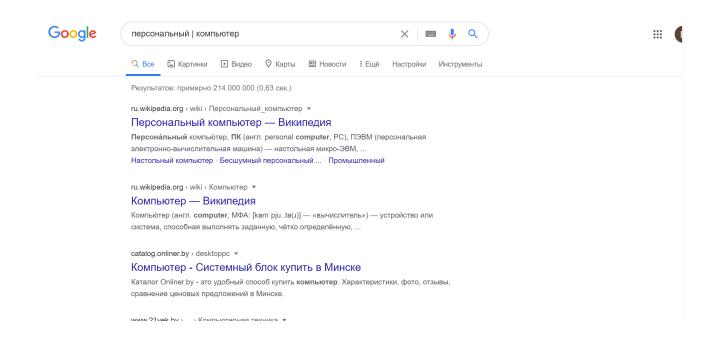


Рисунок 4.2.1 – результаты поиска *Google* 

Яндекс выдал 110 млн результатов (рисунок 4.2.2).

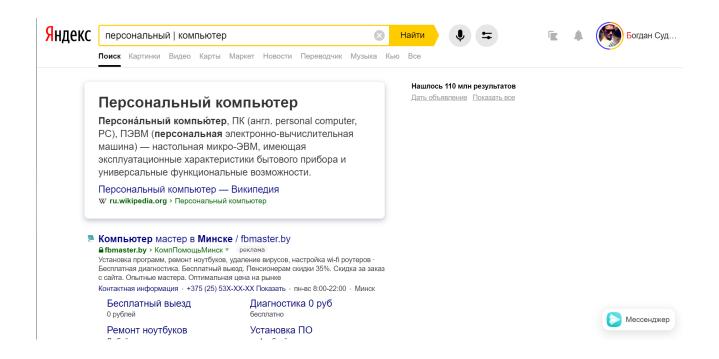


Рисунок 4.2.2 – результаты поиска Яндекс

Рамблер – 155 млн результатов (рисунок 4.2.3).

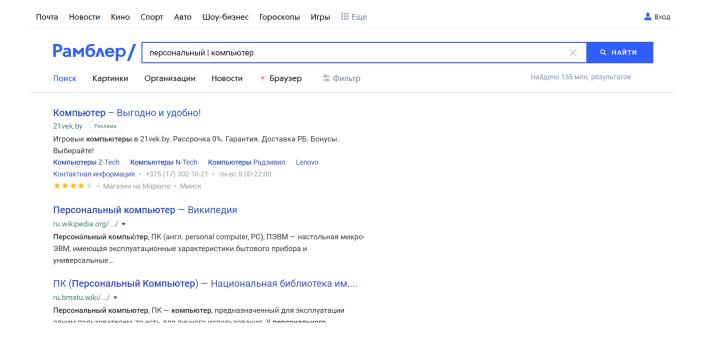
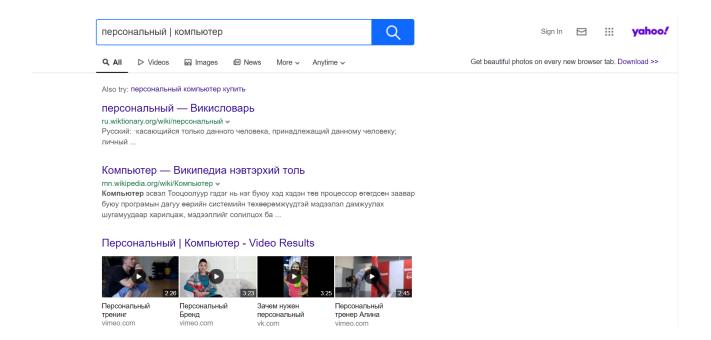


Рисунок 4.2.3 – результаты поиска Рамблер

## Yahoo! предоставил 20 800 000 результатов (рисунок 4.2.4).



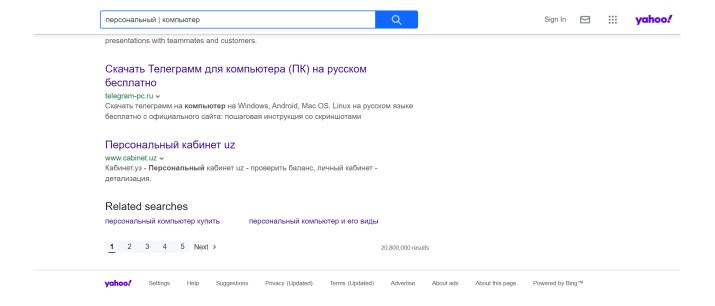
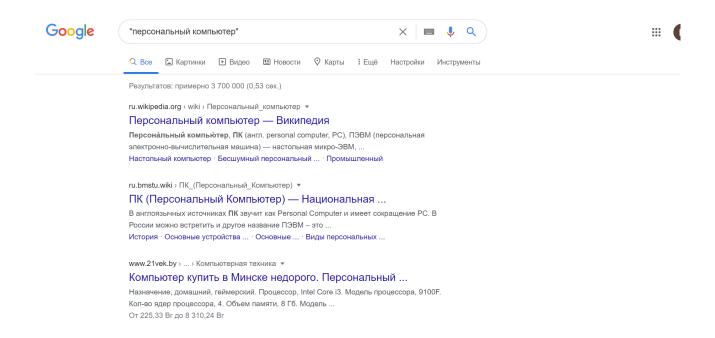


Рисунок 4.2.4 – результаты поиска *Yahoo!* 

#### 4.3 Кавычки

При использовании кавычек количество ссылок заметно сократилось, т.к. слова по той части запроса, которая заключена в двойные кавычки, ищутся в документах именно так, как они в этих кавычках представлены.

Результаты запроса в Google - 3700000, но тем не менее, это снова лучший результат из всех представленных поисковых систем (рисунок 4.3.1).



### Рисунок 4.3.1 – результаты поиска *Google*

Яндекс смог найти 24 тыс. результатов (рисунок 4.3.2).

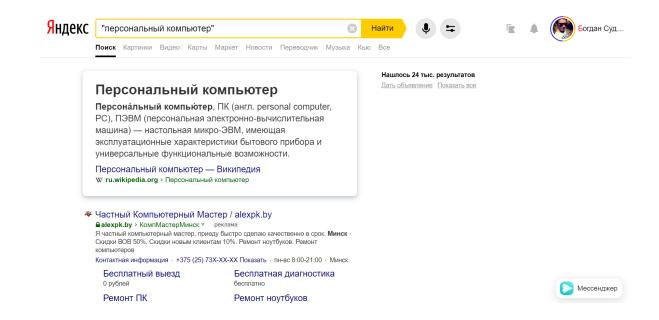
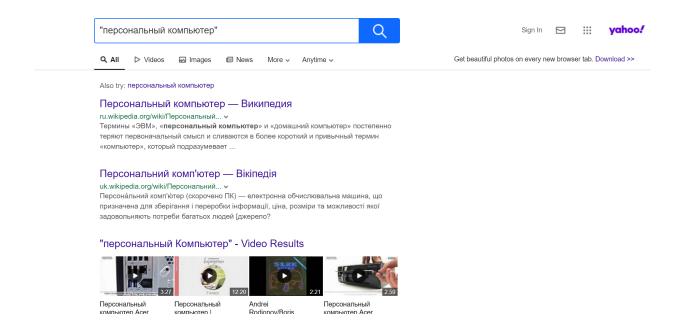


Рисунок 4.3.2 – результаты поиска Яндекс

*Yahoo!* показал 198 000 ссылок и даже при другом запросе по-прежнему занял второе место (рисунок 4.3.3).



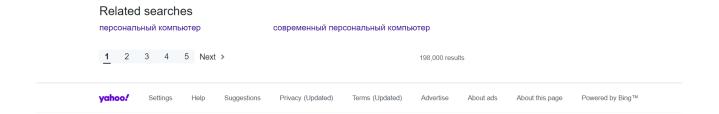


Рисунок 4.3.3 – результаты поиска Yahoo!

Рамблер вывел 2 результата (рисунок 4.3.5).

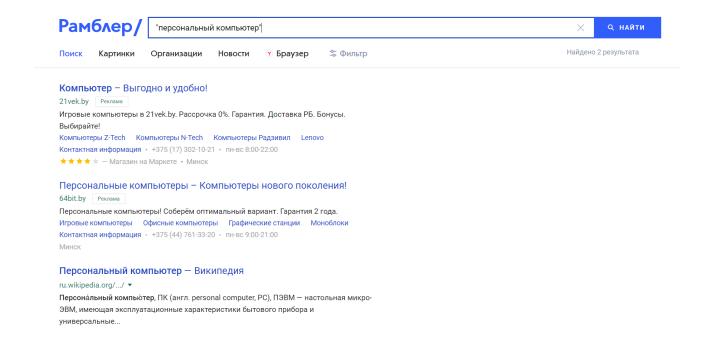


Рисунок 4.3.5 – результаты поиска Рамблер

#### **4.4 SITE**

Попробуем найти шаблоны заявлений на сайте bsuir.by. Google -2830 результатов. (рисунок 4.4.1).

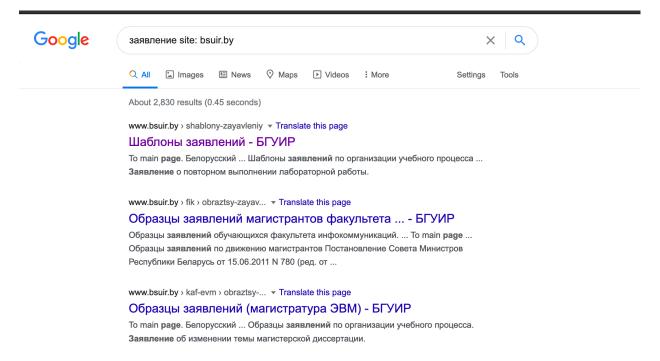


Рисунок 4.4.1 – результаты поиска *Google* 

Попробуем найти шаблоны заявлений на сайте *bsuir.by* Google – 2830 результатов. (рисунок 4.4.1).

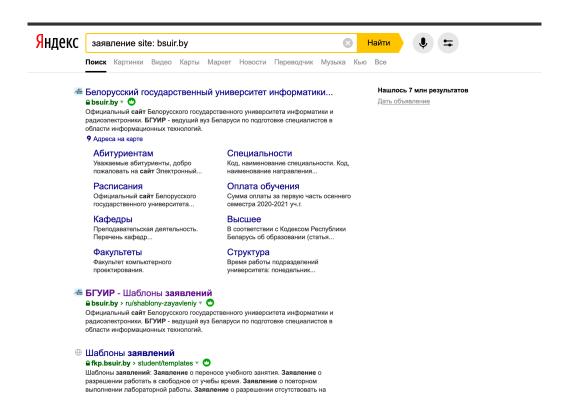


Рисунок 4.4.2 – результаты поиска *Yandex.ru* 

## 5 Вывод

На основании полученных результатов можно сделать вывод, что лидером оказалась поисковая система Google, на втором месте — Яндекс, на третьем — Yahoo!, на четвертом — Рамблер.