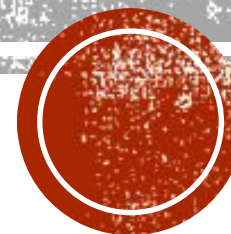


ПОИСКОВЫЕ СИСТЕМЫ В СЕТИ ИНТЕРНЕТ

Выполнил: Жумагалыев М

Сейілхан А

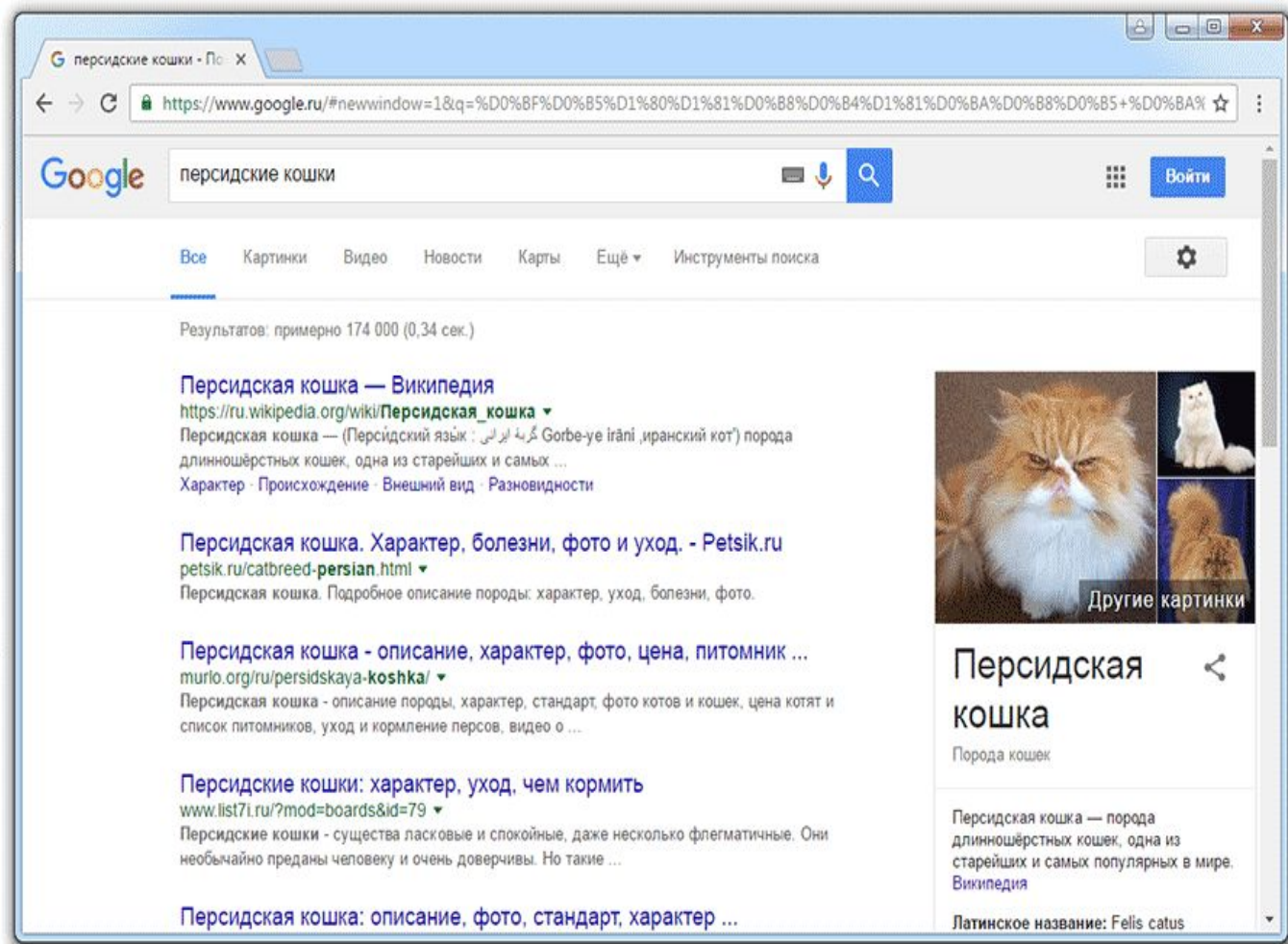
Проверял: Жумагалиев Б



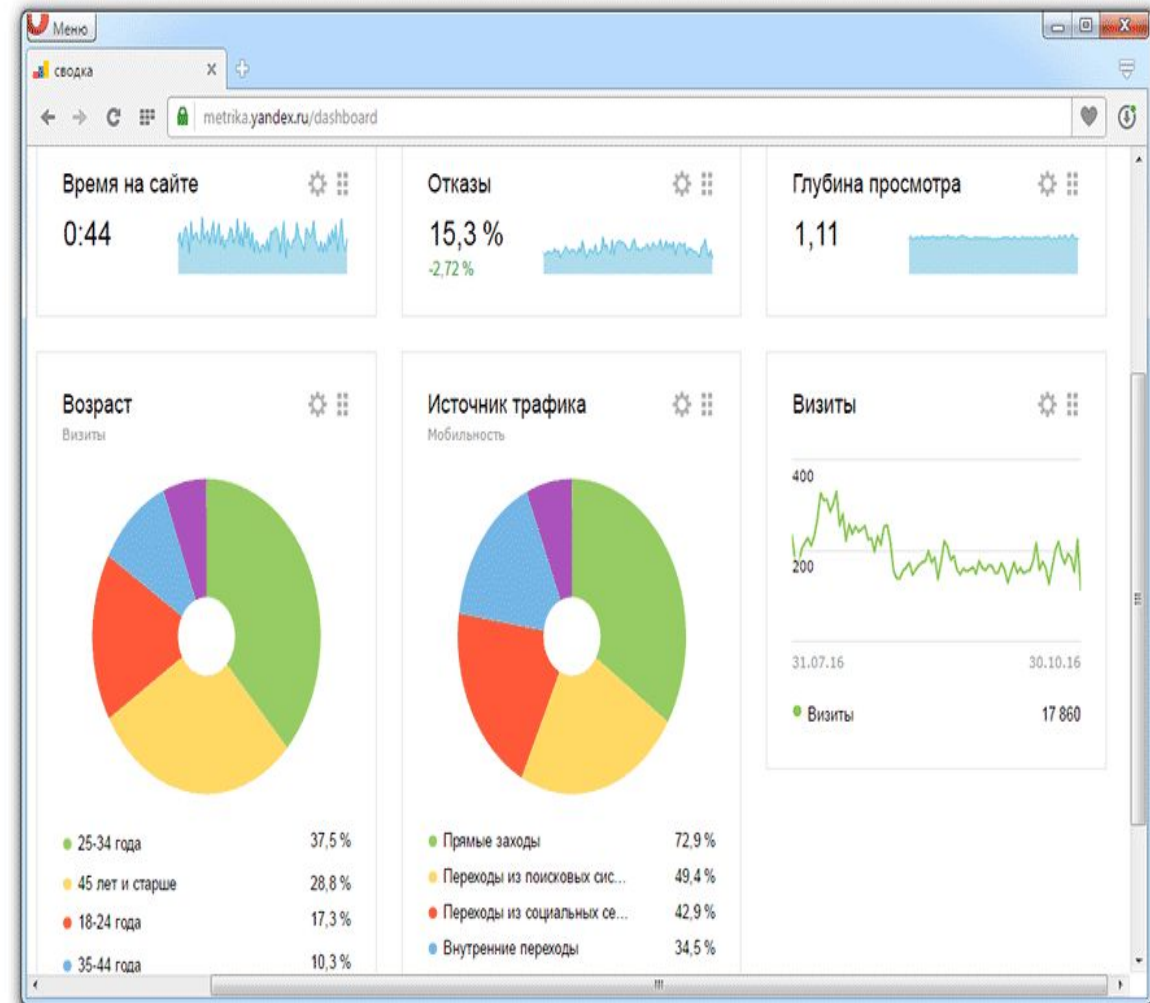
- Поисковые системы (ПС) уже приличное время являются обязательной частью интернета. Сегодня они громадные и сложнейшие механизмы, которые представляют собой не только инструмент для нахождения любой необходимой информации, но и довольно увлекательные сферы для бизнеса.



- **Поисковая система** – это специальный веб узел или по-другому сайт, который помогает пользователю находить нужную ему информацию в Интернете. Работает он так: вы печатаете то, что хотите найти, и сайт выдает список подходящих под этот запрос страниц.
- Получается, вы указываете ключевое слово или несколько слов, а система быстро анализирует все страницы, которые знает, и показывает подходящие результаты. Лучшие из них она ставит в начало списка.



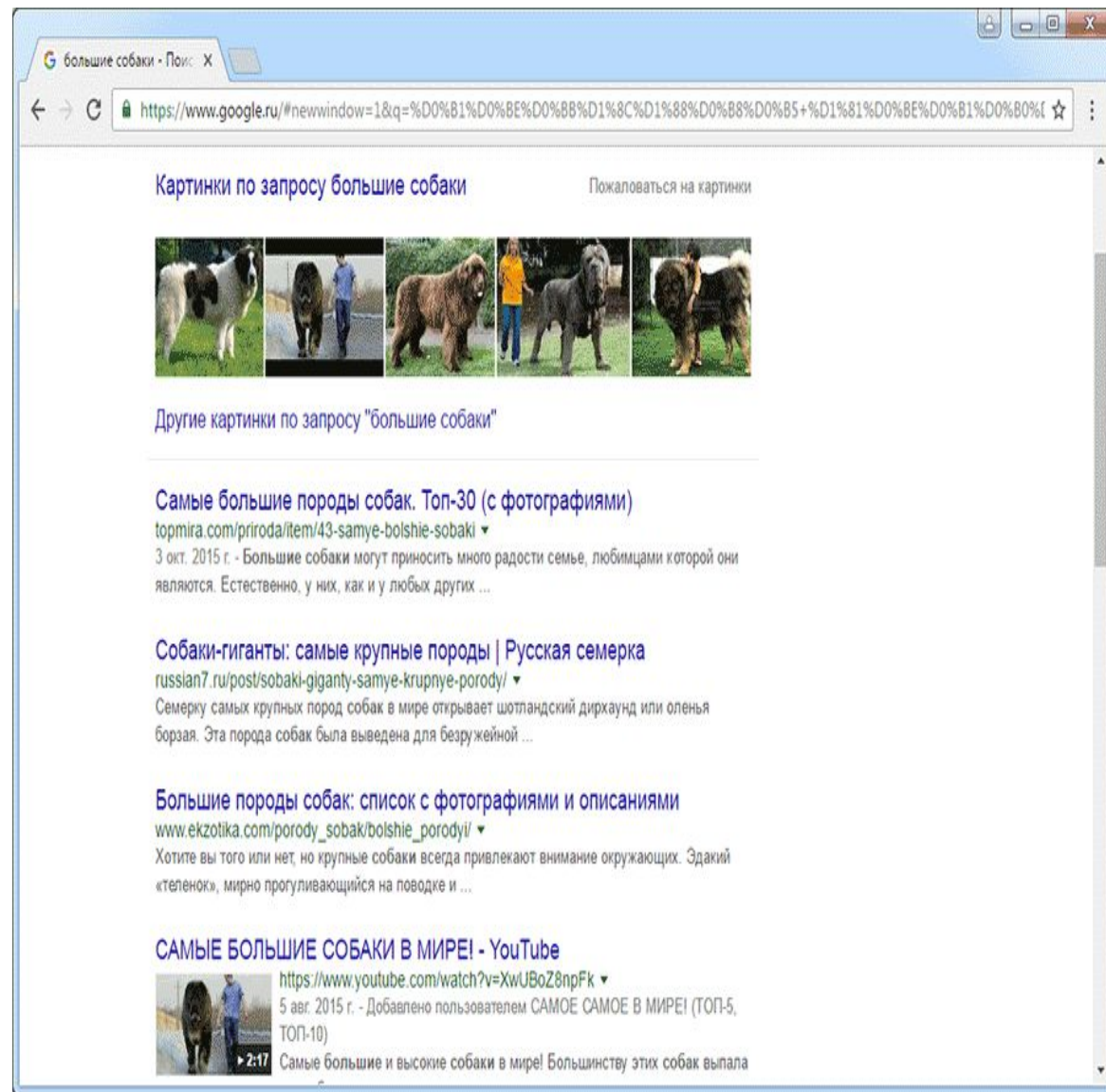
- *Речь идет не о сайтах, а именно о страницах. Возьмем, к примеру, сайт про кошек. На нем опубликовано двадцать статей о разных породах. Получается, на этом Интернет-ресурсе есть двадцать страниц с разной информацией. Поисковая система отдельно учитывает содержание каждой из них и показывает в своих результатах по подходящим запросам.*
- **Каждую страницу в сети поисковик тщательно изучает: анализирует, насколько статья полная, есть ли иллюстрации, видео. Потом он «смотрит» как ведут себя люди, которые заходят на эту страницу. Читают ли они ее до конца, делают это медленно и вдумчиво или быстро и поверхностно, куда щелкают мышкой и с какой скоростью крутят колесиком.**
- **Система собирает все эти данные и сравнивает с другими страницами Интернета. Плюс анализирует дополнительную информацию: делятся ли люди этим материалом в социальных сетях, оставляют ли комментарии и многое другое. И вот на основе всего этого поисковик выстраивает список результатов по запросу. В начале он ставит страницы, которые показали себя лучше остальных.**



- Последняя тенденция – персонализированный поиск. Система следит за каждым пользователем и анализирует конкретно его предпочтения. Поисковик учитывает, что пользователь смотрит в Интернете, какие темы ему интересны, о чем он пишет в комментариях. На основе этих данных предлагает то, что с наибольшей вероятностью понравится конкретно этому человеку.
- *Любителям теорий заговоров скажу, что никому эта информация не интересна и никуда она не уходит. Конечно, поисковики используют ее и в своих целях тоже, чтобы крутить наиболее подходящую рекламу. Но не более того.*



- **Поисковые запросы или ключевые слова** – это те слова и фразы, которые печатают люди, чтобы найти нужную им информацию.
- *Бывают они самыми разными. Например, я могу написать «большие собаки», а могу «крупные породы собак». В обоих случаях я хочу получить одну и ту же информацию.*
- По каждому запросу поисковик должен понять, что именно нужно человеку, и выдать подходящие результаты. Пользователь хочет прочитать статью или посмотреть фотографии? А, может, он хочет посмотреть видео? Или вообще купить? Система должна это моментально «подсекать» и предлагать варианты, которые устроили бы каждого.



The screenshot displays a Google search interface for the query "большие собаки". At the top, there are navigation buttons and the search bar. Below the search bar, the text "Картинки по запросу большие собаки" is visible, followed by a "Пожаловаться на картинку" link. A row of six image thumbnails shows various large dog breeds. Below the images, the text "Другие картинки по запросу 'большие собаки'" is present. The main search results section includes several links: "Самые большие породы собак. Топ-30 (с фотографиями)" from topmira.com, dated October 3, 2015; "Собаки-гиганты: самые крупные породы | Русская семерка" from russian7.ru; "Большие породы собак: список с фотографиями и описаниями" from www.ekzotika.com; and a YouTube video titled "САМЫЕ БОЛЬШИЕ СОБАКИ В МИРЕ!". The video thumbnail shows a person standing next to a large black dog, and the video duration is 2:17.



Категории поиска:

- по ключевым словам
- с булевой логикой объединения слов
- по словосочетаниям
- с учетом расстояния между словами
- с учетом регистра
- по семантике (основанная на использовании *СМЫСЛОВОГО* значения запрашиваемых фраз)
- по шаблону
- по полям документа

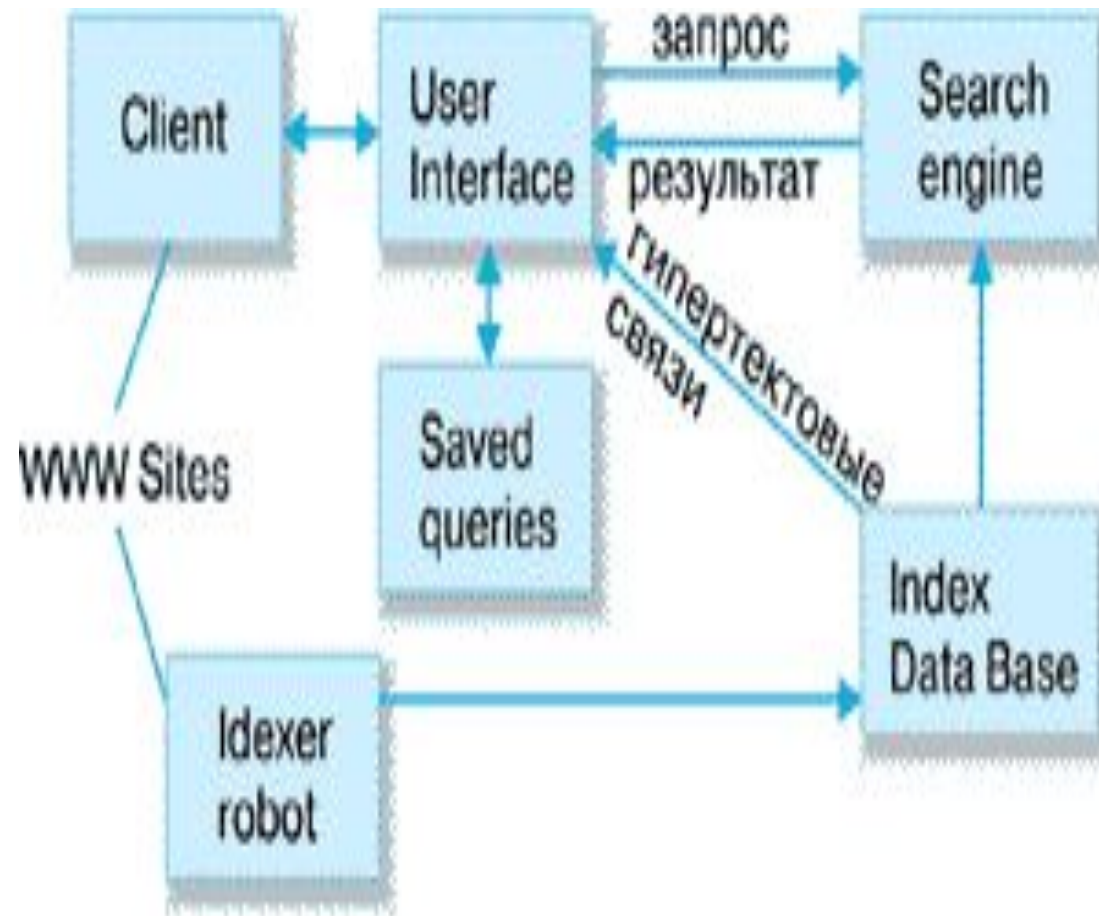


ПРАВИЛА ЭФФЕКТИВНОГО ПОИСКА В ИНТЕРНЕТЕ

- Проверяйте орфографию в написании слова.
- Используйте синонимы, если список найденных страниц слишком мал.
- Ищите больше, чем по одному слову. Максимально сужайте предмет поиска.
- Не начинайте обычные слова с прописной буквы, кроме имён собственных.
- Используйте ссылку «найти похожие документы», если один из найденных документов наиболее близок к искомому.
- Чтобы исключить документы, где встречается определённое слово, поставьте перед ним знак минус. И наоборот, если это слово наиболее важно — знак плюс.

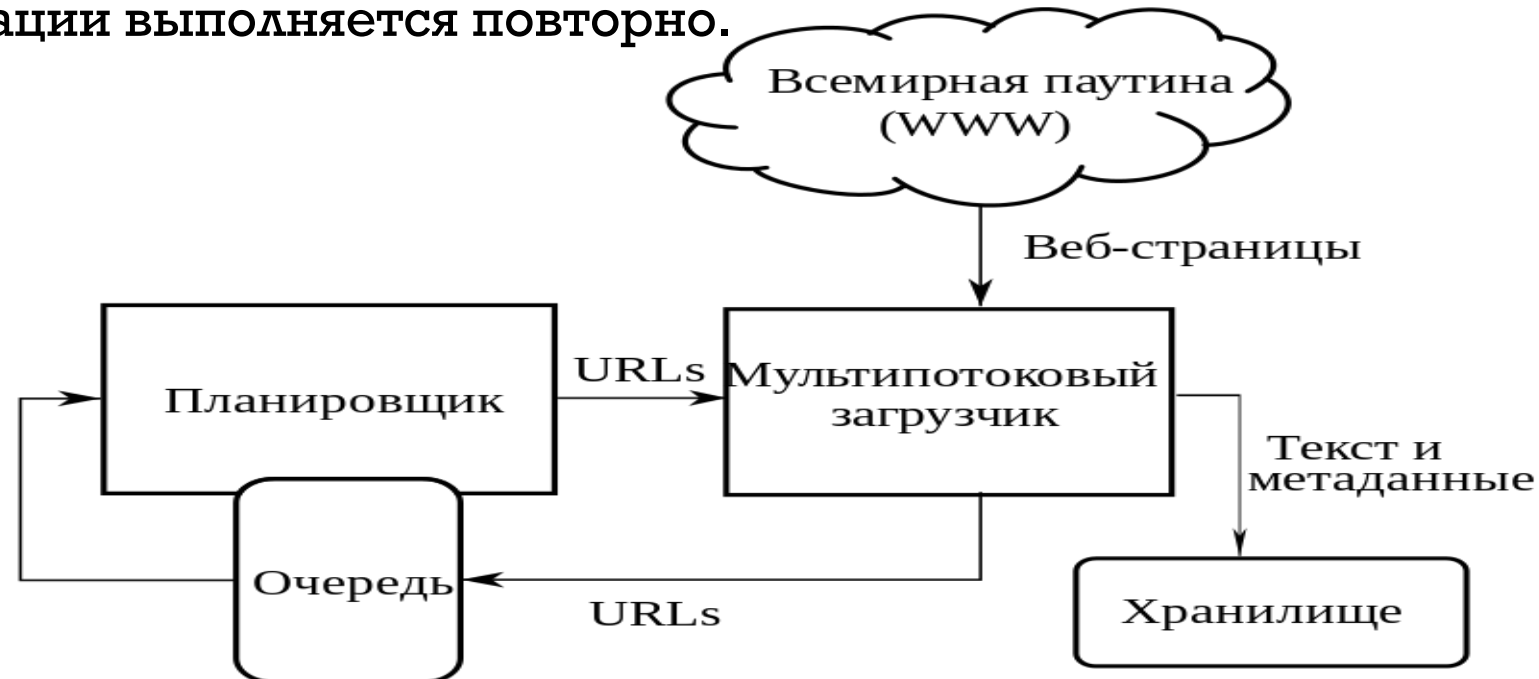


- Основные протоколы, используемые в Интернет не обеспечены достаточными встроенными функциями поиска. Нужную информацию уже не представляется возможным получить сразу, так как в сети сейчас находятся миллиарды документов и все они в распоряжении пользователей Интернет, к тому же сегодня их количество возрастает согласно экспоненциальной зависимости. Для того, чтобы структурировать информацию, накопленную в сети Интернет, и обеспечить ее пользователям удобными средствами поиска необходимых им данных, были созданы информационно-поисковые системы.



КАК РАБОТАЕТ ПОИСКОВАЯ СИСТЕМА

- Основные составляющие поисковой системы: поисковый робот, индексатор, поисковик.
- Как правило, системы работают поэтапно. Сначала поисковый робот получает контент, затем индексатор генерирует доступный для поиска индекс, и наконец, поисковик обеспечивает функциональность для поиска индексируемых данных. Чтобы обновить поисковую систему, этот цикл индексации выполняется повторно.

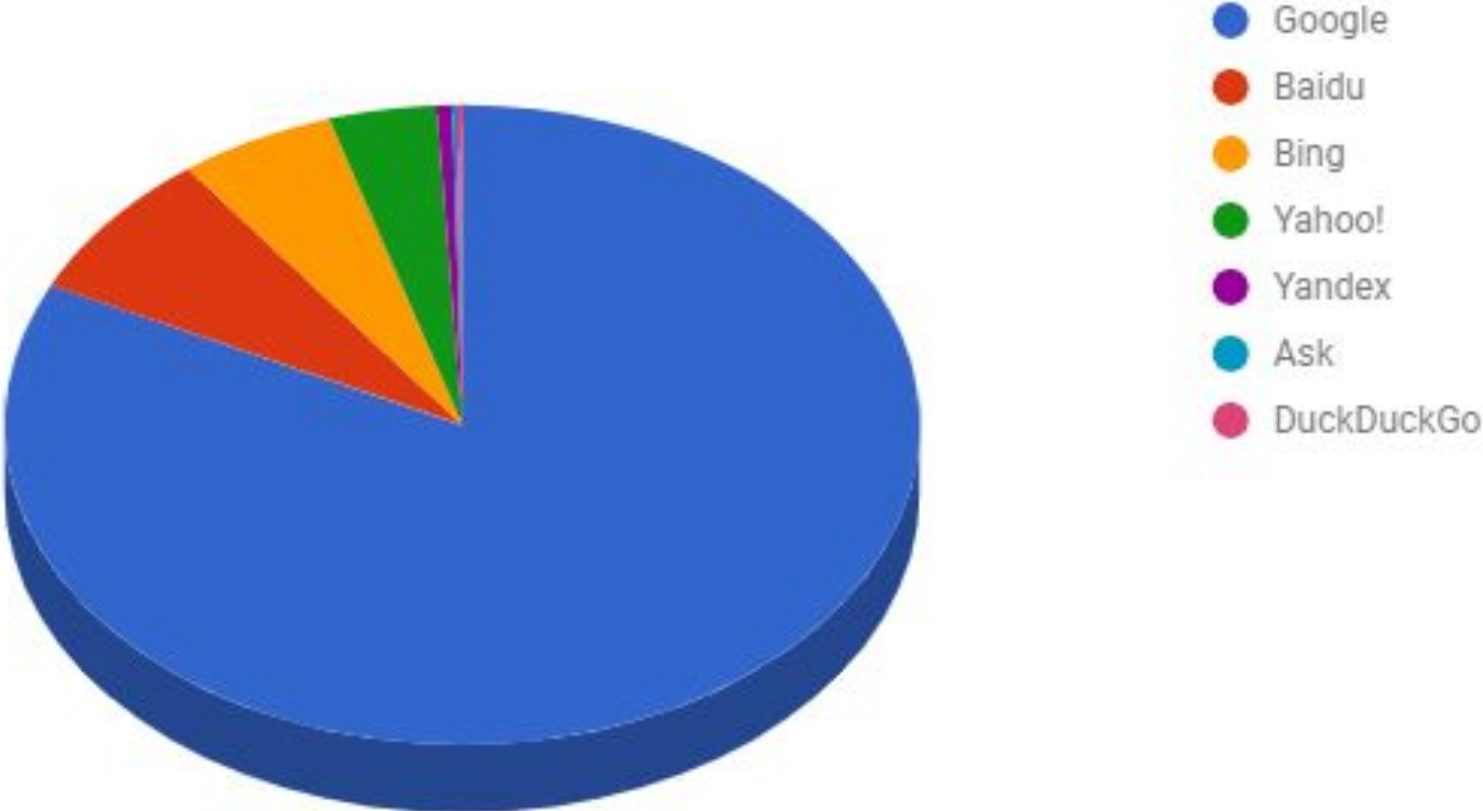


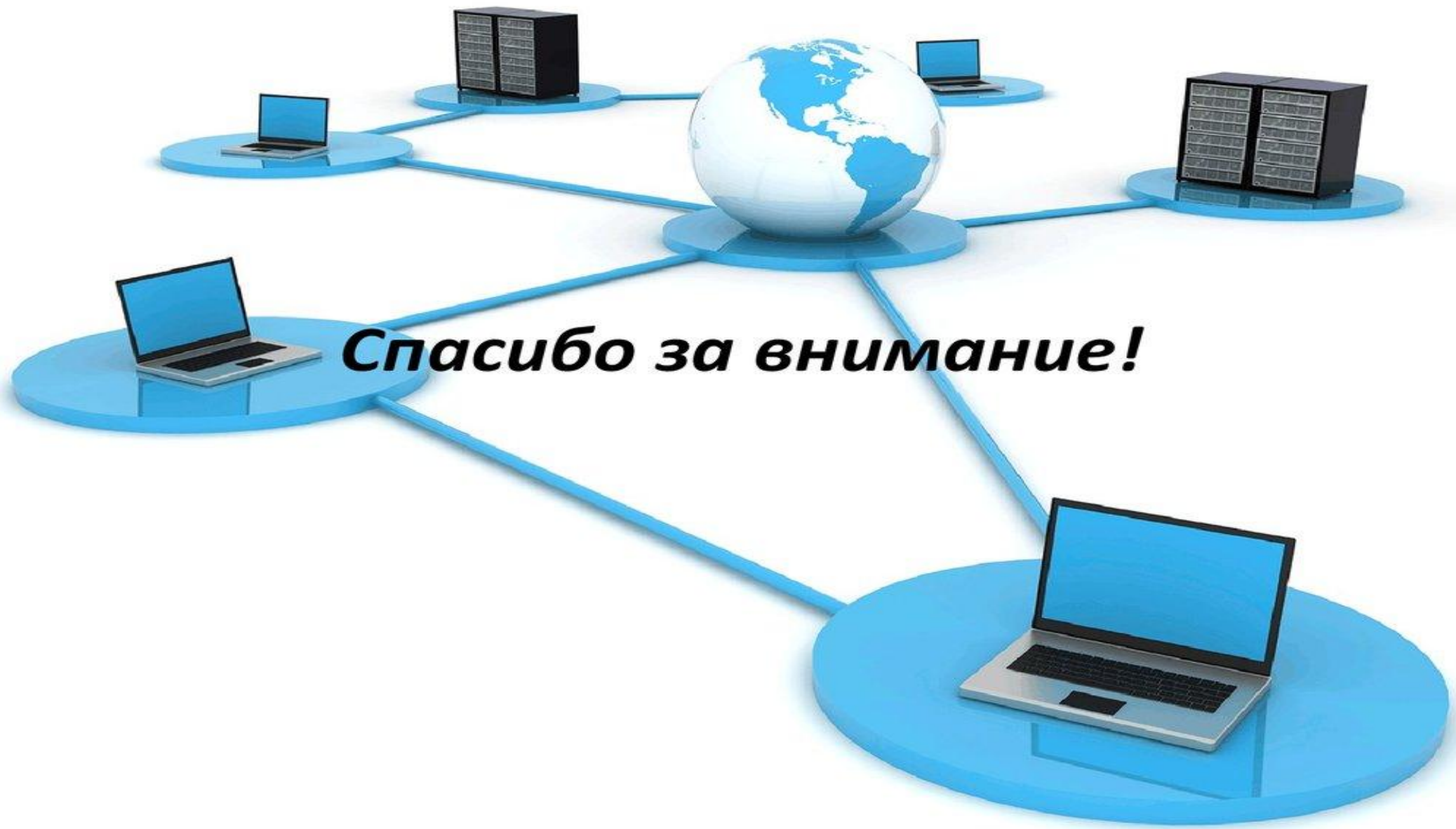
КАК РАБОТАЕТ ПОИСКОВАЯ СИСТЕМА

- Поисковые системы работают, храня информацию о многих веб-страницах, которые они получают из [HTML](#)-страниц. Поисковый робот или «краулер» ([англ. Crawler](#)) — программа, которая автоматически проходит по всем ссылкам, найденным на странице, и выделяет их. Краулер, основываясь на ссылках или исходя из заранее заданного списка адресов, осуществляет поиск новых документов, ещё не известных поисковой системе. Владелец сайта может исключить определённые страницы при помощи [robots.txt](#), используя который можно запретить индексацию файлов, страниц или каталогов сайта.



Популярность поисковых систем в мире





Спасибо за внимание!

