

## **Содержание:**

## **Введение**

В век информационных технологий огромную роль играет интернет. Объём доступной информации увеличивается с каждым днём и любое путешествие по просторам интернета невозможно без специальных поисковых систем, позволяющих комфортно просматривать веб-страницы.

Первоочередной задачей любой поисковой системы является доставление людям именно той информации, которую они ищут. А научить пользователей делать «правильные» запросы к системе, то есть запросы, соответствующие принципам работы поисковых систем, невозможно. Поэтому разработчики должны создавать такие алгоритмы и принципы работы поисковых систем, которые бы позволяли находить пользователям искомую ими информацию.

Цель работы – выбрать лучшие поисковые системы.

Для достижения поставленной цели необходимо решить следующие задачи:

- выбрать поисковые системы для анализа;
- произвести анализ и сравнение поисковых систем;
- в результате сравнения, выбрать наиболее подходящие поисковые системы.

### Глава 1. Поисковые системы

Поисковая система - это сервис, который позволяет пользователям Интернета искать контент через WorldWideWeb (WWW). Пользователь вводит ключевые слова или ключевые фразы в поисковую систему и получает список результатов веб-контента в виде веб-сайтов, изображений, видео или других онлайн-данных. Список контента, возвращаемого через поисковую систему пользователю, называется страницей результатов поисковой системы (SERP).

## **1.1 Задача поисковых систем**

Интернет непрерывно растёт и развивается, вместе с ним эволюционируют и поисковики интернета, но их главная задача остается неизменной – они должны

помочь пользователю найти самый лучший ответ на запрос, который он ввел в поле для поиска. Чем более качественные результаты в выдаче показывает поисковик, тем больше ему доверяют люди.

Поисковые системы постоянно анализируют терабайты информации, размещенной на миллионах web страниц, стараясь при этом определить какие сайты заслуживают попадания в ТОП выдачи, а какие являются лучшими кандидатами для попадания в бан.

## 1.2 Как работает поисковая система

Поисковик – это набор сложных программ и баз данных, которые действуют по определенному алгоритму. Упрощенно, этот алгоритм можно разбить на 3 этапа.

### 1.2.1 Поиск новых страниц

Поисковые системы выдают информацию не о страницах, находящихся в интернете, а о страницах, находящихся в базе данных поисковой машины. То есть, если сайт неизвестен Яндекс или Google, то и в выдаче он не появится.

Задача поисковика на этом этапе заключается в поиске всех возможных адресов страниц в интернете. Выполняет эту работу так называемый робот «паук».

Интернет - это ссылки и этот «паук» просто переходит по всевозможным ссылкам, записывая в свою базу адреса всех найденных страниц.

Попал на главную страницу сайта, на ней нашел ссылки на страницы рубрик, на страницах рубрик нашел ссылки на страницы со статьями, карточками товаров, ссылки на файлы или другую информацию. На каких-то из посещенных страниц одного сайта, он нашел ссылки на другие сайты – поисковая система переходит по ним и сканирует все, что нашла там.

### 1.2.2 Индексация

Как я уже писала выше – в поисковую выдачу попадает информация не с сайтов, находящихся в интернете, а информация из базы данных поисковой системы.

И следующая программка поисковика как раз занимается добавлением информации в базу. Она путешествует по всем известным адресам сайтов и страниц, копируя их содержимое на склады поисковой системы.

Называется этот процесс индексация – попадание информации в индекс поисковой системы.

Первый и второй процессы протекают непрерывно и, зачастую, одновременно. Постоянно пополняется база адресов страниц и база информации с этих страниц.

Кстати, в процессе индексации поисковые системы оценивают качество страниц, и информация некоторых из них не попадает в индекс. Как бы поисковик знает об их существовании, но по каким-то причинам считает их бесполезными для пользователя, поэтому не добавляет в выдачу – зачастую это не уникальный контент или служебные страницы.

### 1.2.3 Определение релевантности и ранжирование

Если то, что мы обсудили в предыдущих пунктах, работает непрерывно и независимо от внешних факторов (действий человека), то третий этап в алгоритме работы поисковых систем начинает действовать только под воздействием человека.

Когда в поисковике задается запрос, система начинает искать на него ответ в наполненной базе знаний по критериям, заданным человеком в этом запросе.

Сначала, система делает выборку, определяя все релевантные запросу страницы из известных (Релевантные – значит соответствующие, подходящие). Например, для запроса «купить холодильник Норд» релевантными будут страницы содержащие слова «купить», «холодильник», «Норд». Все страницы, содержащие одно или несколько из этих слов, попадут в выдачу поисковой системы.

Следующая задача поисковика, определить в какой последовательности пользователь увидит все эти страницы – их необходимо ранжировать. Факторов, которые будут влиять на порядок выдачи много, но если по-простому, то сначала пользователь увидит страницы, содержащие «купить холодильник Норд». Если таких нет, то ему будет предложено «купить холодильник» или «холодильник Норд» и в самом конце будут страницы со словами «купить», «холодильник», «Норд».

## **Глава 2. Метапоисковые системы**

Метапоисковая система — это поисковая система, которая в отличие от классических поисковых машин не имеет собственной базы данных и собственного поискового индекса, а формирует поисковую выдачу за счёт смешивания и переранжирования результатов поиска других поисковых систем.

## **2.1 Задачи метапоисковой системы**

Объем информации в Сети постоянно увеличивается, а часть данных динамически изменяется. По этой причине ни одна существующая поисковая система не может охватить все ресурсы Интернета. Таким образом, может возникнуть ситуация, когда ваша любимая поисковая машина не сможет найти нужный вам документ, хотя в Сети он присутствует. Тогда придется использовать другие поисковые системы в надежде найти то, что вам нужно.

Но есть и альтернативный способ решения этой проблемы – использование метапоисковой системы, которая представляет собой поисковый инструмент, посылающий ваш запрос одновременно нескольким поисковым системам.

Метапоисковая система не имеет собственной базы данных, а пользуется исключительно ресурсами других систем. Полученные результаты метапоисковая система обрабатывает и в соответствии со своим алгоритмом ранжирует в общий список.

## **2.2 Как работает метапоисковая система**

Принцип работы метапоисковика заключается в следующем: запрос пользователя преобразуется в запросы, отформатированные синтаксически и логически в конструкции, оптимальные для каждого отдельного, "традиционного" поисковика, т. е. из одного запроса метапоисковый механизм делает ряд запросов, которые адресуются нескольким "обычным" поискам.

Таким образом, метапоисковые системы объединяют для достижения результатов индексы поисковых серверов различных специализаций. В рамках одной метапоисковой системы можно осуществлять поиск информации различного типа. Метапоисковые системы не предназначены для индексирования и накопления данных, их назначение – чистый поиск и обработка результатов поиска. Все метапоисковые системы сортируют результаты поиска таким образом, чтобы

избежать дублей и одинаковых страниц в выдаче, найденных различными поисковыми системами.

Преимущество метапоисковых систем очевидно – они экономят значительное количество времени, избавляя вас от необходимости вводить запрос в каждой поисковой машине. С другой стороны, эти системы не позволяют вам воспользоваться уникальными возможностями отдельно взятых поисковых машин. Таким образом, поиски в Интернете удобно начинать с метапоисковых систем, а если это не помогло, то стоит попробовать детальный поиск на различных сайтах.

Также метапоисковые интернет-системы используют самые современные технологии обработки информации, например, кластеризацию. Суть такого подхода заключается в том, что полученное в ходе поиска по определенному ключевому слову подмножество страниц дополнительно анализируется, в результате чего выявляются новые ключевые слова, встречающиеся на веб-страницах вместе с терминами запроса. С большой вероятностью они относятся к отдельным аспектам интересующей пользователя темы. Далее найденные веб-страницы разделяются на отдельные кластеры в соответствии с наличием этих дополнительных ключевых терминов. Такая технология имеет ряд преимуществ. Пользователь получает подсказки, позволяющие уточнить запрос, что особенно важно при поиске по незнакомой предметной области. Применение кластеров также помогает поднять в результатах поиска узкотематические сайты, которые при обычном поиске могли оказаться далеко за пределами первой десятки результатов и, следовательно, вне поля зрения подавляющего большинства пользователей.

## **Глава 3. Сравнительный анализ поисковых систем**

### **3.1 Описание популярных поисковых систем**

Наиболее известные международные поисковые системы: Google, Yahoo, Baidu (лидер среди китайских поисковых систем), Bing, AOL, Lycos, Ask.

Русскоязычные — большинство «русскоязычных» поисковых систем индексируют и ищут тексты на многих языках — украинском, белорусском, английском, татарском и др.

Отличаются же они от «всеязычных» систем, индексирующих все документы подряд, тем, что в основном индексируют ресурсы, расположенные в доменных зонах, где доминирует русский язык или другими способами ограничивают своих роботов русскоязычными сайтами.

В России основной поисковой системой является «Яндекс», далее - Rambler, Aport, Mail.ru, Нигма.

А теперь более подробно об этих поисковых системах.

## **Google**

Google — поисковая система, принадлежащая корпорации Google Inc.

Первая по популярности (84,65 %), обрабатывает 41 млрд 345 млн запросов в месяц (доля рынка 62,4 %), индексирует более 8 миллиардов веб-страниц, может находить информацию на 191 языке (с 15 октября 2009).

Поддерживает поиск в документах форматов PDF, RTF, PostScript, Microsoft Word, Microsoft Excel, Microsoft PowerPoint и других.

В качестве учебного проекта двое студентов Стендфорского университета, Лари Пейдж и Сергей Брин, предложили новую поисковую систему, которая на сегодняшний день стала одной из самых влиятельных компания во всемирной сети Интернет - Google.

Своё название система и соответствующая компания получили от математической величины "гугол" (googol), которая равна 10<sup>100</sup>. Создатели посчитали, что такое название лучше всего передаст основную идею их системы - организовать миллиарды байтов информации, содержащейся в Интернете.

Поисковая система Google представляет собой мощный механизм. Без таких поисковых систем найти информацию в глобальной сети Интернет было бы практически невозможно. Подобно всем поисковым серверам Google использует специальный поисковый алгоритм для получения результатов поиска. Некоторые основные характеристики алгоритма компания не скрывает, особенность же алгоритма является строгой тайной компании. Именно благодаря этому Google сохраняет свое лидерство в сети Интернет и не позволяет никому взломать ее систему.

Как и большинство поисковых серверов, Google использует программный поисковый механизм, называемый «паук» или «ползунок», для автоматического выбора всех документов, на которые есть ссылки в первом выбранном документе. Ключевые слова вписываются в специальную строку ввода, после чего начинается поиск.

Особенность Google заключается в том, как и по какому критерию эта поисковая система классифицирует результаты поиска на своей странице. Используемый алгоритм PageRank сортирует все веб-страницы по критерию смыслового соответствия.

Алгоритм PageRank зависит от нескольких факторов:

От частоты повторения и местоположения ключевых слов на веб-странице – если искомое слово или фраза только один раз встречается на сайте, то страница получает низкий бал.

От времени существования страницы – каждый день в Интернете появляются новые сайты, но не многие задерживаются надолго. Поэтому преимущество отдается уже зарекомендовавшим себя сайтам, существующим долгое время.

От количества веб-страниц, связанных с «главной страницей» - Google «смотрит» на количество веб-страниц, относящихся к определенному сайту и определяет ее рейтинг среди всех остальных.

Систему Google практически невозможно обхитрить, потому что все ссылки на веб-страницы она воспринимает как «голоса». Самый лучший способ обеспечить своему сайту первые места, это напичкать его самой разнообразной информацией, которая будет привлекать больше людей. Чем больше ссылок на веб-странице, тем выше будет оценка поискового агента PageRank.

Для некоторых результатов поиска Google ранее предоставлял повторное поле для поиска, которое позволяет пользователю найти то, что он ищет внутри конкретного веб-сайта. С 2011 года эта опция недоступна, что весьма отрицательно сказалось на удобстве, популярности, и востребованности Google.

«Википоиск» – поисковая технология, позволяющая пользователю настраивать результаты выдачи по поисковым запросам. Пользователь может удалять результаты из списка и поднимать вверх списка. Технология была запущена компанией «Google» весной 2009 года и проработала до осени.

Google предоставляет не только возможность эффективного поиска, но и ряд бесплатных сервисов, большинство из которых не требует установки дополнительного ПО на компьютер пользователя.

Наиболее популярными веб-приложениями от Google являются «Gmail» и «Gtalk». Оба проекта прекрасно дополняют друг друга, но могут работать и по отдельности.

«Gmail» представляет собой почтовый сервис с автоматической фильтрацией спама, мобильным доступом и большим объемом почтового ящика (на апрель 2009 г. он составляет более 7 Гб, и это значение неуклонно растет).

«Gtalk» – сервис, позволяющий обмениваться голосовыми и текстовыми сообщениями, как в окне браузера, так и с помощью специального ПО.

Для владельцев популярных сайтов Google предлагает несколько программ контекстной рекламы. Сервисы «AdSense» и «AdWords» позволяют заработать на посещаемости своих Интернет-страниц и привлечь новых посетителей.

Обычным пользователям «Google» предлагает доступ к справочной информации, новостям, обмену фотографиями, и многим другим ресурсам.

## **Yahoo!**

Yahoo! (произносится как яху!) — американская компания, владеющая второй по популярности (6.35 % ) в мире поисковой системой и предоставляющая ряд сервисов, объединённых интернет-порталом «Yahoo! Directory»; портал включает в себя популярный сервис электронной почты «Yahoo! Mail», один из старейших и наиболее популярных в Интернете. Не так давно была запущена новая версия почтового интерфейса, основанная на AJAX (русскоязычный обзор нового интерфейса).

В январе 1994 года аспиранты Стэнфордского университета Дэвид Файло и Джерри Янг создали веб-сайт, который назывался «Путеводитель Джерри по Всемирной Паутине». «Путеводитель» представлял собой каталог других сайтов.

В апреле 1994 года сайт был переименован в Yahoo!. Существует две версии происхождения названия. Согласно первой, слово было взято из книги Джонатана Свифта «Путешествия Гулливера» (в русском переводе Йеху, еху), где обозначает расу грубых и тупых человекообразных существ. Именно на этой версии настаивают основатели компании. Согласно второй, Yahoo! — акроним,

образованный от фразы «Еще один иерархический неотесанный (неофициальный) прорицатель» (англ. Yet Another Hierarchical Officious Oracle). Но есть и третья версия происхождения названия. Вполне возможно, Yahoo произошло от японского Yahhoo, что в переводе означает «Привет» (неформальный вариант). Однако к тому времени Yahoo уже был зарегистрированной торговой маркой соуса для барбекю, поэтому к названию был добавлен восклицательный знак. Янг и Файло быстро оценили коммерческий потенциал проекта и 2 марта 1995 года основали корпорацию Yahoo!

Согласно статистике Alexa Internet, на сегодняшний день Yahoo! — четвертый по посещаемости веб-сайт в сети Интернет, и примерно 28% посещений состоят из просмотра только одной страницы.

## **Bing**

Bing — поисковая система, разработанная международной корпорацией Microsoft. Доступна по адресу <http://www.bing.com/>. Ранее имела следующие наименования и адреса:

MSN Search (<http://search.msn.com/>) — с момента появления и до 11 сентября 2006;

Windows Live Search (<http://search.live.com/>) — до 21 марта 2007;

Live Search (<http://www.live.com/>) — до 1 июня 2009.

В настоящее время Bing занимает третье место по счёту самых используемых поисковых систем, в отличие от которых, обладает рядом эксклюзивных возможностей, таких как просмотр результатов поиска на одной странице (вместо пролистывания многочисленных страниц результатов поиска), а также динамическое корректирование объёма информации отображаемой для каждого результата поиска (например, только название, краткая или большая сводка).

В американской версии Bing присутствуют некоторые интересные нововведения, касающиеся интернет-поиска, например:

- ежедневно изменяющиеся темы оформления стартовой страницы с информационными блоками;
- вывод уточняющих вариантов поисковых запросов по отдельным категориям;
- видео с автоматически запускающимся предварительным просмотром;
- дополнительные данные по каждому результату поиска;

- встроенный сервис для поиска маршрутов (другие специальные поисковые сервисы появятся в скором времени);
- функции, повышающие удобство в использовании при поиске информации, изображений и видео.

При том, что Bing неплохо ищет по англоязычным ресурсам, для российских пользователей он фактически бесполезен. «Яндекс» и Google дают гораздо более релевантные результаты.

## **Яндекс**

Поисковая система "Яндекс" является восьмой среди крупнейших поисковых сайтов мира по количеству обработанных поисковых запросов (1,892 млрд., статистика за декабрь 2009 г.) и первым крупнейшим неанглоязычным поисковым сервером.

Ежедневная посещаемость этой поисковой системы – более 9 миллионов человек в сутки (январь 2010). В поисковой системе представлено более 5,4 миллиона сайтов с количеством страниц более 2,5 миллиарда, что составляет более 62 терабайта информации.

Поисковая система Yandex.ru была официально анонсирована 23 сентября 1997 года, и первое время развивалась в рамках компании CompTek International. Как отдельная компания "Яндекс" образовался в 2000 году.

"Яндекс" постоянно совершенствует свои поисковые алгоритмы, что позволяет ему соответствовать современным критериям поиска и становиться на уровень Google хотя бы в России. А пока так и есть, если судить по уровню освоения обеих поисковых систем аудиторией СНГ.

"Яндекс", как и Google, работает на кластерной системе организации компьютерных вычислительных сетей. Каждый кластер отвечает за определённый сегмент сохранённой информации.

Сканирующие роботы поисковой системы бывают 2 видов:

1) основной сканирующий робот

2) быстрый робот, который регулярно сканирует сайты, на которых информация обновляется очень часто, и вносит информацию с этих сайтов в поисковую систему, что обеспечивает быстрое обновление индекса поисковой системы.

Обновления (апдейты) поисковой системы бывают 2 видов:

1) Обновления поисковой базы. Происходят несколько раз в месяц. В поиске начинают выдаваться обновлённые страницы сайтов, собранные основным сканирующим роботом.

2) Обновления программной части (движка) поисковой системы. Сутью данных обновлений является изменение алгоритмов ранжирования документов в поисковой системе. Обычно подобные обновления анонсируются и им даются определённые названия.

"Яндекс" учитывает морфологию русского языка (мощная система определения словоформ), имеет возможность сузить запрос до максимально точного при помощи специальных поисковых формул и геотаргетинга. Также имеет качественно разработанный алгоритм оценки релевантности документов (точность отношения текста документа к поисковому запросу), обладает высокой скоростью реакции на запросы и практически полным отсутствием перегрузок серверов.

Совсем недавно (с 2009 года) с вводом нового алгоритма "Снежинск" поисковая система "Яндекс" научилась определять региональную принадлежность сайтов и формировать результаты поиска по географическому расположению пользователя. Сегодня это самый точный географический поисковик Рунета.

Вместе с тем, интернет-портал «Яндекс» представляет собой не только мощную поисковую систему. Под этим брендом собрано множество самых разнообразных сервисов.

Так, с помощью «Яндекса» можно:

- узнать свежие новости («Новости»),
- пообщаться с друзьями и коллегами («Блоги», «Почта»),
- заработать («Мой Круг», «Директ», «Рекламная сеть»),
- продать или приобрести различные товары («Маркет», «Авто»),
- получить море полезной информации (пробки, погода, котировки валют, телепрограмма, афиша).

Не менее популярна у русскоязычных пользователей и платежная система «Яндекс Деньги», которая позволяет осуществлять электронные платежи с помощью веб-интерфейса или Интернет-кошелька.

**Рамблер**

Rambler – поисковая система, разработанная в 1996 году. До недавнего времени Rambler был первым по популярности поисковиком Рунета. И хотя сейчас виртуальная пальма первенства «перешла» к Яндекс и Google, Rambler прочно обосновался на ведущих позициях. По оценкам SpyLog на Rambler приходится 20-25% от всех поисковых запросов Рунета.

Название поисковика как нельзя лучше отражает суть деятельности компании («rambler» - по-англ. «бродяга»). Rambler позволяет искать документы на английском, русском, украинском и многих других языках, а при формировании списка к выдаче учитываются не только ключевые запросы, но и их разнообразные словоформы. Полученные в результате поиска данные структурируются по степени соответствия запросу (релевантности) и распределяются по сайтам.

Сравнительно недавно (в начале 2009 г.) разработчики Rambler'a начали внедрение вертикального поиска, в основе которого лежит технология XAG (eXtended AGgregator). Благодаря этой системе, запрашиваемая информация сортируется по темам, а работать с поисковой машиной становится несколько проще.

Использование вертикалей поисковиком имеет и еще одно преимущество — при недостатке информации в одном документе, она может дополняться данными из другого. Что очень удобно для «очистки» данных от повторов, спама и так далее. Например, при поиске вакансий по телефонному номеру той или иной компании система определяет ее название. Таким образом, поисковая система Rambler может отслеживать на соответствующих сайтах дублирующиеся объявления, сомнительные вакансии и т.д.

Особенностью поисковой системы является то, что Rambler обслуживает только сайты, находящиеся в следующих доменах первого уровня: Российская Федерация: .ru, .su; Украина: .ua; Белоруссия: .by; Казахстан: .kz; Киргизия: .kg; Узбекистан: .uz; Грузия: .ge.

Вместе с тем, Rambler — это не только поисковая машина. Под эгидой Rambler организовано множество Интернет-проектов, с помощью которых можно узнать новости, посетить наиболее популярные сайты, посмотреть видео, послушать музыку, получить полезную информацию, а также развлечься и познакомиться с интересными людьми. Помимо таких традиционных и знакомых всем функций, как Топ 100, почта, новости и игры, Rambler предлагает и другие сервисы.

**Нигма**

Нигма.РФ – российская интеллектуальная метапоисковая система, первая кластеризующая поисковая система в Рунете. Проект создан при поддержке факультетов ВМиК и психологии МГУ, а также Стэнфордского университета.

Название — Nigma (один из родов пауков семейства Dictynidae, en:Nigma) было выбрано по ассоциации со Всемирной паутиной.

Nigma осуществляет поиск как по своему индексу, так и по индексам Google, Yahoo, Bing, Яндекс, Rambler, AltaVista, Aport. По состоянию на 28 февраля 2009 в суммарном индексе всех этих поисковых систем находится более [7 160 000 000](#) русскоязычных документов.

На основе введённого пользовательского запроса Нигма формирует список документов, разделённых на несколько множеств (кластеров). Пользователь может уточнить в каком множестве продолжить поиск, тем самым улучшив релевантность результатов поиска. Пользователь также может исключить ненужные ему множества сайтов, например, документы, пришедшие с интернет-магазинов (для них формируется специальный кластер). Список кластеров выводится слева от списка результатов поиска. Для каждого кластера указывается образующая его фраза и количество документов в кластере. Пользователь может управлять кластерами при помощи специальных ссылок под списком кластеров.

Nigma поддерживает русскую морфологию.

Используется морфологический модуль для русского языка собственной разработки.

Нигма также может выводить ход решения многих математических задач.

Система позволяет производить поиск по более чем 12000 неорганических реакций.

В ответ на запрос пользователя увидят три самые свежие новости прямо на странице результатов поиска. Разработчики проиндексировали и обработали данные более 3500 rss-лент СМИ и популярных блогов. На сайте также есть форма, через которую пользователи сами могут добавлять новостные ресурсы в индекс поисковой системы. База данных новостей обновляется каждые 5 минут.

Поддерживается орфография на базе словаря, составленного из реальной русской орфографии, которую авторы веб-сайтов используют на своих страницах.

Система предлагает варианты на основе предыдущих запросов пользователей Нигма.РФ.

В улучшенной версии поисковик отвечает на вопросы пользователя до того, как он полностью введёт свой вопрос в строку поиска. Система предлагает варианты заполнения, учитывая популярности предыдущих запросов пользователей. Справа от варианта выдаётся ответ на вопрос.

При выборе нужного варианта в поисковой подсказке пользователь может увидеть определения понятий и слов, которые хочет найти. База данных содержит 340928 определения. Информация берется из Википедии. Предполагается подключение и других источников информации.

## **Mail.ru и GoGo.ru**

Mail.ru — крупный коммуникационный портал российского Интернета. Его ежемесячная аудитория превышает 50 миллионов уникальных посетителей. Ресурс принадлежит инвестиционной группе Mail.ru Group.

В 1998 году программисты, работающие в петербургском офисе американской софтверной компании DataArt, создали новое программное обеспечение для почтового веб-сервера, которое в дальнейшем предполагалось продавать западным компаниям. Чтобы протестировать сервис, его временно выложили в открытый доступ для российских пользователей, а сервис вдруг стал стремительно набирать популярность.

GoGo.ru – относительно новый и вполне автономный проект от создателей Mail.ru. Несмотря на то, что соответствующий домен был зарегистрирован компанией еще в 2000-м году, работы над созданием поисковика начались только в 2006. Руководство над новым «детищем» Mail.ru доверила организатору поисковой системы Апорт – Михаилу Костину.

Поисковик был запущен в июне 2007 года, и первое время имел небольшой охват и выдавал довольно странные результаты. Однако работа над поисковой системой продолжалась, и к февралю 2008 г. GoGo.ru осуществляла поиск по более чем 2,5 млрд. документов, 140 млн. изображений, 2 млн. видео-файлов, и т.д.

Успехи новой поисковой системы были отмечены высшей наградой на выставке SofTool-2007. Именно там, в рамках конкурса «Продукт года», в номинации «Интернет-технологии», GoGo.ru получила первое место.

От остальных поисковых систем GoGo.ru отличается возможностью поиска по русскоязычным видеороликам, WAP-сайтам, анализа базы данных Ответы@Mail.ru, а также набор функций для веб-мастеров и владельцев собственных Интернет-проектов.

В настоящее время, основными видео-хостерами, по материалам которых осуществляет поиск GoGo.ru, являются такие ресурсы, как Video.Mail.ru, RuTube.ru, LiveInternet.ru, Teledu.ru, Smotri.com, Myvi.ru, Video.i.ua и некоторые другие. Кроме того, GoGo.ru может искать видео-ролики по размещенным на YouTube или Metacafe файлам с описанием на русском языке.

Следует отметить, что при поиске роликов в выдаче оказываются не только видео, отвечающее запросу, но и изображения. При этом поисковый бот GoGo.ru имеет встроенный фильтр, позволяющий отделять фотографии от всех остальных картинок.

Высокая эффективность графического поиска в GoGo.ru обеспечивается наличием XML-синдикации с ведущими фотохостерами Рунета.

Заслуживает внимания и обработка поисковой системой разнообразных синонимов. Данной проблеме разработчики GoGo.ru уделяли особое внимание, и в настоящее время проект имеет обширный словарь словоформ и синонимии, пополняющийся в полуавтоматическом режиме.

На данный момент, практически любой веб-мастер или владелец популярного Интернет-ресурса может поставить на своих страницах специальную поисковую форму от GoGo.ru, с помощью которой посетители смогут искать информацию, изображения или видео на сайте и в Интернете.

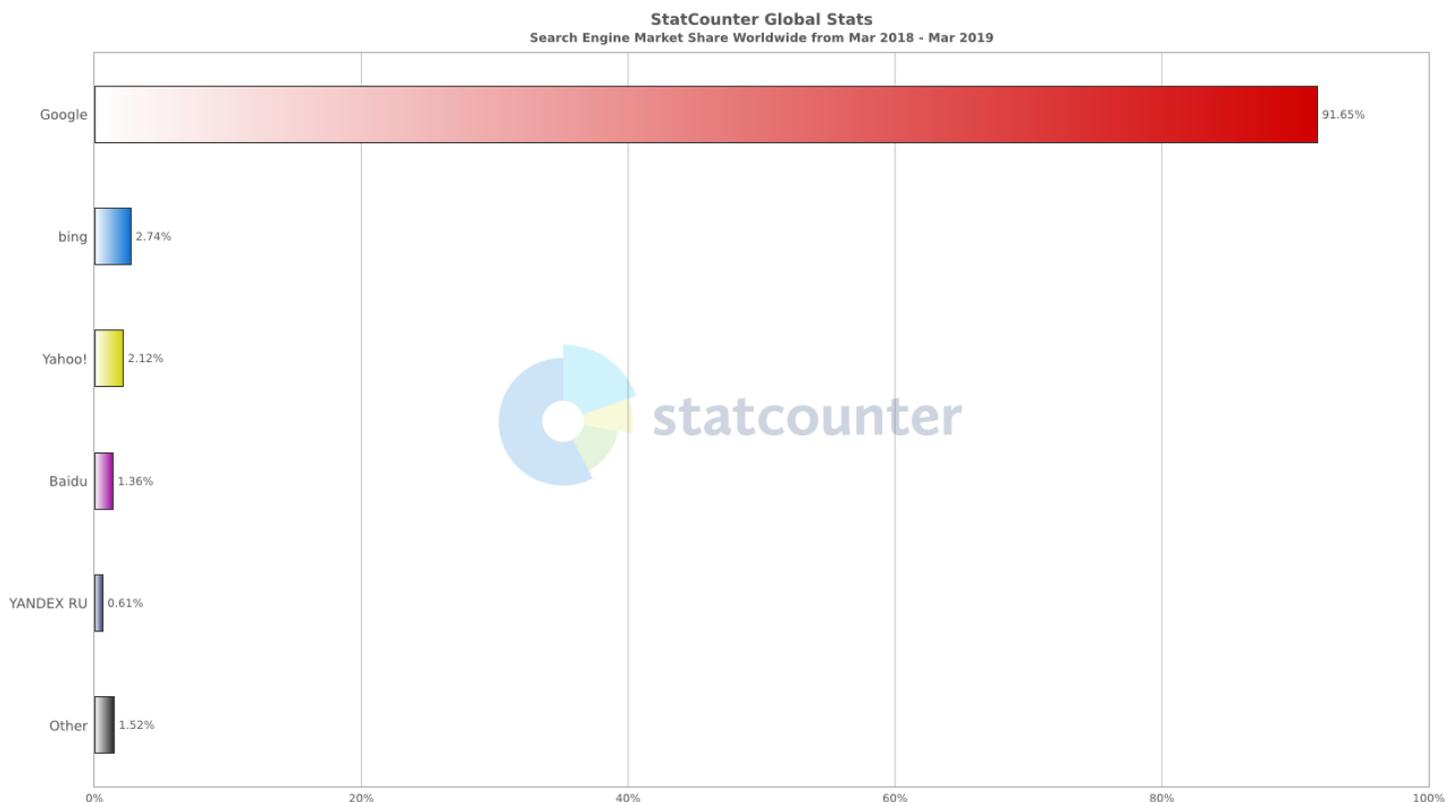
При этом совершенно бесплатно можно обеспечить удобный поиск по своим страницам или организовать мета-поисковую систему. Главное требование GoGo.ru – не более 10 000 поисковых запросов с одного сайта в сутки.

## **3.2 Анализ популярности поисковых систем в мире**

По последним данным StatCounter, самый популярный поисковик в мире — Google.

После него идет Bing от компании Microsoft, затем американский Yahoo, китайский Baidu и российский Яндекс.

В топ-10 входят поисковики Aol, Ask.com, Mail.ru, Yahoo, DuckDuckGo и другие.

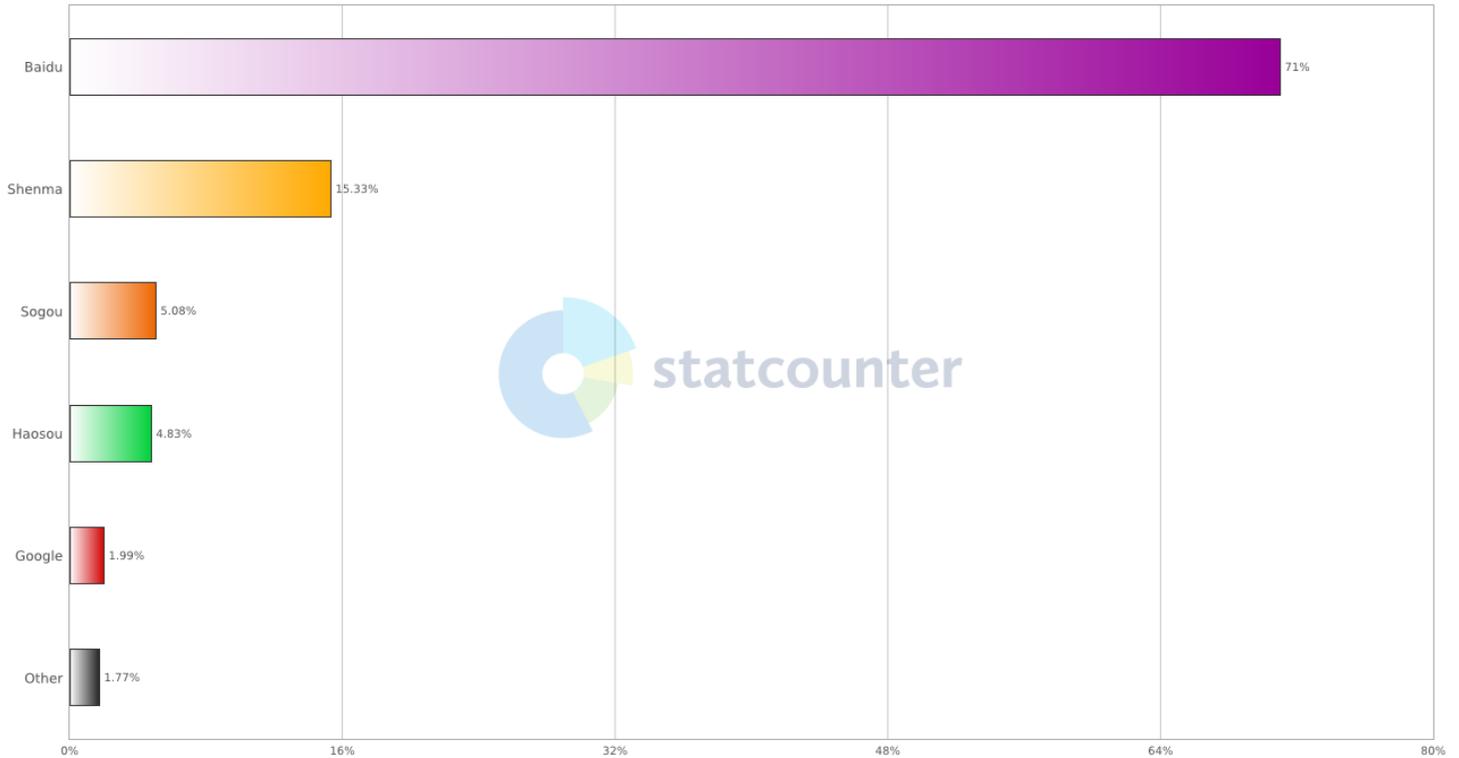


**Рисунок 1. Популярный поисковик в мире.**

Однако общемировая статистика не всегда совпадает с региональными рейтингами и не отражает популярность нескольких местных поисковиков. Чтобы в этом убедиться, проанализируем крупнейшие регионы по отдельности.

## **Китай**

Самый популярный поисковик в Китае — Baidu, затем идет Shenma, Sogou, Haosou и только потом Google.

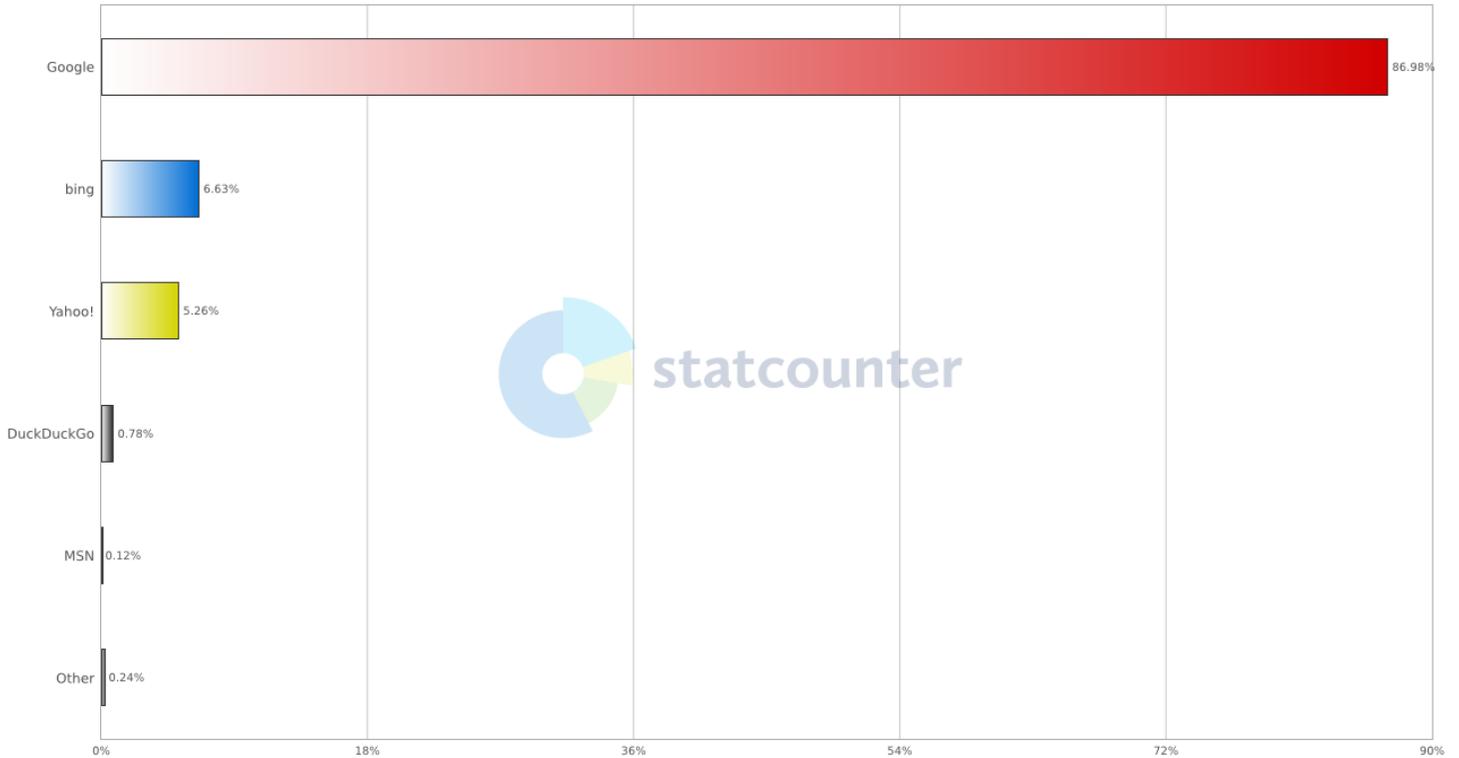


## Рисунок 2. Популярный поисковик в Китае.

### Соединенные Штаты Америки

Самый популярный поисковик в США — Google. Компания была создана в Америке, где расположен ее главный офис. По этой же причине в Штатах популярны поисковики Bing, Yahoo и MSN.

StatCounter Global Stats  
Search Engine Market Share United States Of America from Mar 2018 - Mar 2019



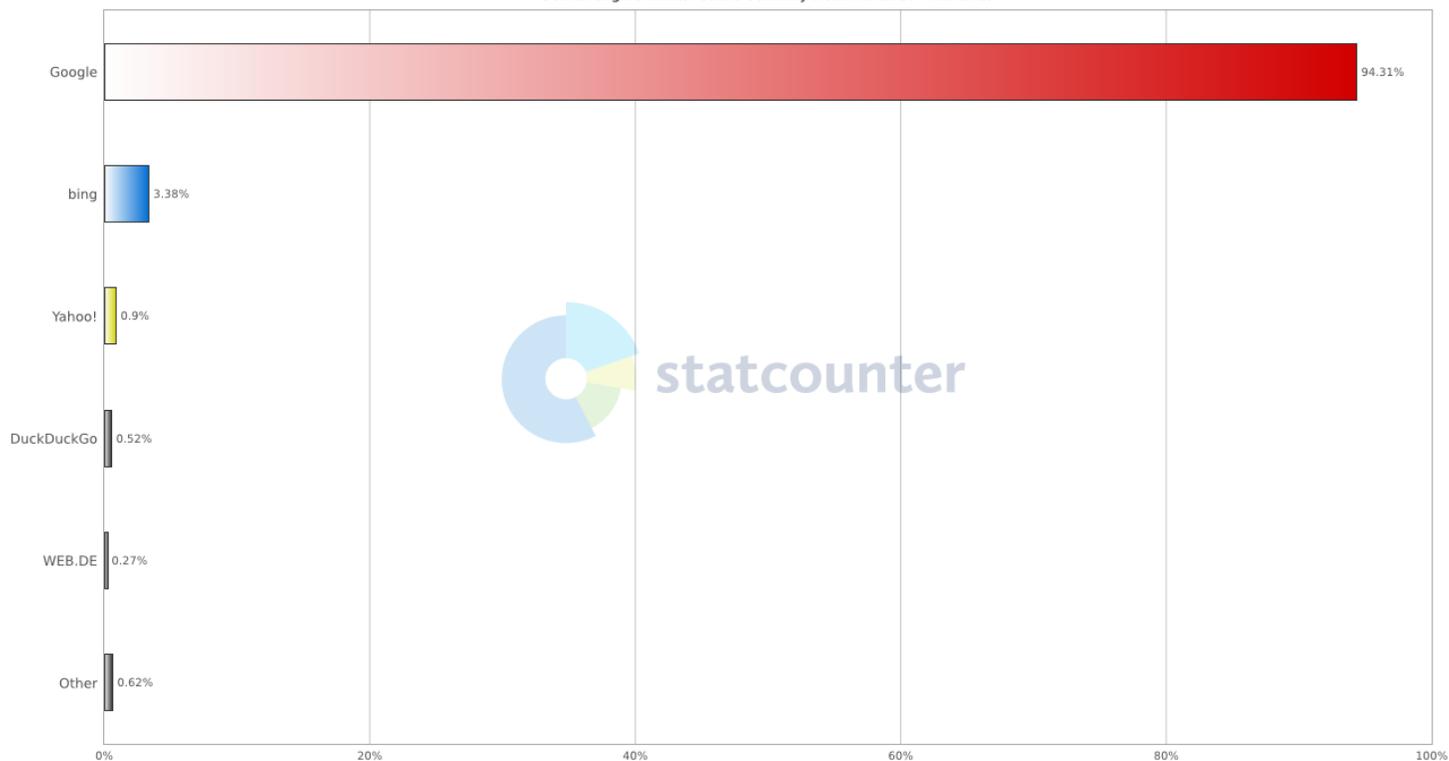
### Рисунок 3. Популярный поисковик в США

#### Европа и СНГ

Самый популярный поисковик в Германии — Google, на второй позиции находится Bing, а на третьей — Yahoo.

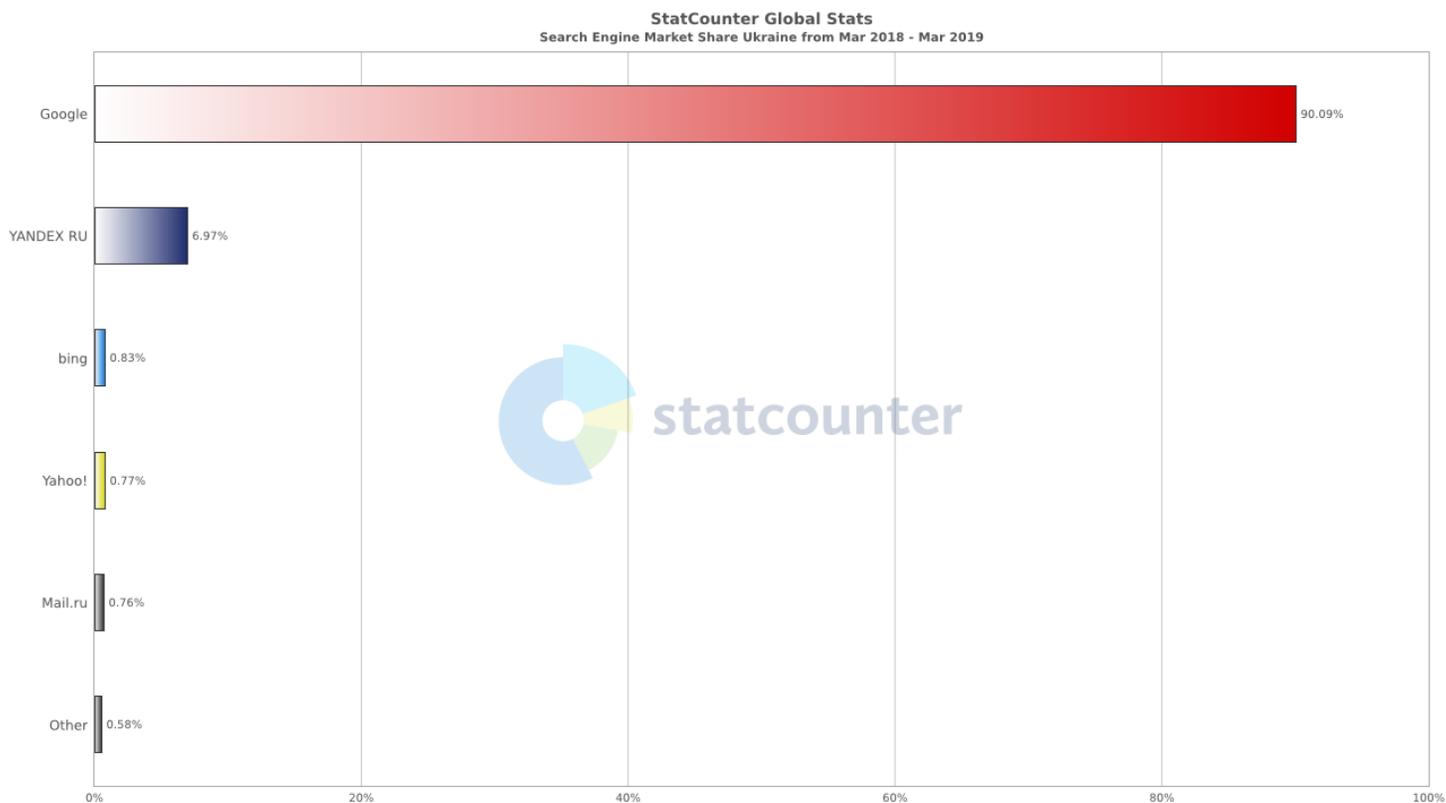
#### StatCounter Global Stats

Search Engine Market Share Germany from Mar 2018 - Mar 2019



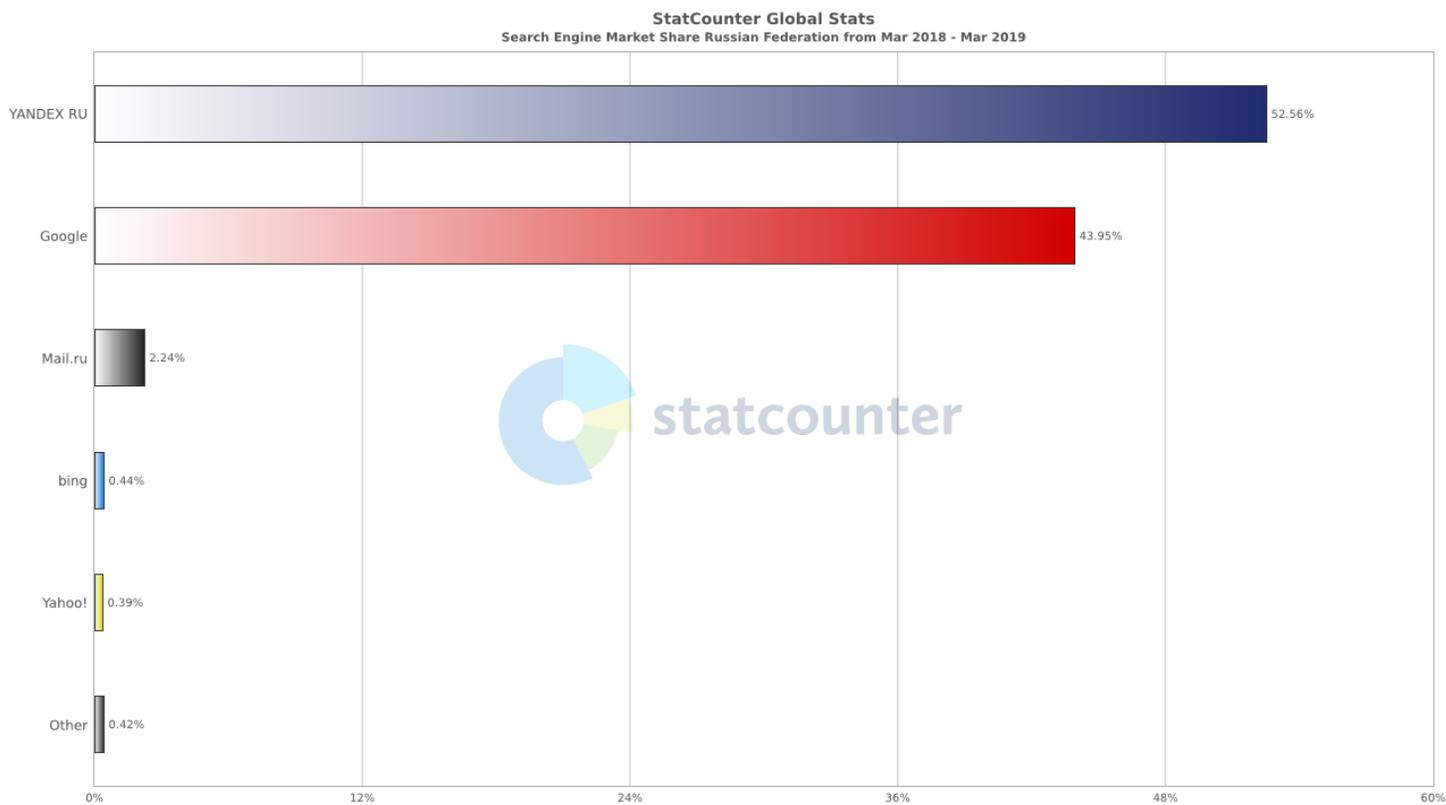
#### Рисунок 4. Популярный поисковик в Германии.

Самый популярный поисковик в Украине — снова Google. Несмотря на блокировку, вторая позиция принадлежит Яндекс, а третья — поисковой системе Bing.



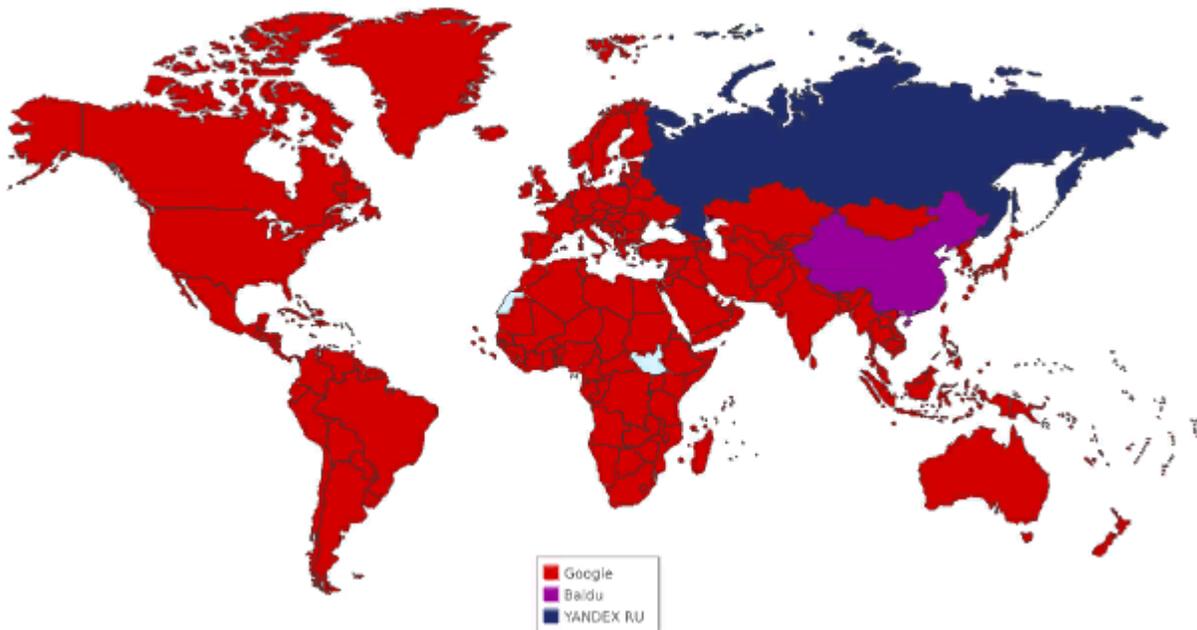
### Рисунок 5. Популярный поисковик в Украине.

Самый популярный поисковик в России — Яндекс, который был создан в РФ. Второе место занимает Google, а третье принадлежит российской поисковой системе Mail.ru.



**Рисунок 6. Популярный поисковик в России.**

Разницу в популярности поисковых систем можно проследить на карте мира.



**Рисунок 7. Карта популярности поисковых систем в мире.**

### **3.3 Сравнение популярных поисковых систем**

Сравним три самые популярные поисковые системы в России: Google, Yandex, Rambler.

**Таблица 1. Анализ поисковых систем**

**Название Достоинства**

**Недостатки**

Специализированные роботы сканируют и оценивают содержание страниц для поискового индекса;

Главный плюс – очень короткое время отклика системы на запрос;

Поиск информации по запросу в PDF-файлах;

## **Google**

Google организует поиск на десяти различных языках – это максимальный результат;

Поисковик показывает наиболее высокие результаты по чистоте, релевантности и точности поиска;

Поисковая система имеет собственную почту, облачное хранилище, браузер, карты и прочие сервисы.

Постоянное усовершенствование систем и добавление различных фильтров в значительной мере усложняет работу и затрудняет продвижение сайтов;

Обширная информационная база препятствует точному поиску, засоряя «эфир» огромным количеством несоответствующих запросу ссылок.

Огромное количество сервисов системы, самые популярные из которых – почта, маркет, карты, диск, браузер и другие, значительно превосходящие основных конкурентов;

Создание виджетов (блоков новостей) на главной странице Yandex;

Высокая информативность Яндекса;

Удобное и компактное расположение результатов поиска;

База поисковой системы – индекс – которую формирует поисковый робот. Он анализирует и собирает информацию в Сети с определенной периодичностью. При выдаче информации пользователю робот учитывает язык пользователя, его местоположение, прошлую активность и историю запросов. Каждый результат создается индивидуально! Такого подхода не встретишь ни у одного другого поисковика;

Необходимо добавить, что Yandex обладает еще одной очень ценной характеристикой: он исправляет орфографические ошибки и предлагает возможные варианты. Это крайне удобно, если вы ищете информацию по термину, в написании которого вы не уверены. Google в этом отношении не надежен и запросто может предложить несколько десятков страниц результатов на неверный запрос. Так что плюсы пользования системой Yandex очевидны;

Ну, и по мелочи: удобный и красивый интерфейс, надежный хостинг,

Постоянно появляющаяся САРТСНА отнимает время и раздражает пользователей;

Для тех, кто сайты «делает» и продвигает очень сложно раскрутить новый ресурс именно на Яндексе, система крайне подозрительно воспринимает «юные» сайты;

Сбои в поиске порою приводят к тому, что в топе оказываются откровенно никчемные сайты, а популярные еще вчера площадки стремительно скатываются в рейтинге. Впрочем, так же неожиданно они возвращаются на ранее занимаемые места.

**Yandex**

Большой опыт работы;

**Rambler** Новая технология анализирует поведение пользователя и определяет его пол и возраст.

Поисковик морально устарел, не выдержав конкуренции со стороны отечественных и западных поисковых систем;

Тестирование Рамблера указывает на низкое индексирование динамических сайтов.

Итак, по результатам сравнения все поисковые системы показали хороший результат, а опросы пользователей показывают, что предпочтения у всех разные.

Можно составлять сколько угодно сводных таблиц, сопоставляя результаты работы различных поисковых систем, но, так или иначе каждый человек сам делает свой выбор.

Привлекает скорость – Google к вашим услугам, нужна точность и широкая вариативность запросов – к вам на помощь придет Yandex, а Rambler для тех, кто хочет отойти от шаблона.

## **Заключение**

В результате выполненной работы были решены следующие задачи:

- выбраны поисковые системы для анализа;
- произведён анализ и сравнение поисковых систем;
- в результате сравнения, были выбраны наиболее подходящие поисковые системы.

Была достигнута цель данной работы – найдены лучшие поисковые системы.

## **Список использованной литературы**

1. Афанасьева, Е. В. Знакомства в Интернете. Просто как дважды два / Е.В. Афанасьева. - М.: Эксмо, 2007. - 128 с.

2. Блам, Эндрю Сеть. Как устроен и как работает Интернет / Эндрю Блам. - М.: АСТ, Neoclassic, 2014. - 320 с.
3. Васильева, В.М. Шаг за шагом в мир глобальной науки. Руководство по использованию Web of Knowledge, Journal Citation Report, EndNote Web, Researcher ID, Google Scholar, Истина / В.М. Васильева. - М.: Московский государственный университет имени М.В. Ломоносова (МГУ), 2012. - **457** с.
4. Габбасов, Ю. Internet 2000. Эффективные технологии работы в Сети / Ю. Габбасов. - М.: БХВ-Петербург, **2005**. - 448 с.
5. Гультяев, А. К. Самое главное о... Поиск в Интернете / А.К. Гультяев. - М.: Книга по Требованию, 2012. - 144 с.
6. Кенин, Александр Окно в мир Интернета / Александр Кенин. - М.: У-Фактория, 2003. - 360 с.
7. Кутовенко, Алексей Профессиональный поиск в Интернете / Алексей Кутовенко. - М.: Книга по Требованию, 2011. - 254 с.
8. Паризер, Эли За стеной фильтров. Что Интернет скрывает от вас / Эли Паризер. - М.: Альпина Бизнес Букс, 2012. - 304 с.
9. Райтман, М. Как найти и скачать в Интернете любые файлы / М. Райтман. - М.: БХВ-Петербург, 2015. - **319** с.
10. Севостьянов, Иван Олегович Поисковая оптимизация. Практическое руководство по продвижению сайта в Интернете / Севостьянов Иван Олегович. - М.: Питер, 2017. - **383** с.
11. Топорков, С. С. Альтернативные браузеры / С.С. Топорков. - М.: ДМК Пресс, 2006. - 320 с.
12. Халявин, В. Секретная книга internet-пирата / В. Халявин. - М.: Мартин, 2015. - **114** с.
13. Щербаков, Илья Интернет как средство получения денег, бесплатных товаров и услуг / Илья Щербаков. - М.: БХВ-Петербург, 2002. - 368 с.
14. Яковлев, А. А. Раскрутка сайтов. Основы, секреты, трюки / А.А. Яковлев, В.Г. Ткачев. - М.: БХВ-Петербург, 2015. - 352 с.
15. [www.seor.ru](http://www.seor.ru) - Search engine optimization project, рейтинг основных поисков
16. [www.gpntb.ru](http://www.gpntb.ru) - Перспективы развития поисковых систем